Options for Short-Term Price Determination in the Brazilian Wholesale Electricity Market: Report Prepared for Câmara de Comercialização de Energia Elétrica (CCEE)

by

Frank A. Wolak
Department of Economics
Stanford University
Stanford, CA 94305-6072
wolak@zia.stanford.edu

July 15, 2008

1. Introduction

This report was commissioned by the Câmara de Comercialização de Energia Elétrica (CCEE), the market operator for the Brazilian electricity supply industry, to study the options for short-term price determination in the Brazilian electricity market. As part of the process of preparing this report, I read background documents on the market structure and rules governing the Brazilian electricity supply industry, analyzed actual market outcome data on the performance of the short-term wholesale electricity market in Brazil, and participated in a conference held in São Paulo on May 20, 2008 where this topic was discussed by stakeholders from all segments of the Brazilian electricity supply industry.

This report describes the results of my analysis of the options for short-term price determination in the Brazilian electricity supply industry. The three major questions considered are: What are the initial conditions necessary for the introduction of bid-based short-term market for the Brazilian electricity supply industry? What should be the transition process from the current cost-based market to the final bid-based market. What is the recommended form for the final bid-based short-term market in Brazil? To provide a framework for considering these questions, the economic theory of the electricity market design process is first introduced. The two fundamental challenges of the market design process are how to obtain: (1) technically and allocatively efficient production and (2) economically efficient pricing of wholesale electricity.

Six major dimensions of the short-term electricity market design process are then introduced. I then discuss how each of these dimensions is dealt with in the current Brazilian short-term wholesale electricity market and how each might be addressed in my recommended future short-term market. The major issue dealt with in this section of the report is the issue of a cost-based versus bid-based short-term wholesale market. In order to understand the potential market efficiency and system reliability benefits of a bid-based market for Brazil, I then present the results of a comparative empirical analysis of the performance the current Brazilian short-term market and the short-term markets in hydroelectric-dominated industries with bid-based markets in Colombia, New Zealand, and Norway. I believe that the results of these market performance comparisons provide evidence that there are significant market efficiency benefits associated with Brazil adopting a bid-based short-term market.

The next section of the report describes the initial conditions necessary to implement a bid-based short-term market in Brazil. These necessary conditions are: (1) coverage of close to 100% of final demand in fixed-price forward contract obligations negotiated far enough in advance of delivery to allow new entrants to compete to supply these contracts, (2) a local market power mitigation mechanism that applies to all market participants, (3) a cap and floor on supply offers into the short-term wholesale market, and (4) a prospective market monitoring process with public release of all data necessary to operate the short-term market. A key recommendation from this section of the report is that a bid-based short-term market should not be implemented in Brazil without these necessary pre-conditions.

The report then presents a recommended bid-based short-term market design and suggests a transition process from the current cost-based market design to this market design that initially involves minimal changes in the current cost-based market. Although I believe that this transition process should take between 12 to 18 months to complete, I do not think that this timetable should be adhered to without regard to events in the short-term market. In particular, further moves towards introducing flexible market mechanisms should not be made without the appropriate safeguards against the exercise of unilateral market power in place and validation that these safeguards are working as intended.

2. Market Design in Wholesale Electricity Markets

The goal of electricity industry restructuring is to reduce the amount of economic inefficiencies that existed in the former vertically-integrated, state-owned monopoly regime. There are two major sources of economic inefficiencies: (1) productive inefficiencies and (2) pricing inefficiencies. As discussed below, it is impossible to eliminate completely both of these sources of economic inefficiencies. Ideally, electricity restructuring finds the optimal combination of explicit regulation and market mechanisms to minimize these two sources of economic inefficiencies. Market design is shorthand for the process of determining this optimal combination of explicit regulation and market mechanisms.

The technology of energy production and delivery imposes a hard constraint on the market design process. Specifically, how many units of each input—capital, labor, input energy, and materials—is required to produce one unit of energy or how many megawatt-hours (MWh) of electricity produced can be delivered to final consumers through a given configuration of the transmission and distribution networks. However, firms in the electricity supply industry have considerable discretion over how they utilize these technological relationships. Because of differences in the incentives their owners and managers face, two firms may use the same production technologies in substantially different ways. For example, if a firm's long-term survival depends on the number of people that work at the firm then it may employ more labor than is necessary to produce a given level of output. If the firm's management is able to curry political favor by making certain capital investments, it may do so despite the fact that these investments are unnecessary to serve the firm's demand. In summary, there are many technologically feasible ways to produce a given level of output with different levels of implied productive inefficiencies.

This logic implies that the market designer faces the following challenge with respect to a given production technology: How can it cause firms to supply their output in a technically and allocatively efficient manner. Technical efficiency implies that the firm is producing the maximum amount of output technically possible for a given quantity of each input. Allocative efficiency implies that the firm is producing that chosen level of output in a least-cost manner given the input prices that it faces. A privately-owned, profit-maximizing firm that is unrestricted in its input choices will produce in a technically and allocatively efficient manner. This result follows from the fact that once the firm's level of output is chosen, its total revenues are independent of its own actions, so that the only way for the firm to maximize profits is to choose its inputs to minimize the total cost of producing this level of output.

Allowing a privately-owned, profit-maximizing firm that owns a significant fraction of the available generation capacity to determine its output level can result in market at prices vastly in excess of economically efficient levels. Consequently, the second market design challenge is how to cause producers to set the lowest possible price consistent with the long-term financial viability of the industry. Economically efficient pricing requires that the market price equal the short-run minimum marginal cost of producing the last unit of output sold. In defining the efficient price, it is important to make the distinction between a firm's minimum cost function and its incurred cost function. The former implies that the firm is producing in a technically and allocatively efficient manner, while the latter does not. The efficient price is the short-run marginal cost of producing the last unit sold assuming that all firms in the industry are producing along their minimum cost function.

The two market design goals of technically and allocatively efficient production and economically efficient pricing often conflict because incentives that cause a firm to produce in an economically efficient manner may cause the firm to set prices far in excess of economically efficient levels. A privately-owned, profit-maximizing firm has a strong incentive to produce in a least-cost manner, but little incentive to set a price that only recovers the marginal cost of the last unit of output sold. Alternatively, regulation that sets an output price that only recovers the firm's actual production costs may cause it to produce a given level of output at significantly higher cost than the technically and allocatively efficient mode production. These higher costs are translated into higher output prices through the regulatory process, which implies significant deviations from economically efficient pricing.

The primary constraint on the market designer's choice of market rules is that all market participants will choose their strategies to maximize their payoffs given these rules. For example, a privately-owned firm has very strong incentives to choose its actions to maximize its profits given the market rules for determining its revenues. This constraint on the market design process is often called the "individual rationality constraint." This constraint implies that firms do not produce in a technically or allocatively efficient manner or price to recover only their production costs unless they have a financial incentive to do so. As noted above, many mechanisms that provide strong incentives for least-cost production also provide strong incentives for prices that yield revenues far in excess of production costs. Conversely, many mechanisms that set prices to recover only the firm's incurred cost of production provide strong incentives for the incurred cost to produce a given level of output to be significantly larger than the minimum cost to produce that same level of output.

This individual rationality constraint implies that the market designer has a choice between two imperfect mechanisms. The first is an imperfectly competitive market where some participants possess unilateral market power. The second is an imperfect regulatory process where firms can extract informational rents because of their superior knowledge, relative to that of the regulator, of the technology of production or form of demand for their output. The market design process involves setting a mechanism for compensating each market participant for their actions so that each participant's unilaterally rational response to the mechanism results in market outcomes that achieve the market designer's goals. In this sense, the market design process has many features of what economists call a principal-agent problem.

There are many real-world examples of the principal-agent problem. These include the lawyer-client, doctor-patient, firm owner-firm manager, and regulator-firm relationships. The principal does not observed everything that the agent observes about the underlying economic environment, but the principal's payoff depends on the actions of the agents. For example, in the lawyer-client relationship, the client (the principal) does not know as much about the law as the lawyer (the agent). Consequently, the client attempts to design a mechanism for compensating the lawyer for his or her actions to achieve the client's desired outcome. The mechanism must recognize the individual rationality constraint on the lawyer's behavior that once the compensation scheme is designed, the lawyer will take actions to maximize his or her payoff function subject to this compensation scheme. For the regulator-firm relationship, the regulator (the principal) does not know as much about the firm's production process or demand as the firm (the agent), but it must design a mechanism for compensating the firm for its actions that comes as close as possible to inducing the firm to produce in an efficient manner and to set efficient prices.

The wholesale market design process is far more complex than the simple principal agent problem described above because it implies a principal-agent relationship at multiple levels in hierarchy with multiple principals and agents at each level. The first level of the principal-agent relationship involves the market designer and the firms participating in the market. Here the market designer is the principal and each firm is an agent. The market designer attempts to design a mechanism for compensating or charging each market participant for their actions to achieve its desired market outcomes. The second level of the principal-agent relationship arises between the owner of each firm and the management of that firm. The owners of the firm attempt to design a mechanism for compensating the management of the firm to achieve behavior that maximizes the owners' payoffs. Explicit solution for the optimal market design is an infeasible computational task given the number of market participants and complexity of the technology of production and delivery of wholesale electricity. Nevertheless, viewing the market design process as a multi-level, multiple-principal and multiple-agent problem emphasizes the importance of the individual rationality constraint that the market designer faces at each level of the principal-agent interaction.

This perspective also clarifies distinct market design challenges of market versus regulatory mechanisms. For bid-based market mechanisms, the fundamental challenge is limiting the exercise of unilateral market power. Short-term wholesale electricity markets are extremely susceptible to the exercise of unilateral market power. Unilateral market power is the ability of a firm to influence the market price through its own actions and to profit from this price change. Both are necessary for a firm to possess unilateral market power. It is also important to emphasize that a firm exercising all available unilateral market power subject to the market rules is equivalent to the firm maximizing its profits. Moreover, a firm maximizing its profits subject to the market rules is equivalent to the firm's management serving its fiduciary responsibility to its shareholders. This logic implies that a firm exercising all available unilateral market power is equivalent to the firm's management serving its fiduciary responsibility to shareholders to earn the highest possible return on their investment in the firm. Consequently, a key issue for regulatory oversight is not whether a firm possesses unilateral market power or exercises it, but if the exercise of this unilateral market power results in market outcomes that cause sufficient harm to consumers to justify explicit regulatory intervention. To prevent the need for explicit regulatory intervention, the market designer should implement mechanisms for compensating firms that cause their unilateral actions to serve their fiduciary responsibility to their shareholders (exercise all available unilateral market power) to result in market outcomes that achieve the market designer's goals.

There are a number of factors that cause short-term bid-based markets to be so susceptible to the exercise of unilateral market power. First, demand must equal supply at every instant of time at every location in the transmission network. Second, all electricity must be delivered through a transmission network with finite capacity between each link in the network. Third, it is very costly to store electricity for consumption at a later date. Fourth, production is subject to severe capacity constraints in the sense that a generation unit with a nameplate capacity of 500 MW can only produce slightly more than 500 MWh in an hour. Finally, how electricity is priced on consumers makes the real-time demand extremely price inelastic. All of these factors enhance the ability of electricity suppliers to exercise unilateral market power in the short-term market. There are numerous examples from wholesale electricity markets around the world of firms exercising enormous amounts of unilateral market power in very short periods of time because of these features of the product sold.

The fundamental market design challenge associated with explicit mechanisms that regulate the price a firm receives for its output is how to cause the firm to produce in a technically and allocatively efficient manner. It is a relatively straightforward accounting exercise for the regulator to set a price that recovers the firm's incurred cost of production. However, the commitment by the regulator to set a price that recovers the firm's incurred cost provides strong incentives for these incurred costs to exceed the minimum cost mode of production. This divergence can be especially large over long time horizons. A regulated firm has very little incentive to undertake investments that reduce or slow the growth of its production costs if the regulatory process commits to translating any future production cost reductions into lower output prices. That is because cost reductions require effort by the firm's managers and employees and the promise to reduce prices in the future when the firm's production costs fall, eliminates any payoff that these agents might receive from production cost reductions. In contrast, market mechanisms provide strong incentives for firms to make cost-reducing investments. Even a firm that possesses no unilateral market power can realize high profits from a cost-reducing investment as long as these cost reductions are not immediately duplicated by all of its competitors. If this is the case, the market price will not be impacted by the firm's investments and therefore it will achieve a profit increase equal to the reduction in its total production costs. As the number of firms that undertake this cost-reducing investment increases, the greater the likelihood the market price will fall because of competition between these lower cost firms to sell their output.

