Transformation and Modernization of the Wholesale Electricity Market in Colombia

by

Frank A. Wolak
Director, Program on Energy and Sustainable Development (PESD)
Holbrook Working Professor of Commodity Price Studies
Department of Economics
Stanford University
Stanford, CA 94305-6072
wolak@zia.stanford.edu

Current Draft: August 28, 2019

1. Introduction

This report provides recommendations on the six topic areas in the transformation and modernization theme “Competition, participation and structure of the electricity market.” These are: (1) investment, reliability charges, and contracts; (2) generation diversification, of Non-Conventional Renewable Energy Sources (NCRES) and greater number of agents; (3) new services and agents: storage systems and aggregators; (4) restrictions, nodal prices and infrastructure; (5) market structure; and (6) pathways to de-carbonization and implications for market design. These recommendations are aimed at enhancing the efficiency of the short-term electricity market design and the long-term resource adequacy process in Colombia. They also provide policy pathways for the government of Colombia to support the deployment of NCRES, the entry of new market participants and technologies, and the active participation of final consumers in the wholesale market in manner that increase the competitiveness of wholesale and retail market outcomes.

1.1 Short-Term Market Design

The experience of the past thirty years with restructured electricity supply industries has led to a general consensus among scholars on the most efficient short-term market design. This was not always case, particularly in the United States (US), where there were significant differences between the initial short-term market designs implemented in the late 1990s and early 2000s by the California Independent System Operator (ISO), PJM Interconnection, ISO-New England, New York ISO, and Electricity Reliability Council of Texas (ERCOT) wholesale electricity markets. As I discuss in more detail below, the economic and reliability challenges faced by the initial market designs in these regions were the result of a common set of short-term market design flaws, many of which are exacerbated by an increasing share of intermittent renewable generation resources.

Starting with PJM in 1998 and ending with ERCOT in late 2010, all US market designs have adopted a multi-settlement locational marginal pricing (LMP) market design that co-optimizes the procurement of energy and ancillary services. This market design has a day-ahead financial market which jointly clears the markets for energy and each ancillary service simultaneously for all 24 hours of the following day. A real-time market then operates using the same network model as the day-ahead market adjusted to real-time system conditions. Deviations from purchases and sales in the day-ahead market are cleared using these real-time prices. Both of these markets price all relevant transmission network and other relevant operating constraints on the transmission network and generation units.

Because most European countries began the wholesale market regime with significantly more extensive and modern transmission networks than the regions in the US with wholesale markets, similar magnitudes of economic and reliability challenges did not arise until recently, even though most current European market designs share many features with the initial US market designs. Recently, the increasing amounts of intermittent renewable generation capacity in many European markets have made these same market design flaws more apparent and consequential. So much so that a number of European markets are now considering implementing the consensus market design from the US—a multi-settlement LMP market that co-optimizes energy and ancillary services.

The Colombian short-term market design shares many features with the early US market designs and existing European market designs, and for that reason, it faces similar economic and reliability challenges. If Colombia significantly increases the share of intermittent renewable generation in its electricity supply industry the magnitude of these economic and reliability challenges is also likely to increase. A multi-settlement LMP market design in Colombia would support the least cost operation of the current mix of generation capacity as well as a resource mix that contains a significantly larger share of intermittent renewable generation units.

This reports summarizes the underlying causes of the current economic and reliability challenges facing the Colombian short-term market and how they are likely to change in the future. The basic features of a multi-settlement locational marginal pricing market is described and reasons it should improve short-term market performance explained. A recommended process for implementing this market design in Colombia is provided.

1.2 Long-Term Resource Adequacy

There is considerably less agreement among scholars on the efficient mechanism for ensuring long-term resource adequacy for a wholesale electricity market, particularly in regions with significant hydroelectric and intermittent renewable resources. However, as I explain below, a long-term resource adequacy mechanism is necessary for all current wholesale markets to address what I call the "reliability externality." I then survey the two dominant approaches to ensuring long-resource adequacy—long-term energy contracts and capacity payment markets/

mechanisms—and explain why a capacity-based approach makes very little sense for a hydroelectric energy dominated industry like Colombia. The most recent El Nino event in Colombia described in McRae and Wolak (2019) provides a real-world example of the unsuitability of the current reliability payment mechanism in Colombia for ensuring long-term reliability in a hydroelectric-dominated system like Colombia. A capacity payment mechanism is even less suitable for a wholesale market with a significant fraction of intermittent renewable resources, where the major reliability challenge is no longer sufficient installed capacity to meet system demand peaks, but inadequate energy to meet demand during all possible future system conditions.