One final implication of viewing the electricity market design process as a multi-level principal-agent problem is that the market design process must be forward-looking and adaptive. Because it is impossible to determine in advance for a given market structure the optimal compensation mechanism for a each market participant, it is essential to recognize that all market designs have flaws, particularly at the start of the re-structuring process. The best performing markets from around the world continually adapt their market rules to changes in the number and size of each market participant, the level and pattern of demand throughout the year, the characteristics of the transmission network, and other features of the market structure.

It is also impossible to eliminate all market design flaws before the start of the market. Because of the complexity of the market design process, it is often only possible to identify many market design flaws by actually running the market. For this reason, it is crucial that a market monitoring process for continuous improvement be implemented at the start of the market to allow all to determine what is working, what is not working, and how to correct any market design flaws. Wolak (2004) outlines the rationale and basic features of such a market monitoring process.

..

3. Six Dimensions of Short-Term Wholesale Market Design

There are six major characteristics of the short-term electricity market design process that impact the incentive firms have to produce in an efficient manner and to set efficient market prices. The first concerns the degree of temporal granularity in pricing. The second characteristic is the degree of spatial aggregation in market prices. The third characteristic is the number of settlement periods--the number of firm financial forward market commitments that market participants make to buy and sell energy and ancillary services before real-time system operation. The fourth market design dimension is ex ante versus ex post pricing. The fifth dimension is the degree of integration between system and market operation. The final short-term market design dimension is a cost-based versus bid-based dispatch and pricing mechanism. The remainder of this section describes each of these dimensions of the market design process.

3.1. Temporal Granularity of Pricing

As noted earlier, economic efficiency implies that the market price should equal the minimum-cost, marginal cost of the highest cost resource necessary to meet demand at all times. Fixing the short-term price of wholesale electricity for a long period of time implies charging higher than the efficient price in some periods and lower than the efficient price in others. This has several negative consequences. First, it overpays suppliers for less valuable electricity. Second, it underpays suppliers for more valuable electricity. Third, it provides incentives for over-consumption during under-priced periods and over-consumption in over-priced periods. One rationale for fixing short-term prices for an extended period of time is that it is also costly to adjust prices on a second-by-second basis. A far more extensive metering infrastructure is required to record the consumption and production of all market participants at this degree of time resolution. Consequently, the market designer faces a trade-off in increasing the temporal granularity of short-term prices between introducing the pricing inefficiencies described above and the up-front costs of implementing greater temporal granularity.

Most wholesale markets set day-ahead prices on an hourly or half-hourly basis. Real-time prices are often set at an even finer level of temporal granularity. This is necessitated by the fact that the system operator often needs to move generation units within an hour or half-hour period to maintain real-time system balance in spite of the existence of generation units on automatic generation control (AGC). In order to issue a real-time dispatch instruction to a generation unit in a bid-based short-term market, the system operator must increase or decrease the market price along that supplier's offer curve into the wholesale market.

An alternative approach for maintaining real-time system balance is to require firms to follow the dispatch instructions of the system operator under the threat of significant penalties. However, this approach raises the question of how to set the optimal penalty for failing to follow the system operator's dispatch instructions. Setting this penalty too high can cause market participants to exhibit excessive caution in responding to the system operator's instructions. For example, the supplier may take actions that ultimately endanger system reliability in order to avoid incurring a significant penalty for failing to respond to the system operator's dispatch instructions. Setting these penalties too low can result in suppliers failing to respond to dispatch instructions, which can also endanger real-time system balance. The advantage of setting prices with a high degree of temporal granularity is that the system operator always has the option to set the price at the level necessary to achieve system balance during the time intervals that prices are set. For example, if the system operator is required to maintain system balance every ten minutes (with automatic generation control maintaining the second-to-second balance within that time period), setting prices every five minutes ensures that the system operator can meet this criterion.

Consequently, a high degree of temporal granularity in pricing in a bid-based wholesale market allows the system operator the greatest flexibility in moving generation units to maintain real-time system balance. Applying this logic to the case of Brazil suggests that the market efficiency costs of infrequent price changes are likely to be smaller when the vast majority of electricity is produced from hydroelectric units which share very close to or the same opportunity cost of water if transmission constraints across regions of the country are rarely binding. Under these conditions, the opportunity cost of moving any generation unit relative to its scheduled dispatch level is not likely to differ significantly across hours of the day or across generation units. However, as the share of fossil fuel generation units in the Brazilian system increases, there is likely to be significant differences across generation units in the variable cost of moving these units, and which of these units must be moved is likely to differ across hours of the day. This logic argues in favor of greater temporal granularity in pricing as the share and diversity of fossil fuel generation units in the Brazilian capacity mix increases.

3.2. Spatial Granularity of Pricing

Economically efficient pricing also requires that the market price equal the minimum-cost marginal cost of the highest cost resource necessary to meet demand at each location in the transmission network. This logic implies that when there is congestion along a transmission path the market price must increase on the constrained side of the transmission path and the market price must fall on unconstrained side of the transmission path. This will reward generation resources located close to load centers with higher average prices than resources located far from the load centers. Greater spatial granularity in pricing also provides incentives for loads to reduce their consumption in the higher priced regions.

Similar to the case of greater temporal granularity in pricing, greater spatial granularity may require a more extensive metering infrastructure and greater market participant sophistication, both of which impose costs on these entities. Consequently, if the transmission network within a large geographic region or zone with multiple locations where generation units can inject and loads can withdraw electricity has sufficient transfer capacity between these locations to allow all generation units to be equally effective at meeting load at all locations in this zone, then the cost of setting a single price for all nodes in this geographic area is likely to be small. This logic implies that a successful zonal market design requires a commitment by the transmission network owner to have sufficient capacity within each zone to make all generation units equally effective at meeting demand at all locations in this zone. This implies that more transmission network investments are necessary within each zone than would be the case if the system operator chose to employ a nodal-pricing model, which allows different prices to be set at all locations in the transmission network.

Applying this logic to the case of the Brazilian electricity market implies that greater spatial granularity in pricing than the current zonal market design is unnecessary if sufficient transmission investments are made within each zone to ensure that all generation units in each zone are equally effective at meeting load at all locations in that zone. Price differences across these zones can be used to compensate or charge market participants for differences in the effectiveness of generation units located outside of each zone at meeting loads at all locations within the zone. It is also important to emphasize that if this commitment to have adequate transmission capacity within each zone is not honored by the transmission network owner it may be necessary to transition at some future date to greater spatial granularity in pricing in order to reliably operate the transmission network.

3.3. Number of Settlement Periods

Multi-settlement systems allow market participants to make firm financial commitments to produce and consume electricity in advance of delivery. This can provide long-start generation units and units with ramping constraints or minimum operating times with greater certainty about their real-time production levels. Under a single settlement market, prices are set based on actual or estimated real-time system conditions and suppliers are paid for all of their output and consumers are charged for all of their consumption at these prices. In a two-settlement market, participants typically first make sales and purchases in the day-ahead forward market for each hour of the following day. In real-time, participants then buy and sell imbalances between what they bought or sold in the day-ahead market and what they actually produced or consumed in real-time.

A two-settlement system can enhance system reliability and market efficiency because it limits the need for the system operator to issue a substantial volume of dispatch instructions in the real-time market because market participants have very strong incentives to limit their participation in the real-time market by limiting the magnitude of uninstructed deviations from their day-ahead schedules. For example, if a supplier produces less than its day-ahead schedule, then it must purchase the difference between this day-ahead schedule and its actual production from the real-time market at the real-time price. A supplier that injects less energy into the transmission network than its day-ahead schedule is very likely to face a real-time price that is substantially above the day-ahead price if the system operator had to call upon generation units with higher offer prices to provide the energy this supplier failed to provide. Consequently, any supplier that produces less than their day-ahead schedule faces the risk of buying this shortfall at a very high real-time price and any load that consumes more than its day-ahead schedule faces the risk of buying the additional energy at a higher real-time price. This property of multi-settlement systems provides strong incentives for participants to limit the amount of imbalance energy they must buy or sell from the real-time market.

A two-settlement system also provides suppliers with a strong incentive to respond to dispatch instructions issued by the system operator. In most two-settlement markets, dispatch instructions alter the supplier's final dispatch schedule. For example, if a supplier with a 500 MWh final schedule going into the real-time market receives a dispatch instruction to produce an additional 40 MWh of energy, its final schedule is now equal to 540 MWh. If its actual production is less 540 MWh then it must purchase the remaining energy from the real-time market at the real-time price. Consequently, if a supplier fails to respond to a dispatch instruction then it will be subject to the risk of buying the energy it did not provide out of the real-time market at a very high price. Equivalently, a supplier that fails to respond to a request from the system operator to produce less than its day-ahead schedule faces the risk of selling the energy it does provide beyond its final schedule at a very low or even a negative price, meaning that the supplier is paying to produce this incremental amount electricity.

There is little need for a multi-settlement system in a market where most of the electricity is produced from hydroelectric units that can start up and can ramp their level of production up and down very quickly. However, as the share of fossil fuel generation units with long-start times, ramping constraints or minimum run times increases, there is a greater need for a multi-settlement market to allow market participants and the system operator sufficient advance notice to manage these constraints without significantly degrading system reliability.

3.4. Ex Post versus Ex Ante Pricing in the Real-Time Market

Under ex ante pricing, real-time prices are set when dispatch instructions are issued. With ex post pricing, real-time prices are set after the dispatch instructions are issued and the injections and withdrawals from the transmission network are known. There are market efficiency and system reliability advantages and disadvantages to each kind of pricing mechanism.

An ex post market provides greater flexibility to the system operator because it can run the system in real-time by issuing dispatch instructions to generation units and loads according to the reliability criteria it determines are necessary to maintain system balance. Then the system operator can record the actual withdrawals of all market participants at all locations in the transmission network and run the market with these locational demand levels and the offers submitted by all generation unit owners to determine ex post market prices. A downside of an ex post pricing mechanism is that market participants do not know the price they will be paid when real-time dispatch instructions are issued. Typically, it will be the case that if a generation unit's actual output is close to the operating point set by the system operator the price it is paid will be above the generation unit's offer price. However, there is no guarantee this will be the case if the unit owner is producing at a level far from the operating point set by the system operator. There are various approaches to handling this problem. Some markets simply pay the unit owner the ex post market-clearing price that is below its offer price for the energy the unit provides. Others at least pay the supplier's offer price. Others have more complicated rules that attempt to penalize large differences between the unit owner's actual output and the real-time operating point set by the system operator.

The advantage of ex ante pricing is that the market participant knows the price it will be paid when a dispatch instruction is issued. This implies that a supplier will always be paid a price at least equal to its offer price to provide additional energy or pay a price that is no larger than its offer price to supply less energy. The downside of ex ante pricing mechanism is that the transmission network model used to set prices must more closely agree with the reliability criteria used by the system operator to manage the transmission network in real-time. If there are significant differences between the economic model used to set ex ante prices and these reliability criteria, market participants can exploit these differences to raise the price they are paid for responding to the system operator's dispatch instructions. Most multi-settlement markets in the United States use ex post pricing because it is extremely difficult to eliminate all of the possible differences between the network model used to set prices and the reliability criteria used to operate the transmission network. If a two-settlement short-term market is adopted in Brazil, then ex post pricing should be favored initially for the same reasons that it is used in most multi-settlement markets in the US.

3.5. Degree of Integration Between System and Market Operation

Some short-term wholesale markets have separate entities that undertake the wholesale market operation and the real-time system operation function, whereas others have these two functions contained in the same entity. A separate market operator and separate system operator requires the system operator to rely on explicit penalties to ensure that market participants respond to its dispatch instructions. Integration between system operation and market operation allows the system operator to use market prices to provide the incentives needed to ensure that market participants respond to the system operator' dispatch instructions.

As discussed earlier, these market mechanisms typically provide the system operator with greater flexibility to provide the precise financial incentive necessary to cause market participants to follow its dispatch instructions during each hour of the day. In contrast, fixed financial penalties for failing to follow the system operator's dispatch instructions are typically either too high or too low for the reasons described above. Consequently, if a two-settlement system for the short-term market is adopted in Brazil, then there appear to be clear system reliability and market efficiency gains from integrating the market operation and system operation functions.

3.6. Cost-Based versus Bid-Based Markets

A cost-based market uses the system operator's estimate of the variable cost of each generation unit to set the market-clearing price, which is typically equal to the variable cost of the highest cost generation unit necessary to meet demand. A bid-based market uses each generation unit owner's willingness-to-supply curve and each demander's willingness-to-purchase curve to set the market-clearing price of electricity, which is equal to the price at the point of intersection of the aggregate willingness-to-supply curve with the aggregate willingness-to-purchase curve. Under both regimes, all suppliers receive and all loads pay the market-clearing price.

For the case of a hydro-dominated electricity supply industry such as Brazil, the price-setting process for a cost-based market is considerably more complex than simply intersecting an aggregate marginal cost curve with the level of demand. The system operator must compute an opportunity cost of water for all hydroelectric generation units using the variable cost of fossil fuel generation units, the distribution of future water inflows, an administratively determined cost-of-deficit parameter, and some model of the characteristics of the transmission network. This locational opportunity cost of water is then treated as the hydroelectric generation unit's variable cost in determining the dispatch levels of all generation units and the market-clearing price in each congestion zone.

Choosing between these two approaches for setting prices in the short-term market requires taking into account both the politics and economics of electricity prices. Because all voters purchase electricity, the wholesale price of electricity is very politically visible. Moreover, as noted earlier, bid-based markets are extremely susceptible to the exercise of unilateral market power. During certain system conditions, the individual rationality constraint on supplier behavior described earlier can result in offer prices far above the variable cost of production of each generation unit. This can result in market-clearing prices vastly in excess of efficient levels. These market outcomes can arise for a variety of reasons, but all periods of high electricity prices are very politically visible and usually cause consumers and politicians to have a negative view of the future viability of the short-term electricity market.

The configuration of the transmission network and a supplier's location in this network can further enhance its ability to exercise unilateral market power. For example, system conditions can arise when one firm or small number of firms may be the only market participants able to meet a local energy need. Under these circumstances a supplier is said to possess local market power. Consider the following two-node example of this phenomenon given in Figure 1. There is 100 MWh of load and 80 MW of generation at node L1. There is 100 MW of generation and no load at node G1. A transmission line with capacity equal to 60 MW connects G1 and L1. The generation capacity at L1 is owned by many small suppliers and each unit has a regulated variable cost of $50/MWh. All of the generation capacity at node L1 is owned by a single firm and has a regulated variable cost of $80/MWh. If a bid-based market is used to set the prices at G1 and L1, then there is a no limit on the price that the generation unit owner at L1 can charge to supply 40 MWh of energy. That is because a maximum of 60 MWh can be transferred from G1 to L1 to meet the 100 MWh demand at L1. This implies that at least 40 MWh of energy must be supplied by the generation unit at L1, regardless of its offer price, or demand will not equal supply at L1.

In a bid-based market, this local market power problem is avoided by the fact that the regulated variable cost at L1 is $80/MWh and the regulated variable cost G1 is $50/MWh so that a cost-based market will set a marking-clearing price of $50/MWh at G1 and $80/MWh at L1 and 60 MW will flow from G1 to L1. The above example illustrates the economic and political advantage of a cost-based market. It limits the ability of suppliers to raise wholesale electricity prices very far above efficient levels even in the extreme circumstances when a supplier would possess enormous unilateral market power if the market price was set through a bid-based market mechanism.

This example illustrates the essential need for a local market power mitigation mechanism in any bid-based short-term market for wholesale electricity. This mechanism determines when a supplier possesses local market power worthy of mitigation because of the characteristics of the transmission network. It then determines the supplier's price offer when its generation units are subject to mitigation. Finally, the mechanism determines what price the market participant will receive for output sold from mitigated generation units. All bid-based wholesale markets in the United States have local market power mitigation mechanisms to limit the ability of suppliers to exercise local market power. If Brazil decides to adopt a bid-based short-term market it should have as a pre-condition the existence of a stringent local market power mitigation mechanism.

Bid-based markets can also allow significant wealth transfers and economic disruption to occur in a very short time period as a result of participants exploiting market design flaws. For example, in California unilateral market power enabled by a market design flaw (inadequate fixed-price forward contracting by the three largest load-serving entities in California) led to approximately $5 billion of transfers from consumers to producers of electricity during the last six months of 2000. Borenstein, Bushnell and Wolak (2002) quantifies the magnitude of these market inefficiencies. For similar reasons, significant wealth transfers and economic disruptions occurred during the winter of 2001 and autumn of 2003 in New Zealand. Virtually all short-term wholesale electricity markets in industrialized and developing countries have experienced significant market performance problems due to market design flaws that have resulted in substantial wealth transfers from consumers to producers.

It is important to emphasize that market rules that prohibit and/or penalize market power abuse are extremely unlikely to prevent these instances of substantial wealth transfers. That is because it is extremely difficult to distinguish between the illegal abuse of market power and the legal exercise of unilateral market power. Market power abuse as seen by one player may be superior business acumen as seen by another. Moreover, the exercise of unilateral market power is typically not prohibited under antitrust or competition law. Despite the enormous wealth transfers from electricity consumers to electricity producers in California during the last six months of 2000 mentioned above, no supplier was convicted of abuse of market power, market manipulation, or collusion. Similar statements hold for New Zealand and other markets around the world that have experienced periods of significant unilateral market power.

These facts emphasize a significant downside associated with bid-based markets if the individual rationality constraint on market participant behavior is not adequately addressed. A bid-based market can provide suppliers with significant opportunities to exercise both system-wide and local market power. Prohibitions on market manipulation can only prevent the most egregious forms of the exercise of both system-wide and local market power. Massive wealth transfers can occur in a very short period of time as a result of the exercise of unilateral market power. Consequently, a bid-based short-term market should not be adopted without the necessary regulatory safeguards and strong incentives for market participants to realize benefits beyond those achievable from a short-term cost-based market.

There are a number of mechanisms for limiting the ability and incentive of suppliers to exercise unilateral market power in bid-based wholesale electricity markets. The primary mechanism for limiting the incentive to exercise unilateral market power in the short-term market is fixed-price, long-term contracts between suppliers and retailers negotiated far in advance of delivery. As discussed in Wolak (2000), the larger the quantity of fixed-price forward contract obligations a supplier has relative to its actual output level, the smaller is the incentive that supplier has to exercise unilateral market power in the short-term market. In addition, if a supplier's fixed-price forward contract obligations exceed its output level then the supplier has an incentive to take actions to reduce the short-term wholesale market price below its variable cost. The primary mechanism for limiting the ability of suppliers to exercise unilateral market power is the local market power mitigation mechanism described above. This mechanism automatically mitigates the bids of suppliers that are deemed to possess substantial local market power.

Other mechanisms for limiting the ability of suppliers to exercise unilateral market power are bid caps on the short-term market and price caps on the short-term market. A bid cap places a limit on the maximum price offer that any supplier can submit. For example, all United States (US) wholesale markets have bid caps on their energy and ancillary services markets. For all of the short-term wholesale markets in the eastern US, the current bid cap on their short-term energy market is equal to $1,000/MWh. A price cap limits the maximum value of the market-clearing price. Price caps provide greater certainty with respect to the maximum market-clearing price, but bid caps are typically favored in zonal-pricing or nodal-pricing markets because it is difficult to limit the maximum price and still properly price transmission congestion in a manner that reflects loop flow constraints. For example, even if the offers of all suppliers are below $1,000/MWh it is possible to have a zonal or nodal price above this level because of loop flow constraints that require backing down one generation unit and increasing another generation unit in order to meet demand in a zone or at a node.

4. Benefits of a Bid-Based Short-Term Market for Brazil

This section considers the benefits to Brazil from implementing a bid-based short-term market assuming that the market power problems described above have been addressed. Brazil currently requires 100% coverage of final demand by fixed-price forward contracts, so one remaining regulatory safeguard for this issue would be imposing the requirement that a significant fraction of these contracts are signed far enough in advance of delivery to allow new entrants to compete to supply them. The second outstanding issue is establishing a local market power mitigation mechanism. There are a number of mechanisms in place in US markets that can be easily transferred to the Brazilian context. The ones in place in the PJM market and the California market in the US appear to be the most effective and best-suited to Brazil.

4.1. A Comparison of Brazil to Bid-Based Hydroelectric Dominated Markets

In order to determine if Brazil would benefit from a bid-based market, this section compares the performance of the cost-based Brazilian market to the performance of bid-based markets with a substantial share of hydroelectric generation capacity in New Zealand, Colombia and Norway. We believe that a key determinant of the difference in performance between the Brazilian market and the markets in these countries is the fact that Brazil uses a cost-based market instead of a bid-based market.

The first major difference between the Brazilian market and the markets in New Zealand, Colombia, and Norway is in the distribution of prices. Figure 2 plots the annual histogram of the natural logarithm of weekly prices, ln(price), from the Brazilian market for the years 2004, 2005, 2006, and 2007. We plot the natural logarithm of average prices instead of the level of prices, because the distribution of prices for all of the markets is positively skewed. This transformation requires that I omit infrequently occurring negative and zero prices from the analysis, but has the benefit of producing a nearly symmetric annual distribution of prices for three bid-based markets for several years. Figure 3 plots annual histogram of the natural logarithm of the daily average price for the New Zealand market for 2001 to 2005. This price is computed as mean of the 48 half-hourly volume-weighted average nodal prices. Figure 4 plots the annual histogram of the natural logarithm of the daily-average price for the Colombia market for 2000 to 2005. This average price is the sample mean of the 24 hourly prices from the Colombian market. Finally, Figure 5 plots the annual histogram of the natural logarithm of the daily average price for the Oslo zone from the Elspot market in Norway for 1997 to 2007. This average price is the sample mean of the 24 hourly prices.

Comparing the plots in Figure 2 to those in Figures 3-5 yields the following observations. For all of the bid-based markets, the distribution of the logarithm of prices exhibits a central tendency with the highest frequency of the ln(price) realizations concentrated at the center of the histogram, with lower frequency of ln(price) realizations above and below this level. In contrast, the annual histograms of Brazilian prices do not exhibit any measure of central tendency. The highest frequency value of ln(price) are at the far left of the histogram, with the remaining ln(price) realizations approximately uniformly spread across the range of possible prices. It is difficult to argue that the Brazilian market price histograms provide a useful signal of the value of water. The mean of the ln(price) in Brazil is substantially lower than the median value of ln(price). There is substantially more agreement between these two measures of central tendency for the ln(price) in New Zealand, Colombia and Norway. The dramatically different shape of the annual ln(price) histograms for Brazil relative to the annual histograms for the other countries provides strong evidence in favor of the ability of a bid-based market to provide a more reliable measure of the opportunity cost of water, which should improve both short-term market efficiency and longer-term electricity supply security.

An important role of price in a hydroelectric dominated market is to signal the value of water to all market participants. When water levels are high and future inflows are likely to be high, the price of water should be low to indicate that using water to produce electricity in the current period has a low opportunity cost. When water levels are high and future inflows are likely to be low, the price of water should be high indicating that using water to produce electricity in the current period has a high opportunity cost. This next sequence of plots compares how well the Brazilian cost-based market signals the value of water compared to the bid-based markets in New Zealand, Colombia, and Norway.

Figure 6 plots the relationship between a measure of the daily water level relative to capacity and the ln(price) for the Brazilian market for that day. I have collected information on the water level in all reservoirs in Brazil each day from 2004 to 2007. For each day during this time period, I compute the variable Fraction(t) as follows:

$$\text{Fraction}(t) = (\text{Water\_Level}(t) - \text{Water\_Min})/(\text{Water\_Max} - \text{Water\_Min}),$$

where Water_Level(t) is the water level in day t and Water_Min is the minimum value of Water_Level(t) observed over the entire sample period and Water_Max is the maximum value of Water_Level(t) observed over the entire sample period. Each plot in Figure 6 displays the daily value of Fraction(t) and associated value of ln(price) for that day. The solid line in each figure is a smoothed kernel regression through these points. For two of the four years plotted, this line is upward-sloping, contradicting the logic that when Fraction(t) is low the opportunity cost of water is high so that market prices should be high. Even for the years when this line is downward sloping, few of the points are very close to the line. Figure 7 plots these same daily values for New Zealand. In this case, the smoothed regression line through the points takes on the expected downward slope and the points tend to be clustered around the line. Figure 8 repeats this plot for Colombia. Once again the expected downward slope is present for all but one year and the points tend to be clustered around the smoothed regression line. For Norway, only weekly water levels are available. For this market, Figure 9 plots the natural logarithm of the daily average price against the weekly value of Fraction(t) for that day for each year from 1997 to 2007. These plots also largely confirm the expected downward sloping relationship between Fraction(t) and ln(price) and the points tend to be clustered around the smoothed regression lines.

The analysis in Figures 6 to 9 provides visual evidence that the three bid-based markets provide a more accurate estimate of the current opportunity cost of using water to produce electricity. For all of these markets and virtually all years, there is a clear downward-sloping relationship between the value of Fraction(t), a measure of the water level relative to capacity, and the natural logarithm of the daily average price. The bid-based markets also appear to provide a smoother transition between high prices and low water levels and low prices and high water levels, whereas in Brazil there appear to be discrete jumps between these two states. These empirical results indicate that there are significant market efficiency and system reliability benefits from adopting a bid-based short-term market in Brazil.