The "reliability externality'' that justifies the need for a long-term resource adequacy mechanism arises from the interaction of the political economy of wholesale electricity pricing with the current mechanisms for managing real-time, system-wide supply shortfalls. This reliability externality is shown to give rise to amissing market for long-term contracts for energy at delivery horizons long enough in the future to ensure an adequacy supply of energy under all possible system conditions. Understanding the fundamental cause of the "reliability externality" suggests a solution that involves the minimal interference in market mechanisms yet still internalizes this reliability externality.

I argue that mandated standardized forward contracts for energy are the minimal regulatory intervention approach to ensure long-term resource adequacy. This mechanism requires all retailers and free consumers to purchase hourly-system-load-shape-weighted standardized long-term contracts for energy equal to pre-specified fractions of their eventual demand for energy at various future delivery horizons sufficiently far into the future to ensure there is adequate energy to meet demand under all possible future system conditions. This mechanism does not rule out market participants from engaging in other forward market transactions and short-term price hedging arrangements, only that all retailers and large loads must buy and hold these standardized contracts to delivery. A mechanism for transitioning from the existing reliability payment mechanism to this mechanism is then proposed as well as other regulatory changes needed to implement this mechanism.

1.3. Policies to Support NCRES and Facilitate More Competitive Market Outcomes

A final set of issues concern policies to support the deployment of NCRES generation units and policies to support the entry of other market participants, new technologies and services, as well as active participation of final consumers in the wholesale market. A key concern for supporting the deployment of NCRES generation units is doing so in a manner that minimally interferes with the operation of the short-term energy market and long-term resource adequacy process. The same concern applies to encouraging the entry of new technologies and services and market participants. A multi-settlement LMP market design and standardized energy contracting long-term resource adequancy process encourages the cost-effective deployment of these technologies and services. Finally, I emphasize the crucial role of investments in minimal levels of modern metering infrastructure to encouraging the entry of these new technologies and market participants, particularly the active participation of final demand in the wholesale market.

2. Modernizing the Short-Term Market Design in Colombia

An important lesson from the experience of short-term electricity markets around the world is that a crucial determinant of the efficiency of market outcomes is the extent to which the financial market mechanism used to set dispatch levels and prices is consistent with how the grid is actually operated. In the early stages of wholesale market design in the US, all of the regions attempted to operate short-term markets that used simplified versions of the transmission network. These single-zone or zonal markets assumed infinite transmission capacity between locations in the transmission grid or only recognized transmission constraints between large geographic regions. These simplifications of the transmission network configuration and other relevant operating constraints created opportunities for market participants to increase their profits by taking advantage of the fact that in real time the actual configuration of the transmission network and other operating constraints must be respected.

These markets set a single market-clearing price for a half-hour or hour for an entire country or large geographic region despite the fact that there were generation units with offer prices below the market-clearing price not producing electricity and units with offer prices above the market-clearing price producing electricity. This outcome occurs because of the location of demand and available generation units within the region and the configuration of the transmission network prevents some of these low-offer-price units from producing electricity and requires some of the high-offer-price units to supply electricity. The former units are typically called "constrained-off" units and the latter are called "constrained-on" or "must-run" units. In Colombia, constrained-on units are deemed to provide "positive reconciliations" and constrained-off units provide "negative reconcilations."

A market design challenge arises from this positive and negative reconciliation mechanism because how generation units are compensated for being constrained on or constrained off impacts the offer prices they submit into wholesale energy market. For example, if generation units are paid their offer price for electricity when they are constrained-on and the unit's owner knows that it will be constrained-on, a profit-maximizing unit owner will submit an offer price higher than the variable cost of operating the unit and be paid that price for the incremental energy it supplies, which raises the total cost of electricity supplied to final consumers.

A similar set of circumstances can arise for constrained-off generation units. Constrained-off generation units are usually paid the difference between the market-clearing price and their offer price times the quantity of energy it is unable to supply because of the configuration of the transmission network. This market rule creates an incentive for a profit-maximizing supplier that knows its unit will be constrained off to submit the lowest possible offer price in order to receive the highest possible payment for being constrained-off and raise the total cost of electricity supplied to final consumers. Chapter 8 of McRae and Wolak (2016) documents that these positive and negative reconciliation payments each averaged roughly 20 percent of the total short-term market revenues throughout sample period they studied, and during certain months averaged as much 40 percent of short-term market revenues.