4.2. Sources of Benefits of Bid-Based Market for Brazil

There are three potential sources of benefits from introducing a bid-based short-term market in Brazil. First, is the potential for a lower cost short-term solution to meeting demand despite multiple hydroelectric unit owners on the same river system. Second, is the potential for a lower cost long-term solution due to improved estimates of the current opportunity cost of water relative to the current cost-based system. The final source of benefits is the increased opportunities for active demand-side participation in a bid-based short-term wholesale market.

An often-claimed reason for the use of a cost-based dispatch in Brazil is the need to coordinate production by multiple generation unit owners on the same river system. The production decisions of upstream unit owners can impose costs on unit owners located downstream. This is an example of what economists call a "negative externality," because the action of the upstream firm determines the amount of water available to the downstream firms. The argument in favor of a centralized cost-based dispatch mechanism is that it is the only way the least-cost dispatch of all hydroelectric generation units on the same river system can be found. This argument overlooks a large literature in economics dealing with the question of whether market mechanisms can internalize these negative externalities to find the least-cost dispatch for the entire river system.

The general question of whether negative externalities can be internalized has been extensively dealt with in the economics literature. The Coase Theorem, named after the Nobel prize-winning economist Sir Ronald Coase, deals with precisely this question. Coase (1960) argues that when there are no transactions costs, bargaining will lead to an outcome that internalizes the externality regardless of the initial allocation of property rights. In the present context, the Coase Theorem implies that if the cost of generation unit owners on the same river system bargaining over how to allocate water between their hydroelectric units is small, then the least-cost dispatch can be achieved through a de-centralized market mechanism. Suppliers can bargain among themselves over how to allocate the revenues they earn from selling water as electricity in order to achieve the least-cost dispatch.

In addition, if all suppliers on the river system have fixed-price forward contract obligations for a significant fraction of their expected output, then they all have a common interest in minimizing the total cost of serving these fixed-price forward contract obligations. Consequently, high levels of fixed-price forward contract obligations by generation unit owners on the river system negotiated far enough in advance of delivery to allow new entrants to compete to provide this electricity will provide strong incentives for these generation unit owners to bargain among themselves to find the least-cost dispatch of all of the units on the same river system.

The following example illustrates how this might work. Consider the case of two generation unit owners where Owner U has units upstream from Owner D. There are two possible modes of operation of the river system. Using mode A, U earns $1000 and D earns $500 from selling water as electricity. Using mode B, U earns $900 and D earns $800. Because it achieves a greater total surplus than mode A, mode B is the efficient solution to the water allocation problem. However, U has the unilateral incentive to choose mode A because it earns $1000 as opposed to $900 under mode B. Because U is upstream from D, it can choose the mode of production. The efficient solution can be implemented even if U is allowed to choose the mode of production if D promises to pay U $150 to choose mode B. In this case, U earns $1050, which is more than the $1000 it would earn under mode A. D earns $650, which more than the $500 it would earn under mode A. This example demonstrates how bargaining among U and D would lead U to choose the efficient water allocation scheme.

Given the small number of generation unit owners on each river system, it is very likely that the conditions necessary for the Coase Theorem to hold are at least approximately valid. Moreover, if Brazil keeps its 100% forward contracting requirement on all loads, and adds the requirement that a significant fraction of these contracts must be negotiated far enough in advance of delivery to allow new entrants to compete to sell the contracts, then all suppliers with significant fixed-price forward contract obligations will have a strong incentive to find the efficient solution to allocating the available water across the generation units on the same river system. In fact, there is even an argument that this mechanism will do a better job than the centralized cost-based dispatch model because it uses an opportunity cost of water that is determined from a bid-based market. As we discuss below, this opportunity cost reflects the consensus of the information of all market participants rather than only information of the system operator about the value of water.

There is also a large literature in economics demonstrating that market mechanisms often aggregate all relevant private information possessed by all market participants into the market price. These models demonstrate the conditions necessary for the market price to be a "sufficient statistic" for the private information possessed by market participants about the product sold. In the context of a bid-based market with a substantial hydroelectric capacity share these results imply that the market price aggregates all private information possessed by all market participants about the value of holding water in storage. A major complaint about the current cost-based market is that the model solution sometimes contradicts the consensus of market participants and policymakers on when to use water to produce electricity and how much water to use. For example, the Brazilian Committee on Supply Security (CMSE) recently over-ruled the dispatch model's schedule for thermal generation units. CMSE required that thermal units be turned on when the current dispatch model said that they should not.

There are a number of reasons why the current cost-based dispatch model may not produce the best possible estimate of the opportunity cost of water. First, the stochastic dynamic programming model that determines the opportunity cost of water uses information that is likely to be a poor predictor of future system conditions, such as the distribution of water inflows or the cost of a water shortfall that leads to a rationing event. For example, the historical distribution of water inflows is used as the estimated distribution of future water inflows in the stochastic dynamic programming model. This historical distribution may be a poor predictor of future water inflows for a variety of reasons. For example, global climate change or changes in land use near the river system may cause the future distribution of water inflows to change significantly from the historical distribution. In a bid-based market, participants can take these changes into account in formulating their willingness-to-supply energy curves, but in the current cost-based market this information is ignored in the price-setting process.

The cost of deficit parameter used in the stochastic dynamic programming model is set through an administrative process. This parameter is a key determinant of the level of market prices and the probability that a rationing event will occur. Figure 10 contains the current cost of deficit function. This function shows that for the first five percent load reduction the cost of deficit is 944.51 Reals/MWh, which translates into approximately $600/MWh at current exchange rates. Using this cost-of-deficit parameter in the stochastic dynamic program that computes the value of water assumes that firm load would be curtailed at this price. However, the total political and economic cost of curtailing firm load is likely to be substantially higher than 944.51 R$/MWh.

If the cost-of-deficit parameter is set too low in a cost-based market, this increases the likelihood that the system operator will eventually need to curtail firm load because of a water shortfall. However, a low cost-of-deficit parameter also reduces average electricity prices because it implies an artificially low opportunity cost of using water to produce electricity in the current period. Consequently, there is a clear incentive for a government interested in keeping electricity prices low in a hydro-dominated system to set a cost-of-deficit parameter that leads to an unacceptably high probability that the system operator will need to curtail firm load. Wolak (2003) calls this the "gambling with the weather problem" in cost-based markets.

In a bid-based market, there is no need to set a cost of deficit parameter. All loads submit their willingness-to-purchase function into the wholesale market and the intersection of this function with the aggregate willingness-to-supply function of all generation unit owners determines the market-clearing price. Consumers that are unwilling to pay more than the market-clearing price will have their demand for electricity reduced. There is no need to involuntarily curtail firm load. Unless there is an offer cap on the short-term market, the market-clearing price can always rise to the level necessary to cause supply to equal demand. This aspect of a bid-based short-term market provides further motivation for the requirement for all loads to have a very large fraction of their consumption covered by fixed-price forward contracts signed far in advance of delivery. This will allow the short-term price to rise to the level necessary to cause real-time supply to equal real-time demand without causing significant economic harm to electricity consumers. This logic implies that the portion of final demand that is not covered by fixed-price forward contracts should be limited to the fraction of final demand that is discretionary. All consumption that is essential should be covered by fixed-price forward contracts signed far in advance of delivery to ensure this energy will be supplied in real time.

Other assumptions that determine the market-clearing price in a cost-based market are the distribution of future demand growth and the future availability of all generation units. The stochastic dynamic programming model used to determine the opportunity cost of water in Brazil must make assumptions about the distribution of future values of these two variables. These assumptions can significantly influence market prices, and they are also unnecessary in a bid-based market. Market participants base their willingness-to-supply and willingness-to-purchase curves on their own perceptions about these variables. The resulting market-clearing prices represent the consensus view among all market participants of all of the factors that determine the opportunity cost of using water to produce electricity in the current period.

Because of the substantial uncertainties about the cost of a future supply shortfall, the future distribution water inflows, load growth, and generation unit availability, there are likely to be significant inaccuracies in the future distributions or values of these variables which can lead to a value of water and market-clearing prices of electricity that do not reflect the consensus of market participants about future system conditions in a cost-based market. In contrast, ina bid-based market all market participants have a say in determining the value of water and the price of electricity. Each market participant can also have information about the cost of the future supply shortfalls, the future distribution of water inflows, load growth and generation unit availability. The market-clearing price should reflect all of this information.

This feature of a bid-based market also makes it better able to respond to extreme system conditions that are not representative of historical conditions. A cost-based market will manage water appropriately to the extent that actual future system conditions are well-represented by historical conditions. When extreme conditions arise that are not well-represented by historical conditions, a cost-based market is more likely to make significant mistakes in using and pricing water. That is because it can only respond to circumstances according to the inputs fed into the stochastic dynamic programming model. A bid-based market looks at a far wider set of information in formulating its response to current system conditions. For example, if the consensus of all market participants or significant fraction of market participants, is that water should be conserved because future water inflows are likely to be lower than is predicted by historical data or future load growth is likely to be much higher than is predicted by historical load growth, these market participants will submit willingness-to-supply and willingness-to-purchase functions that are reflective of these beliefs and market prices will rise to the level necessary for less water to be consumed because more fossil-fuel generation units are operated.

The experience of several hydro-dominated systems with bid-based markets versus the former cost-based dispatch regime provides evidence in favor of the validity of this logic. During 1992, before a bid-based market regime was in place in either New Zealand or Colombia, both countries experienced rationing periods because water levels fell below critical levels. However, during the bid-based regime, both countries have experienced periods of low water inflows comparable to 1992. In Colombia, these low water inflows occurred in 1997 and 1998 and in New Zealand these occurred in 2001 and 2003. In all of these cases, the need to rely on rationing was avoided because market prices rose in advance of these events to causes consumers to demand less energy and fossil fuel units to run more intensively. Figure 11 plots the pattern of wholesale prices during the winter of 2001 and Figure 12 plots the pattern of hydro storage levels in New Zealand for 1992 (the year in which rationing occurred) and 1999, 2000 and 2001 (up through the month of June). The rate of decline of water levels in 2001 was very similar to that in 1992 during the first six months of 2001, but the higher prices during the June to September 2001 time period slowed the rate of water use and increased the use of fossil fuel units, which prevented a rationing period from being declared.

This increased flexibility to respond to extreme system conditions provides a strong argument in favor of adopting a bid-based market for a hydro-dominated system such as Brazil, where policymakers are very concerned about preventing rationing periods. Moreover, if the requirement is maintained for loads to have 100% of the final demand covered in fixed-price forward contract obligations, with a substantial fraction signed far enough in advance of delivery to allow new entrants to compete to supply these contracts, then consumers will be insulated from the high short-term prices necessary to dispatch the higher cost fossil fuel units needed to conserve water. In addition, consumers would also have the option to sell back some their fixed-price forward contract obligations at very high prices if they are able to reduce their demand for electricity during this time period.

A final advantage of a bid-based market is that it increases the difficulty that the government or regulator can use the process of pricing wholesale electricity to pursue political ends. As discussed above, there are no modeling assumptions or model inputs such as the cost-of-deficit parameter or distribution of future system conditions that the government or regulator can set to exert a direct influence on prices in a bid-based market. Because market prices are typically set by the highest price offer necessary to meet demand, even if a substantial fraction of generation capacity is in the hands of state-owned firms, as long as some supply from the remaining firms is necessary to meet demand, how the government-owned generation units are offered into market is unlikely to depress prices significantly. However, if these government-owned firms do not offer sufficient capacity into the market then this is likely to increase market prices if these firms control enough of the generation capacity. This outcome can be prevented using the same mechanism that limits the incentive of privately-owned, profit-maximizing suppliers from taking actions to raise market prices, by requiring final demand to purchase virtually all of its requirements in fixed-price forward contracts negotiated far in advance of delivery.

5. Recommended Bid-Based Short-Term Market for Brazil

This section proposes a transition process to a multi-settlement bid-based short-term market for the Brazilian electricity supply industry. I first describe a rudimentary bid-based short-term market that requires minimal changes in the existing Brazilian market design. This short-term market should capture many of the benefits of bid-based markets described above while limiting the potential downside of market mechanisms. I then describe a transition process for achieving my long-term recommendation of multi-settlement, bid-based short-term market for Brazil.

As discussed in Section 2 of this report the fundamental market design challenge of bid-based short-term markets is limiting the ability and incentive firms have to exercise unilateral market power. Consequently, it cannot be emphasized enough that without the appropriate safeguards in place against the exercise of unilateral market power, Brazil should not adopt even the rudimentary bid-based, short-term market proposed below. These necessary safeguards are: (1) coverage of close to 100% of final demand in fixed-price forward contract obligations negotiated far enough in advance of delivery to allow new entrants to compete to supply these contracts, (2) a local market power mitigation mechanism that applies to all market participants, (3) a cap and floor on supply offers into the short-term wholesale market, and (4) a prospective market monitoring process with public release of all data necessary to operate the short-term market. Wolak (2004) provides a detailed discussion of the necessary features of an effective market monitoring process for a bid-based short-term wholesale electricity market. All of these safeguards must be in place before even a rudimentary bid-based short-term market is implemented in Brazil.

Sections 3 and 4 emphasized that the primary advantages of a bid-based short-term market are more efficient management of available water within the same river system and throughout the year across multiple river systems. Arguments were provided for why a bid-based market would be better able to deal with extreme system conditions. The key feature of a bid-based market that leads to these superior market efficiency properties is that each market participant has control over when it sells or consumes electricity and that these decisions collectively influence the market price. Both of these features are necessary for a bid-based market to have favorable market efficiency properties. Bid-based markets that do not have both of these features (bids and offers by market participants impact how much each entity produces or consumes and the aggregation of these bids and offers determine market-clearing prices) are unlikely to achieve the market designer's goals.

5.1. Incomplete Solutions May Be Worse than Existing Cost-Based Model

It is important to recognize that bid-based short-term markets where the collective actions of market participants do not influence both market-clearing prices and the amount of hydroelectric energy produced each pricing period may lead to less efficient water use that the existing cost-based market. For example, one possible proposal for implementing a bid-based short-term market would continue to dispatch generation units using the existing cost-based model, but use a different mechanism to set short-term prices. Market participants would be allocated rights to sell the electricity they produce based on some allocation scheme. For example, the amount of energy reallocation mechanism (MRE) capacity each market participant is allocated could be used to determine the amount energy a supplier is allocated to sell each day.[1] Specifically, each supplier could be given their share of total MRE capacity in energy to sell each day. The total amount of energy each supplier is allowed to sell on an annual basis therefore equals annual electricity production in Brazil times that supplier's share of the total MRE capacity.

[1] My understanding of the operation of the MRE mechanism is based on von der Fehr and Wolak (2003).

Weekly market-clearing prices could then be determined in the following manner. Each week suppliers would submit their willingness to supply energy from the accumulated stock of energy they are allocated. The market operator would aggregate these individual willingness-to-supply curves across all market participants and intersect this aggregate supply curve with actual demand for the week to compute a market-clearing price for the week.

Prices could also be set on a daily or hourly basis by requiring suppliers to submit willingness-to-supply curves on a daily or hourly basis and intersecting the aggregate supply curve with the daily or hourly demand. Although this mechanism would set prices based on the offers of suppliers, these offers would have no impact on how much water is used to produce electricity during each day of the week or hour of the day, because which units are dispatched is still determined by the cost-based market mechanism. This mechanism would only allow market participants to determine when they sell their allocation of energy throughout the year.

This mechanism violates a necessary condition for a bid-based market to improve market efficiency and system reliability because market participants are unable to exert any control over how much water is used to produce electricity each day. If the cost-of-deficit parameter is set incorrectly or the distribution of future hydro conditions differs significantly from the historical distribution of hydro conditions, the existing cost-based dispatch model will use the available water in an efficient manner. Because market participants have no way to influence when the water they have rights to will be used to produce electricity, it is unclear why they would want to sell more or less of their allocated water in one day versus another. Consequently, these prices are unlikely to have any of the favorable market efficiency or system reliability properties described in Section 4. Despite the fact that they are the result of a bid-based market mechanism, these prices are likely to be just as poor predictors of the value of water as those that result from the current cost-based dispatch model, because the market-clearing price does impact the amount of water used to produce electricity. In fact, this bid-based short-term market would attempt to set prices that rationalize the pattern of water utilization resulting from the solution of the cost-based dispatch model.

5.2. Dimensions of Proposed Market Design

This section characterizes the proposed short-term market in terms of each of the dimensions of the market design process described in Section 3. I first describe the ultimate goal of the market design process—a two-settlement, zonal-pricing market with ex post pricing that set prices at hourly time intervals. The distinction is made between market design characteristics that will be implemented immediately and those that will be added as market participants gain experience with a bid-based market mechanism. Rationales will be offered for why certain features should be delayed and why others should be implemented as soon as possible.

The first dimension of the ultimate market design is the degree of temporal price granularity. As noted in section 3, as the amount of thermal generation in the Brazilian system increases there are greater losses in market efficiency associated with fixing prices for longer periods of time. Consequently, the argument for daily or even hourly pricing is stronger than it was even a few years ago. Nevertheless, it seems prudent to implement a process for transitioning from the current weekly-pricing process to more granular pricing. Therefore, the bid-based market would start by setting prices on a weekly basis, then move to daily pricing, and finally to hourly pricing.

Section 3 also described the logic behind introducing more spatial granularity in pricing. However, if sufficient transmission network investments are undertaken to allow all generation units within each existing zone of the Brazilian market to be equally effective at serving load at all locations in this zone, then there is little need to introduce greater spatial granularity in pricing. Even with this commitment to enhance the transmission network within each zone to achieve the goal of equal effectiveness of all suppliers in the zone, there is still the need for a prospective local market power mitigation mechanism to limit the ability of suppliers to exercise unilateral market power. A local market power mitigation mechanism is also needed even if greater spatial granularity in pricing beyond the current zonal market design is introduced into the Brazilian market.

Although there are clear advantages to a multi-settlement market, it is not necessary for this feature to be implemented in the initial bid-based market. A single set of market prices can be set based on offers of generation unit owners. As market participants and the system operator become familiar with the operation of a bid-based market, a transition process to a multi-settlement system with a day-ahead forward market and a real-time imbalance market can be implemented.

Because of the flexibility it provides to the system operator, ex post pricing should be implemented for the real-time market. During the initial stage when a single settlement market is in place, prices should be set using the actual consumption of all market participants in each zone as the level of demand in that zone and the actual offers submitted by all generation unit owners in each zone should be used to construct the aggregate willingness to supply curves in each zone.

In a two-settlement system, all market participants could submit portfolio bids into the day-ahead forward market. A portfolio bid or offer curve implies that the market participant does not have to identify the specific generation unit that will supply the necessary energy or the load that will consume the energy. Day-ahead prices will be determined by the intersection of the hourly aggregate demand curve with the hourly aggregate supply curve. The market operator would then compute each participant's net energy obligations within each congestion zone as the difference between the amount of supply offered at or below the market-clearing price minus the amount of demand offered in at or above the market-clearing price. Each market participant would then be required to submit generation or load schedules in each congestion zone equal to their net energy obligations from the day-ahead market. These generation schedules would be required to be unit-specific, whereas load schedules could be at the zonal level. This means that if a supplier has a net energy obligation at the close of the day-ahead market of 1500 MWh in a congestion zone during one hour of the following day, then it would be required to identify 1500 MWh of unit-specific generation schedules for the following day. This could be accomplished by the supplier scheduling three 500 MW capacity generation units to run at their nameplate capacity for the hour. It is important to emphasize that these generation unit-specific schedules are firm financial commitments in the sense that if the supplier did not provide this amount of MWh from each generation unit, it would be required to purchase or sell the difference in the real-time market at the real-time price.

For the real-time market in a two-settlement system, generation unit owners would be required to submit willingness to supply curves and willingness demand curves relative to their day-ahead schedules that the system operator can use to manage real-time imbalances. For example, if a generation unit owner scheduled 300 MWh from one of its units following the close of the day-ahead market, then this market participant would be required to submit a non-decreasing curve expressing its willingness to increase its output and decrease output of this generation unit relative to its day-ahead schedule in the real-time market. Each generation unit with a day-ahead schedule and any generation unit wishing to sell in the real-time market would be required to submit similar willingness-to-supply curves. For reliability reasons, it is preferable that offers into the real-time market be generation unit-specific in the sense that each offer curve is associated with a generation unit rather than the supplier submitting an aggregate portfolio of offer curves and then deciding how to supply the energy accepted by the market operator in the real-time market from the generation units that it owns. These generation unit-specific offer curves would be used to the set real-time price in each congestion zone based on the actual consumption in that zone using an ex post pricing mechanism.

As noted in Section 3, there are significant reliability and market efficiency benefits from combining the system operation and market operation functions. However, the current separation between these two functions should not be a barrier to implementing a bid-based market, but a transition mechanism should be put in place to integrate these two functions. If ex post pricing is implemented then the market operator can compute prices using the offers submitted by suppliers and actual demand after the fact. As also discussed in Section 3, if there is separation between system operation and market operation, the system operator must penalize market participants for failing to respond to the system operator's instructions. Less rigid penalty mechanisms are necessary if there is greater integration between the system operator and market operator function.

5.3. Transition Path to Proposed Market Design

This section proposes a transition process from the existing market design to this ultimate market design described in the previous section. Most of the dimensions of the first step in this transition process entail small changes from the current regime. The only substantial modification is the use of bids and offers rather than regulated costs and an optimization model to set market-clearing prices. Slowly, more features are added in the transition process until the ultimate goal is achieved. I would expect this transition process to take place over a 12 to 18 month period.

There are number of possible transition paths to the ultimate goal of a two-settlement, zonal, ex post pricing, bid-based market with hourly pricing in the day-ahead and real-time market. The one I favor starts with weekly pricing, consistent with the current cost-based market. Each week, all suppliers would submit technology-specific offers to provide energy in each congestion zone on a weekly basis. For example, each supplier would submit their willingness to sell two types of energy: hydroelectric energy and all other types of energy. Using a forecast of total demand in each congestion zone for the following week, the market operator would determine amount of each type of energy sold in that congestion zone and a preliminary and non-financially binding weekly market-clearing price. Information on the total amount of hydroelectric energy and other energy sold in each congestion zone would be passed on to the system operator, who would then be required to schedule generation units in each congestion zone to achieve these weekly technology-specific production levels. Generation units would receive their generation schedule and real-time dispatch instructions from the system operator throughout the week, just as they do under the current cost-based market. At the end of the week when demand is known, ex post prices would then be set for each congestion zone based on actual consumption in the congestion zone using the same supplier offer curves used to set the weekly technology-specific output levels. Each generation unit would be paid this market-clearing price for their actual output during the week.

To ensure that suppliers follow the dispatch instructions of the system operator the existing penalty scheme for ensuring that suppliers follow dispatch instructions of the system operator should be maintained. As an additional safeguard against the exercise of unilateral market power, the market operator could exclude from the price-setting process the willingness-to-supply curve of a market participant that deviates too much from the system operator's dispatch instructions. For example, suppose that the system operator's schedule for generation units and the dispatch instructions issued within the operating hour for a supplier within a given congestion zone imply that it should have produced 500 MWh during the hour from all of its generation units in that zone. If the supplier is producing more than 110 percent of this magnitude or less than 90 percent of this magnitude the supplier's actual production will be subtracted from the total demand in the zone and that supplier's offers will be excluded from the price-setting process.

This approach to scheduling generation units and pricing can be extended to setting daily prices. Each day suppliers would submit technology-specific willingness-to-supply functions for each congestion zone and the market operator would use a forecast of the demand for the following day for each congestion zone to set zonal generation schedules for hydroelectric energy and all other sources of electricity for the following day. These schedules would be passed to the system operator, who would then dispatch generation units for each hour during the following day within each congestion zone to meet these aggregate zonal schedules for hydroelectric energy and all other energy sources. In real-time, the system operator would dispatch generation units using the existing cost-based model to maintain system balance. Ex post daily market-clearing prices could then be set using the actual demand in each congestion zone and these technology specific aggregate demand curves in each congestion zone. All suppliers would then be paid for their actual output during the day at these market-clearing prices.

The weekly bid-based market should operate for at least six months to familiarize market participants with how their generation units are scheduled to meet the weekly hydroelectric and all other energy sources schedules emerging from the week-ahead market. It should then be straightforward to transition to the day-ahead single settlement market described above. The bid-based market will still only determine the split between hydroelectric versus all other energy sources, but the system will now be constrained to do this on a daily instead of a weekly basis. This will require another learning process for both market participants and the system operator to understand how to achieve these daily aggregate output levels for hydroelectric and all other energy sources.