The experience of many European markets indicates that the cost of making the final schedules that emerge from the current single-zone market in Colombia physically feasible is likely to get even larger as the amount of intermittent generation units in Colombia increases. A number of European markets have found that the cost of making final schedules that emerge from their single-zone (Germany, Great Britain and Spain) and zonal (Italy) markets physically feasible has grown significantly as their share of intermittent renewable energy has grown. In 2017 in Germany these costs were over 1 billion Euros, in Great Britain they were more than 400 million Euros, in Spain they were over 80 million Euros, and in Italy they were approximately 50 million Euros.[1] As mentioned earlier, these costs have led number of European countries to consider adopting more granular approaches to pricing.

[1] See Figure of 90 of ENTSO-E (2018).

2.1. *Locational Marginal Pricing (LMP)*

Almost any divergence between the market model and generation unit operating constraints (ramp rates, minimum load levels, minimum runtimes, etc.) used to set dispatch levels and market prices and the ones used to operate the transmission network and generation units creates an opportunity for market participants to take actions that raise their profits at the expense of overall market efficiency. Multi-settlement wholesale electricity markets that use locational marginal pricing (LMP), also referred to as nodal pricing, largely avoid these constrained-on and constrained-off problems, because all transmission constraints and other relevant operating constraints are respected in the process of determining dispatch levels and prices in the day-ahead market. This means that no generation units are scheduled to produce energy or supply an ancillary service unless it is expected to be physically capable of the doing so given the actual configuration of the transmission network and operating constraints of all generation units in the control area.

All LMP markets in the US co-optimize the procurement of energy and ancillary services.[2] This means that all suppliers submit generation unit-specific willingness-to-supply schedules for energy and any ancillary service the generation unit is capable of providing. Large loads and load-serving entities submit their willingness-to-purchase energy schedules to the wholesale market operator. Locational prices for energy and ancillary services and dispatch levels and ancillary services commitments for generation units at each location in the transmission network are determined by minimizing the as-offered costs of meeting the demand for energy and ancillary services at all locations in the transmission network subject to all network and other relevant operating constraints. No generation unit will be accepted to supply energy or an ancillary service if doing so would violate a transmission or other operating constraint.

The Colombian market design is currently undergoing a reform to its ancillary services markets. Co-optimizing the centralized procurement of energy and ancillary services in the day-ahead and real-time markets could yield significant cost savings to Colombian consumers. LMP markets that co-optimize the procurement of energy and ancillary services ensure that each generation unit is used in the most cost-effective manner based on the energy and ancillary services offers of all generation units in the wholesale market. In addition, the opportunity cost of supplying

[2] Ancillary services are composed of different operating reserves required by the system operator to maintain real-time supply and demand balance. For the California ISO market, these are composed of Regulation Up and Regulation Down Automatic Generation Control (AGC) reserves, Spinning Reserves and Non-Spinning Reserve.

any ancillary service a generation unit is capable of providing will be explicitly taken into account in deciding whether to use the generation unit to provide that ancillary service versus energy or other ancillary service the unit is capable of providing.

For example, if the day-ahead market-clearing price of energy at that generation unit's location is $40/MWh and the unit's offer price for energy is $30/MWh and the unit's offer price for the only ancillary service the unit can supply is $5/MW, the unit will be accepted to supply the ancillary service only if the market-clearing price is greater than or equal to $10/MW, because of the $10/MWh opportunity cost of energy for that unit. In contrast, sequential procurement of ancillary services, after or before the energy market has cleared increases the ability of suppliers to exercise unilateral market power in the subsequent market because generation unit owners know which units were taken in the previous market. At best, sequential procurement of ancillary services and energy relies on individual market participants to make the efficient choice between supplying energy or ancillary services from each generation unit. Co-optimized procurement of energy and ancillary services through an LMP market mechanism ensures that the system-wide least cost solution is chosen based on the offers submitted by all market participants.