The final step in the transition process starts with the integration of the system operation and market operation functions. This should reduce the cost and system reliability consequences of operating a two-settlement short-term market. Once this integration is in place, the day-ahead bid-based market could transition to a two-settlement bid-based short-term market. In the day-ahead market suppliers and loads would submit portfolio offers and bids for each hour of the following days. Following the close of the day-ahead market for the 24 hours of the following days, suppliers and loads would submit generation and load schedules that are financially binding for each hour of the following day to the system/market operator. Before the start of the real-time market for each hour, suppliers and participating loads would submit willingness to supply upward and downward movements in the energy relative to final schedules at the generation unit-level or resource-level in the case of participating loads. These resource-specific offers and bids would be used to manage real-time system imbalances. Then real-time prices would be set on an ex post basis using the actual imbalance energy demand (which could be positive or negative) in that congestion zone. Because of the two-settlement structure there would be less need for explicit penalties for suppliers and loads failing respond to real-time dispatch instructions because they would be required to buy any energy they do not provide at the real-time price or sell any energy they did not have a firm financial commitment to provide at the real-time price.

Clearly, there are many details to be worked out with respect to the final design of this two-settlement, zonal, ex post-pricing bid-based short-term market. However, there are clear market efficiency and system reliability benefits from implementing the rudimentary weekly ex post pricing bid-based market described above and it will only require relatively minor changes in the current market design. Given the potential inefficiencies of the current cost-based market identified in Section 3, implementing this rudimentary bid-based market appears to be a risk worth taking.

6. Conclusions

This report described the initial conditions necessary for the introduction of a bid-based short-term market for the Brazilian electricity supply industry and proposed a transition process to two-settlement, zonal, ex post-pricing bid-based market from the current cost-based market. To provide a theoretical foundation for this discussion, I first introduced the concept of an electricity market design process. Here I pointed out the two fundamental challenges of a market design process, how to obtain: (1) technically and allocatively efficient production, and (2) economically efficient pricing of wholesale electricity.

The six major dimensions of the short-term electricity market design process were introduced and discussed with reference to the Brazilian wholesale electricity market. The major focus of this discussion was the question of a cost-based versus bid-based short-term wholesale market. In order to understand the potential market efficiency and system reliability benefits of a bid-based market for Brazil, I then presented several comparisons of the performance of the Brazilian short-term market with the short-term markets in hydroelectric-dominated industries with short-term bid-based markets in Colombia, New Zealand, and Norway. My interpretation of the results of these market performance comparisons is that there are significant market efficiency benefits associated with Brazil adopting a bid-based short-term market.

The report then described the initial conditions necessary to implement a bid-based short-term market in Brazil. These necessary conditions are: (1) coverage of close to 100% of final demand in fixed-price forward contract obligations negotiated far enough in advance of delivery to allow new entrants to compete to supply these contracts, (2) a local market power mitigation mechanism that applies to all market participants, (3) a cap and floor on supply offers into the short-term wholesale market, and (4) a prospective market monitoring process with public release of all data necessary to operate the short-term market. Because the Brazilian market currently requires 100% contracting by final demand, the major changes necessary to implement this necessary condition for a bid-based market is ensuring that a sufficient fraction of these long-term contracts are signed far enough in advance of delivery to allow new entrants to compete to supply this energy. A fixed-price forward contract obligation without a requirement to purchase a substantial fraction of these obligations far enough in advance of delivery to allow new entrants to compete with existing suppliers will provide no short-term market power mitigation benefits.

The report then presents a recommended bid-based short-term market design and suggests a transition process from the current cost-based market design to this market design. Although I believe that this transition process should take between 12 to 18 months to complete, this process should not be adhered to without regard to events in the short-term market. In particular, further moves towards introducing flexible market mechanisms should not be made without the appropriate safeguards against the exercise of unilateral market power in place and verification that they are working as intended.

References and Other Relevant Background Documents

Borenstein, Severin, Bushnell, James, and Wolak, Frank (2002) "Measuring Market Inefficiencies in California's Restructured Electricity Supply Industry, American Economic Review, 1376-1405.

Coase, Ronald H. "The Problem of Social Cost." Journal of Law and Economics. 3 (1960).

Von der Fehr, Nils-Henrik M. and Wolak, Frank A. (2003) "Power Sector Reform in Brazil—Some Issues (*)

Wolak, Frank A. (2000a) "An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market," International Economic Journal, Summer, 1-40.(*)

Wolak, Frank A. (2001) "Designing a Competitive Electricity Market that Benefits Consumers, November (*).

Wolak, Frank A. (2002) "Competition-Enhancing Local Market Power Mitigation in Wholesale Electricity Markets" November. (*).

Wolak, Frank A. (2003a) "Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market 1998 to 2000," American Economic Review, May 2003, 425-430. (*)

Wolak, Frank A. (2003) "Designing Competitive Wholesale Electricity Market for Latin American Countries." (*)

Wolak, Frank A. (2004) "Managing Unilateral Market Power in Wholesale Electricity," in The Pros and Cons of Antitrust in Deregulated Markets, edited by Mats Bergman, Swedish Competition Authority. (*)

(*) Available at http:/www.stanford.edu/~wolak)

Figure 1: Two-Node Example of Local Market Power Problem

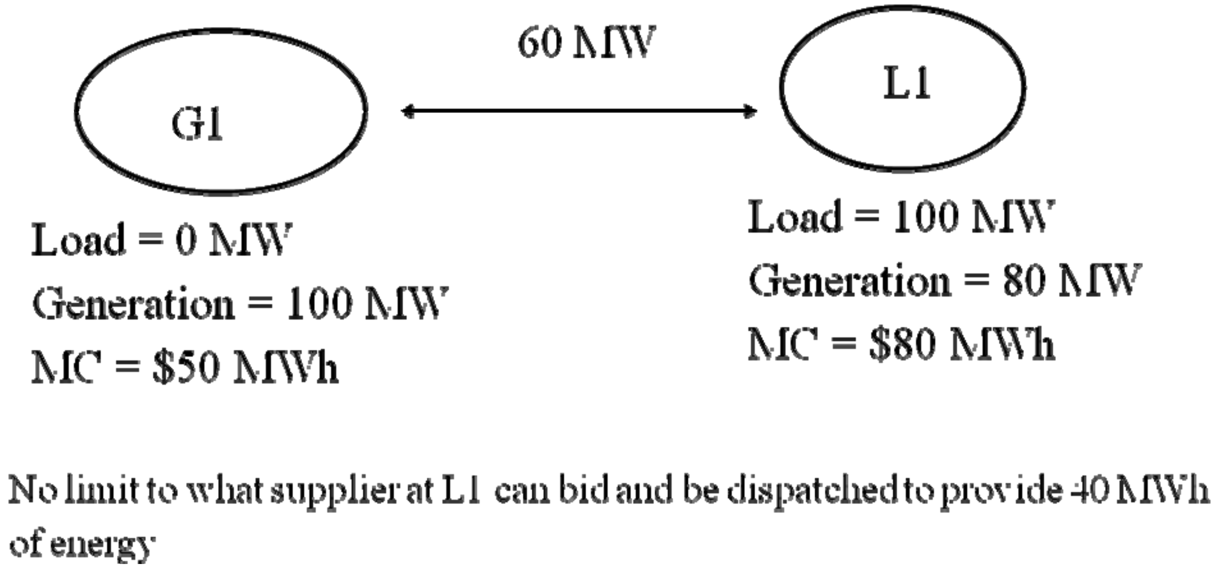

Figure 2: Annual Histograms of Natural Logarithm of Weekly Prices for Brazil

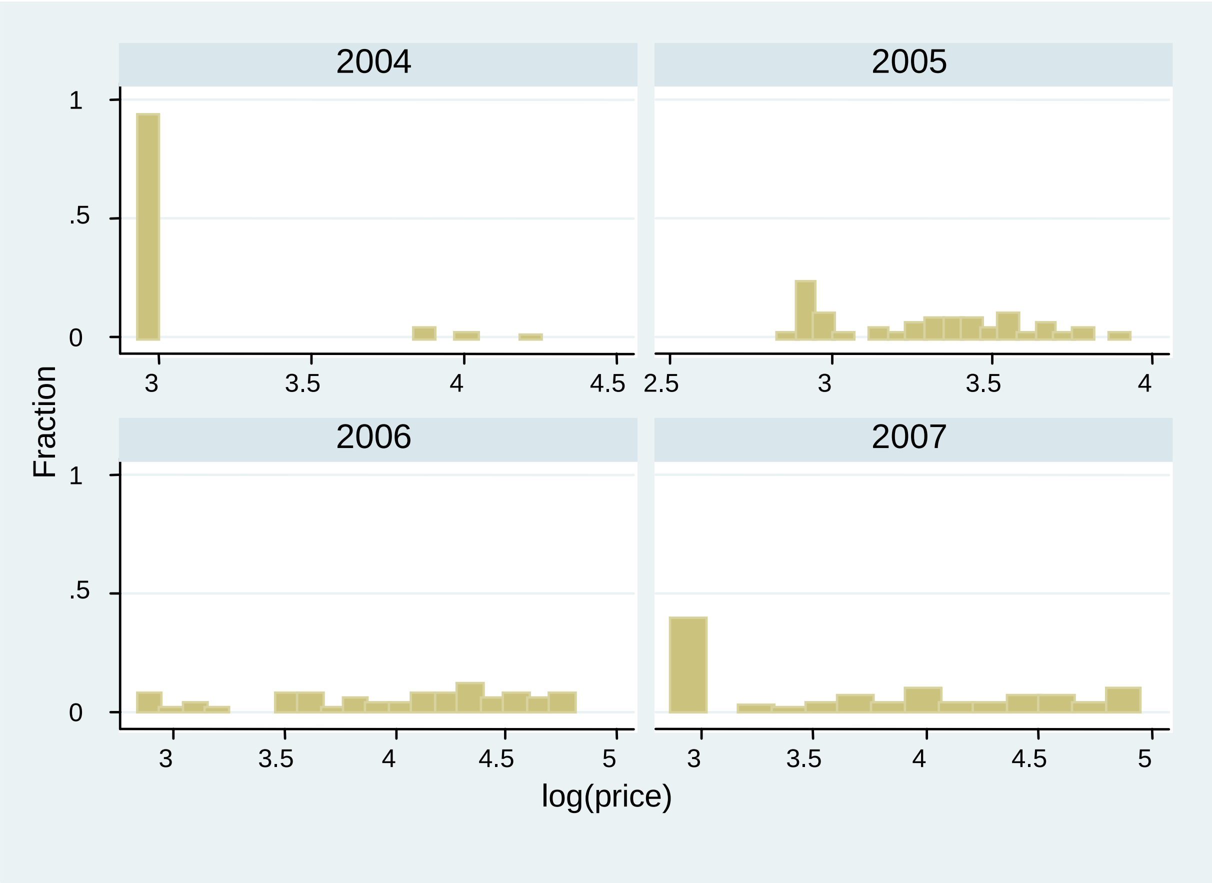

Figure 3: Annual Histograms of Natural Logarithm of Daily Average Nodal Price for New Zealand

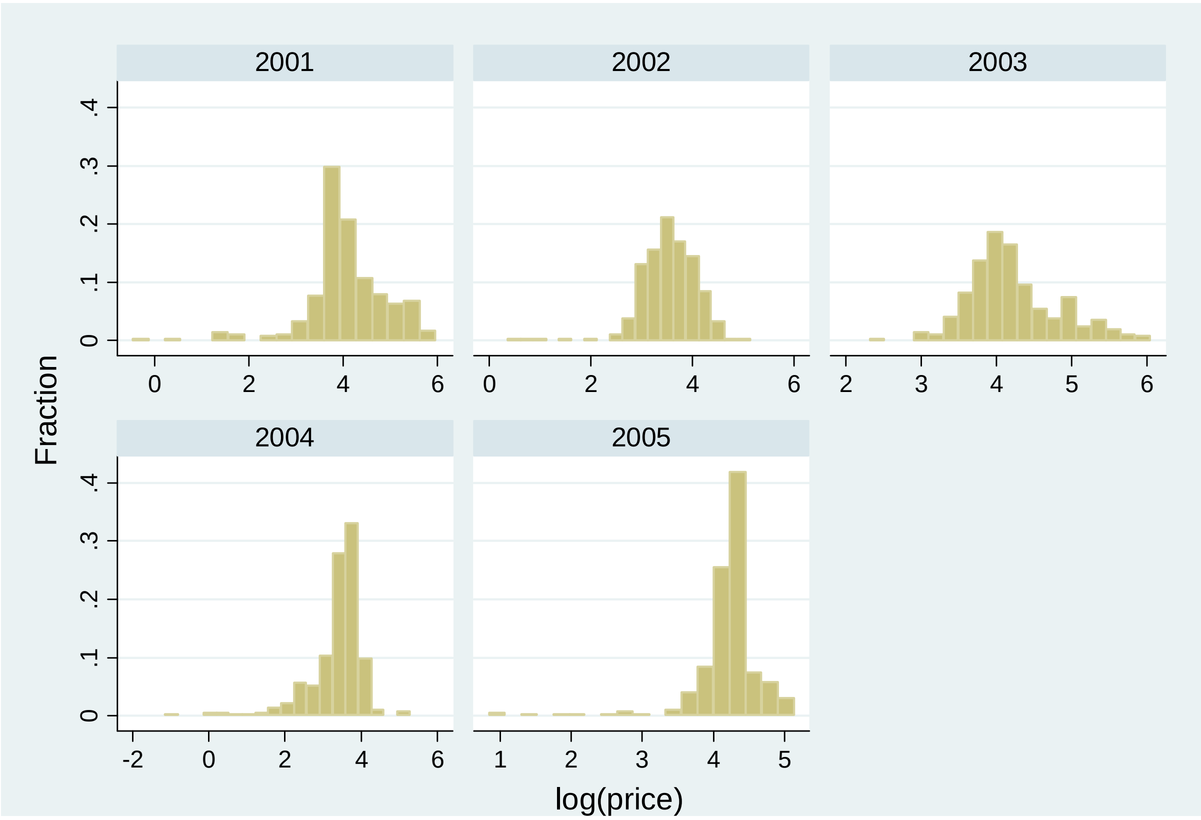

Figure 4: Annual Histograms of the Natural Logarithm of Daily Average Prices for Colombia

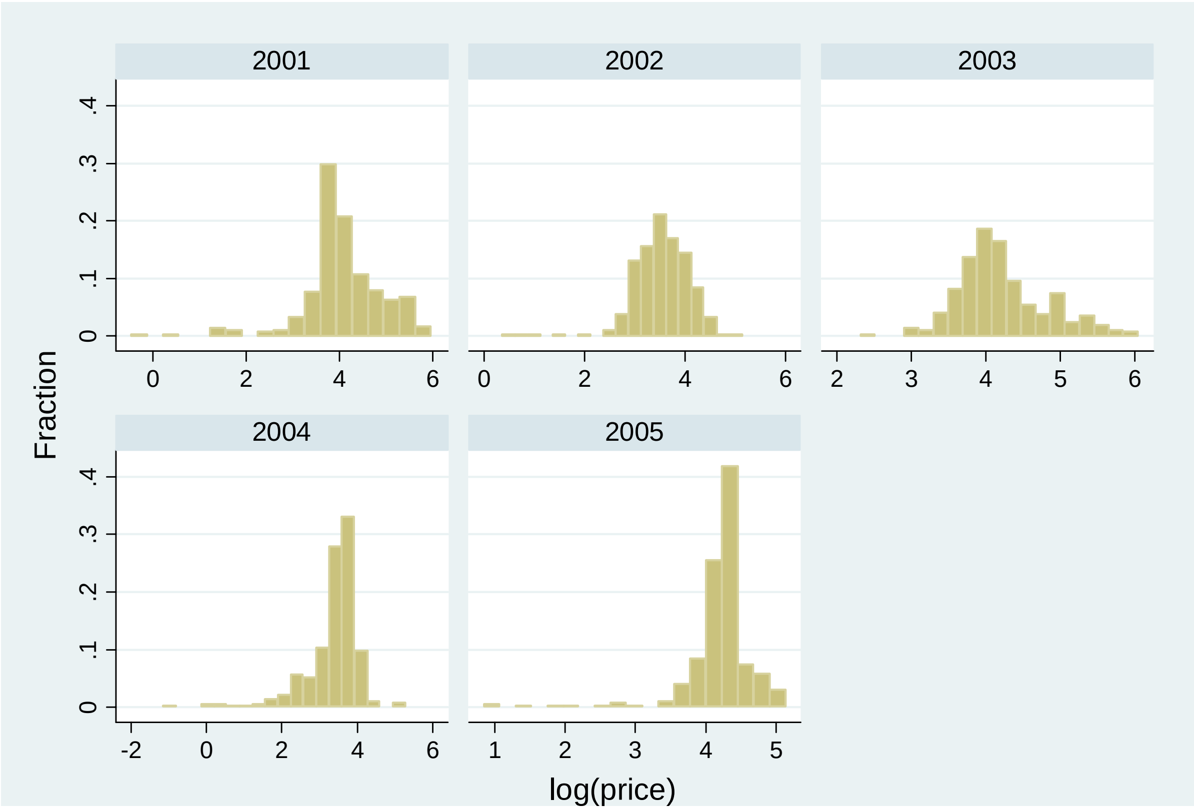

Figure 5: Annual Histograms of Natural Logarithm of Daily Average Prices for Norway

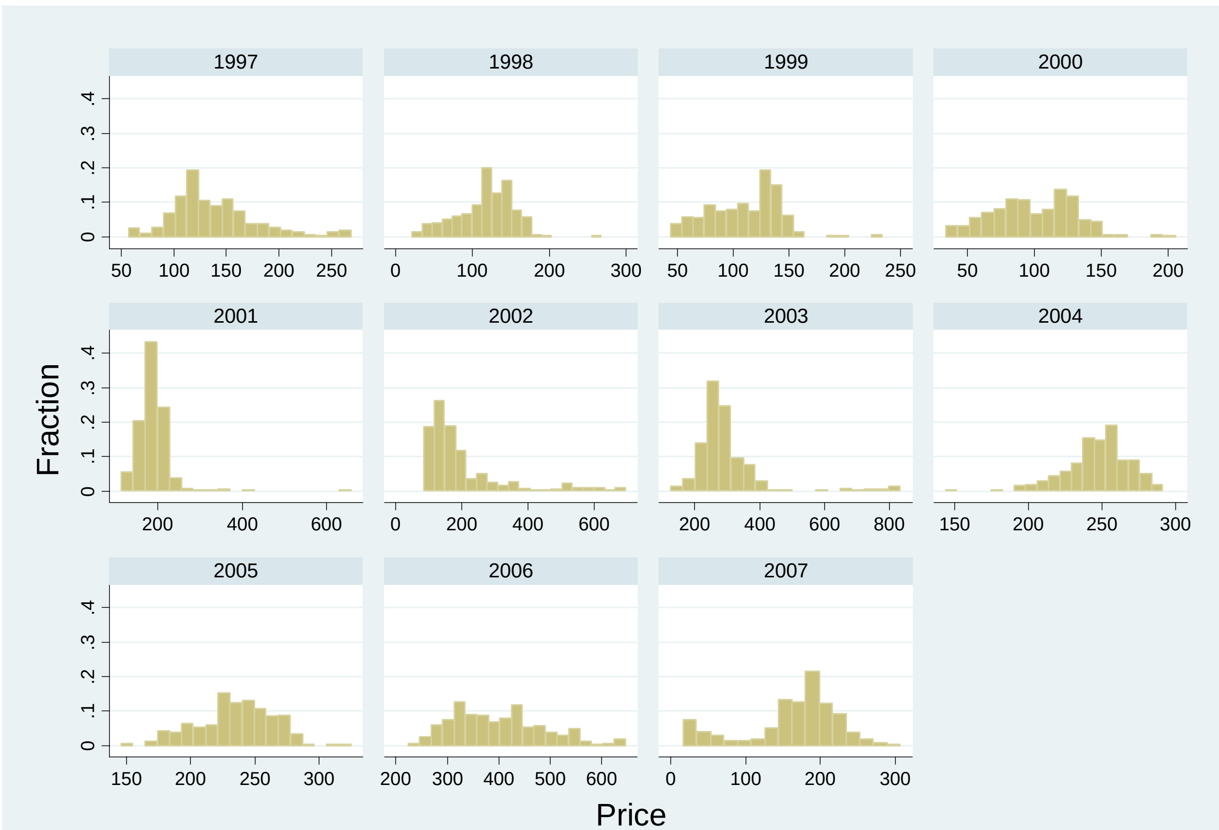

Figure 6: Annual Relationship Between Daily Water Level (Fraction of [Sample Maximum - Sample Minimum]) and Natural Logarithm of Daily Average Price for Brazil

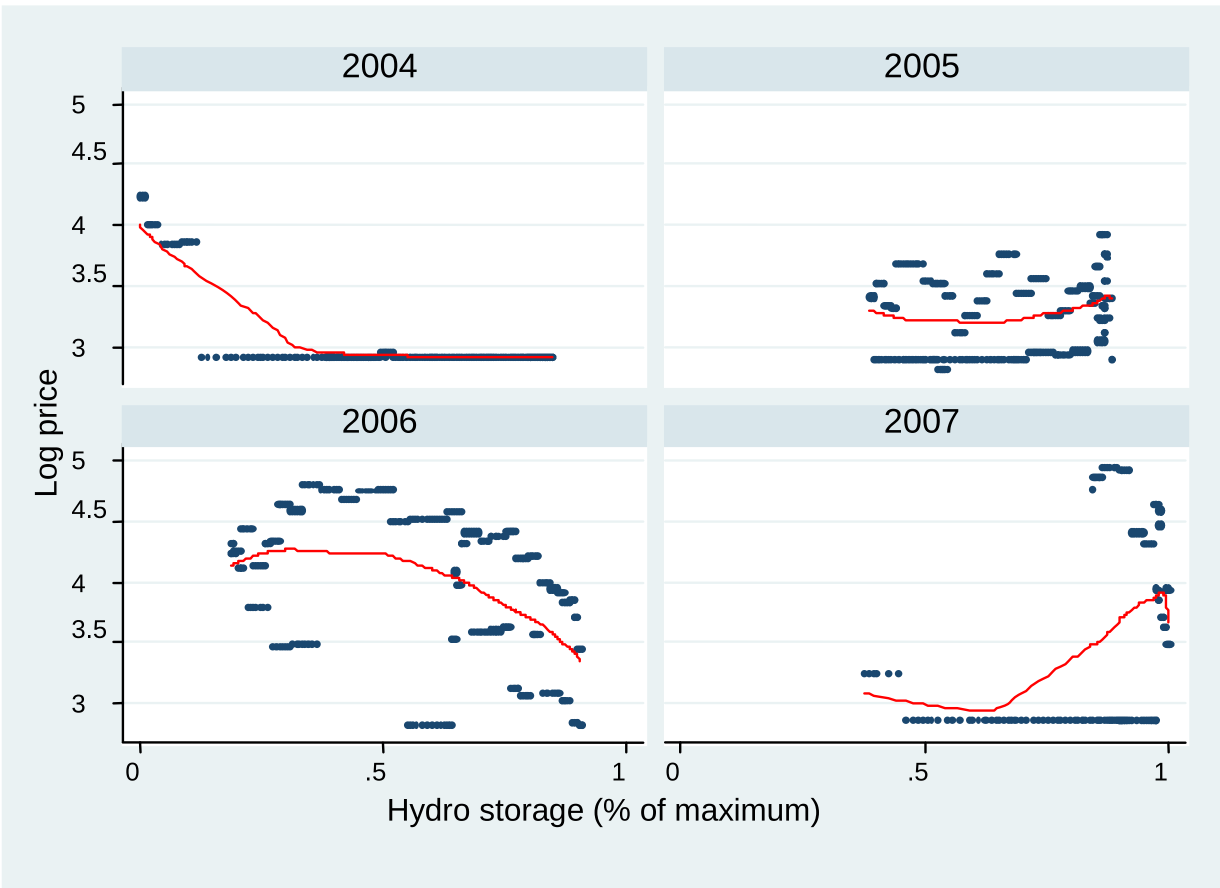

Figure 7: Annual Relationship Between Daily Water Level (Fraction of [Sample Maximum - Sample Minimum]) and Natural Logarithm of Daily Average Price for New Zealand