The nodal price at each location is the increase in the minimized value of the "as-offered costs" objective function as a result of a one unit increase in the amount of energy withdrawn at that location in the transmission network. The price of each ancillary service is defined at the increase in the optimized value of the objective function as a result of a one unit increase in the demand for that ancillary service. In most LMP markets ancillary services are procured at a coarser level of spatial granularity than energy.

The LMP pricing process sets potentially different prices at all locations in the transmission network, depending the configuration of the transmission network and geographic location of demand and the availability of generation units. Because the configuration of the transmission network and the location of generation units and demands is taken into account in operating the market, only generation unit dispatch levels that can actually be produced in real-time will be accepted to serve demand and they will be paid a higher price or lower LMP than other units, depending whether the generation unit is in a generation-deficient or generation-rich region of the transmission network. Bohn, Caramanis, and Schweppe (1984) provide an accessible discussion of the properties of the LMP market mechanism.

Another strength of the LMP market design is the fact that other constraints that the system operator takes into account in operating the transmission network can also be accounted for in setting dispatch levels and locational prices. For example, suppose that reliability studies have shown that a minimum amount of energy must be produced by a group generation units located in a small region of the grid. This operating constraint can be built into the LMP market mechanism and can be reflected in the resulting LMPs. This property of the LMP markets is particularly relevant to the cost-effective integration a significant amount of intermittent renewable generation capacity in the transmission network. Additional reliability constraints may need to be formulated and incorporated into LMP market to account for the fact that this energy supply can quickly disappear and re-appear.

An important lesson from the US experience with LMP markets is that explicitly accounting for the configuration of the transmission network in determining dispatch levels both within and across regions can significantly increase the amount of trade that takes place between regions. Mansur and White (2012) dramatically demonstrate this point by comparing the trades between regions of the eastern US before and after these regions are integrated into a single locational marginal pricing market that accounts for the configuration of the transmission network throughout the entire integrated region. Hourly energy flows between the two regions increased by almost 1,000 MWh immediately following the integration of the two regions into an LMP market. There was no change in the physical configuration of the transmission network for the two regions. This increase in flows was purely due the incorporating the two regions into a formal LMP market that recognizes the configuration of the transmissions network for the two regions in dispatching generation units and setting LMPs.

A number of US LMP markets with significant intermittent renewable resources have introduced fast-ramping ancillary services that hold out generation units with low energy offer prices that are able to quickly ramp up their production in order to have sufficient unloaded generation capacity to meet the steep ramps when solar energy disappears at the end of the daylight hours. The LMP market design makes it straightforward to introduce new products into market mechanism or increase or decrease the degree of spatial granularity in the products demanded. Additional variables are introduced into the objective function used to solve for market outcomes and additional constraints are introduced to ensure that supply equals demand for this new product. The market price is equal to the increase in the optimized value of the objective function associated with increasing the demand for this product by one unit.

2.2. Multi-Settlement Markets

Colombia currently operates a single-settlement, single zone market where individual generation unit owners unilaterally make their commitment decisions. This market design makes no attempt to find the least cost mix of generation units to meet demand throughout the entire day. Consequently, the current Colombian market design leaves considerable room to reduce the cost of serving demand at all locations in the country and improve overall system reliability. A multi-settlement LMP market offers the potential for significant wholesale energy cost savings and system reliability improvements relative to the current short-term market design in Colombia.

Multi-settlement nodal-pricing markets have been adopted by all US jurisdictions with an offer-based short-term wholesale electricity market. A multi-settlement market has a day-ahead forward market that is run one day ahead real-time system operation. This market sets firm financial schedules for all generation units and loads for all 24 hours of the following day. Suppliers submit generation unit-level offer curves for each hour of the following day and electricity retailers submit demand curves for each hour of the following day. The system operator then minimizes the as-offered cost to meet these demands simultaneously for all 24 hours of the following day subject to the anticipated configuration of the transmission network and other relevant transmission network and generation unit operating constraints during all 24 hours of the following day. This gives rise to LMPs and firm financial commitments to buy and sell electricity each hour of the following day for all generation unit and load locations.

The day-ahead market typically allows generation unit owners to submit three part offers to supply energy--start-up costs, no load costs, and energy offer curves. These costs enter the objective function used to compute hourly generation schedules and ancillary service quantities and locational marginal prices for energy and ancillary services for all 24 hours of the following day. This logic implies that a generation unit will not be dispatched in the day-ahead market to provide energy or any ancillary services unless the combination of its offered start-up costs, no-load costs and energy production costs are part of the least cost solution to serving hourly demands for all 24 hours of the following day.