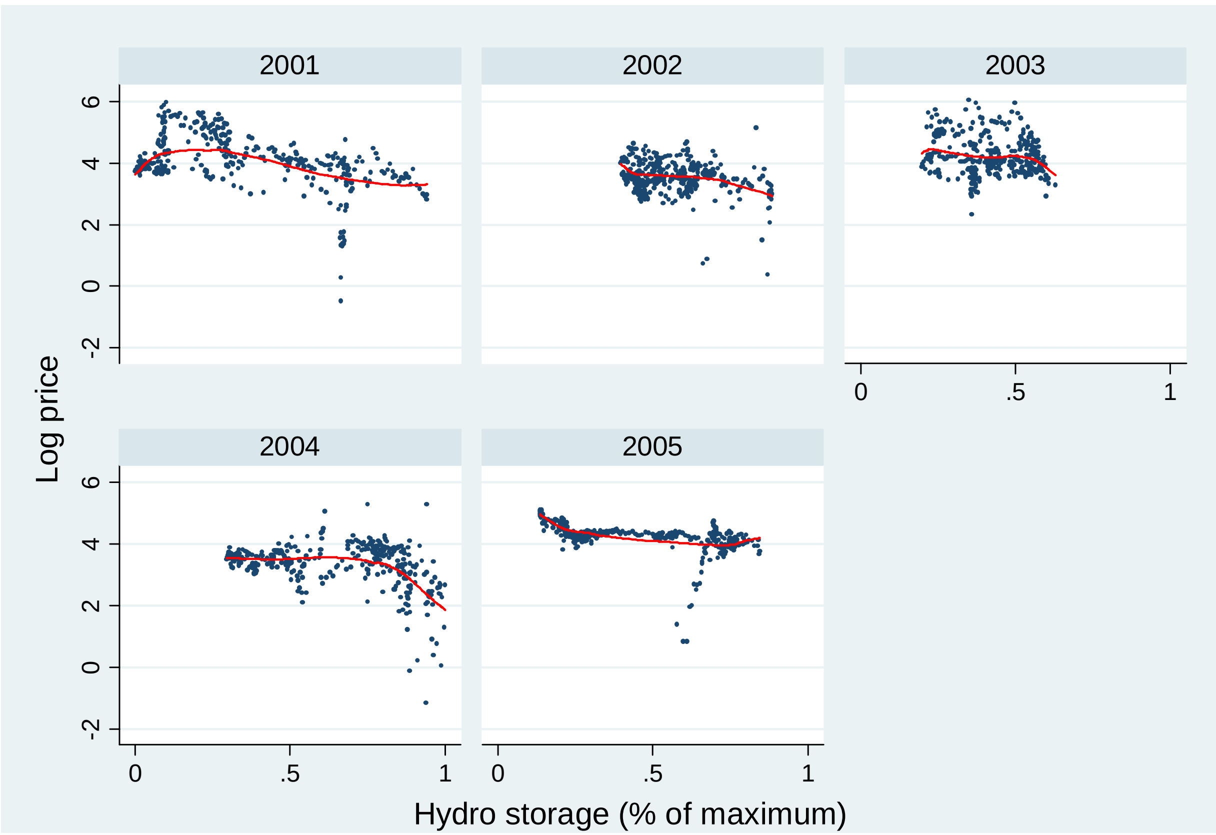

Figure 8: Annual Relationship Between Daily Water Level (Fraction of [Sample Maximum - Sample Minimum]) and Natural Logarithm of Daily Average Price for Colombia

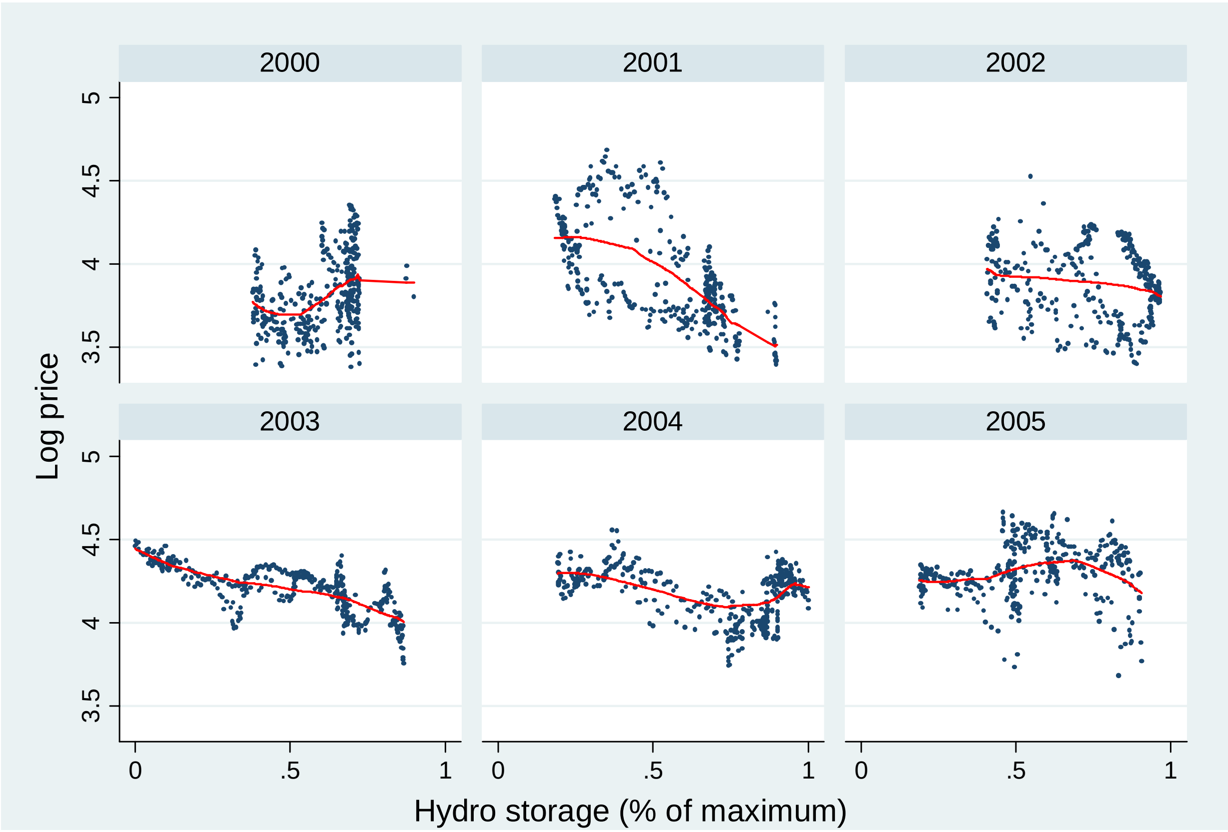

Figure 9: Annual Relationship Between Weekly Water Level (Fraction of [Sample Maximum - Sample Minimum]) and Natural Logarithm of Daily Average Price for Norway

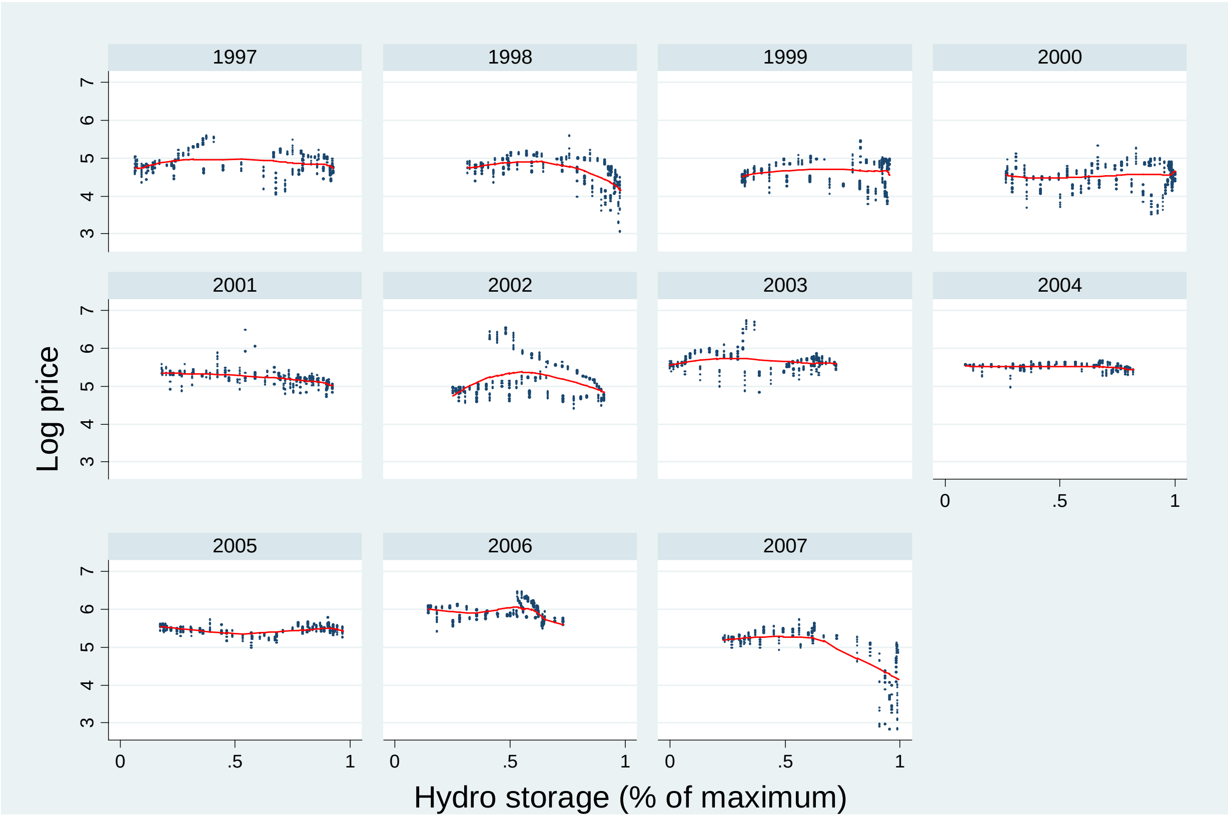

Figure 10: Cost of Deficit Function for Brazilian Market

% Load Reduction (Depth)		Cost of Deficit (R$/MWh)
0 to 5%	→	944.51
5% to 10%	→	2,037.61
10% to 20%	→	4,257.97
Higher than20%	→	4,838.69

Figure 11: Daily Average Wholesale Prices in New Zealand

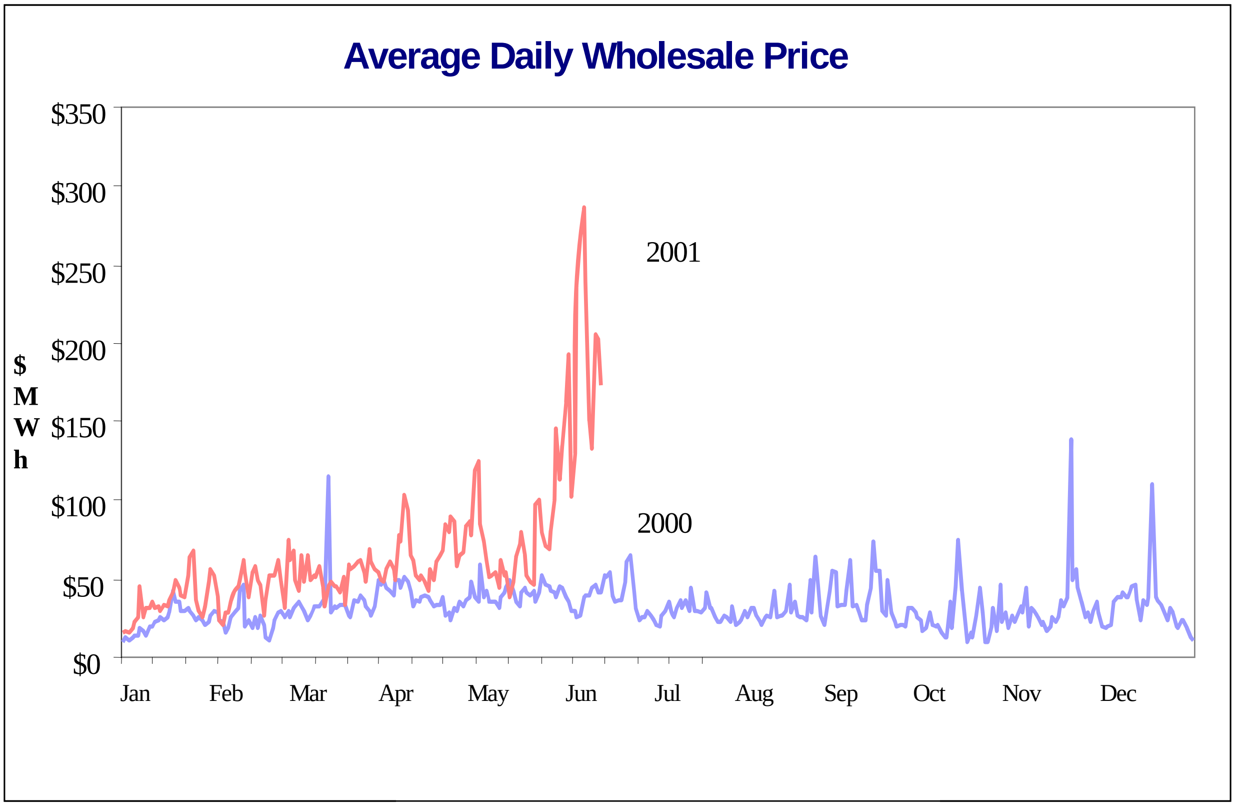

Figure 12: Daily Hydro Storage Throughout the Year