Although this approach to determining energy schedules would represent a significant change from the current Colombia market design where suppliers make their own commitment decisions, it is not as a large of a change as it first might appear for the following reasons. First, all US multi-settlement LMP markets do not prohibit generation unit owners from making their own-unilateral commitment decisions. Instead, these markets rely on the financial incentive created by the market rule that generation units committed through the day-ahead market are guaranteed recovery of their as-offered start-up, no-load, and energy production costs for the following day, whereas generation units that self-commit are not. This mechanism makes it relatively more attractive for generation units to participate in the day-ahead market, but it does not prohibit self-commitment.

This guarantee of as-offered cost recovery from participation in the day-ahead market works as follows. Suppose a generation unit owner submits a start-up cost offer of $10,000, no load cost offer of $5,000 and an energy offer of $20/MWh and it is accepted to supply a total of 1,000 MWh energy and no ancillary services in the day-ahead market at an hourly quantity-weighted average energy price of $30/MWh. This means the unit owner's total energy and ancillary services market revenues for the day are $30,000 = $30/MWh x 1,000 MWh, but its total as-offered cost of producing this energy given that the generation unit needed to be turned on to provide this energy is equal $35,000 = $10,000 + $5,000 + $20/MWh x 1,000 MWh. In this case the unit owner will receive what is typically called a "make-whole payment" of $5,000 to recover its as-offered costs.

One advantage of this as-offered cost recovery promise to generation unit owners committed in the day-ahead market is that it eliminates any incentive a supplier with no ability to exercise unilateral market has to submit an offer price above the unit's marginal cost of producing energy, because the unit owner knows that its fixed-cost of starting up and operating will be recovered if it is committed in the day-ahead market.

To the extent that suppliers submit start-up, no-load, and energy offer curves that are representative of their actual costs, the cost of committing and dispatching the generation units that arise from this centralized unit commitment process are likely to be less than the commitment and dispatch costs that result from a self-commitment market, such as the one that currently exists in Colombia and number of European countries. Consequently, although there no prohibition against self-commitment by generation unit owners in an LMP market, this behavior unlikely to be consistently profitable relative to participation in and commitment through the day-ahead market.

The energy schedules that arise from the day-ahead market do not require a generation unit to supply the amount sold or a load to consume the amount purchased in the day-ahead market. The only requirement is that any shortfall in a day-ahead commitment to supply energy must be purchased from the real-time market at that same location or any production greater than the day-ahead commitment is sold at the real-time price at that same location. For loads, the same logic applies. Additional consumption beyond the load's day-ahead purchase is paid for at the real-time price at that location and the surplus of a day-ahead purchase relative to actual consumption is sold at the real-time price at that location.

In all US wholesale markets, real-time LMPs are determined from the real-time offer curves from all available generation units and dispatchable loads by minimizing the as-offered cost to meet the real-time demand at all locations in the control area taking into account the current configuration of the transmission network and other relevant operating constraints. This process gives rise to LMPs at all locations in the transmission network and actual hourly operating levels for all generation units. Real-time imbalances relative to day-ahead schedules are cleared at these real-time prices.

To understand how a two-settlement market works, suppose that a generation unit owner sells 50 MWh in the day-ahead market at \$60/MWh. It receives a guaranteed \$3,000 in revenues from this sale. However, if the generation unit owner fails to inject 50 MWh of energy into grid during that hour of the following day, it must purchase the energy it fails to inject at the real-time price at that location. Suppose that the real-time price at that location is \$70/MWh and generator only injects 40/MWh of energy during the hour in question. In this case, the unit owner must purchase the 10 MWh shortfall relative to its day-ahead schedule at \$70/MWh. Consequently, the net revenues the generation unit owner earns from selling 50 MWh in the day-ahead market and only injecting 40/MWh is \$2,300, the \$3,000 of revenues earned in the day-ahead market less the \$700 paid for the 10 MWh real-time deviation from the unit's day-ahead schedule.

If a generation unit produces more output than its day-ahead schedule, then this incremental output is sold in the real-time market. For example, if the unit produced 55 MWh, then the additional 5 MWh beyond the unit owner's day-ahead schedule is sold at the real-time price. By the same logic, a load-serving entity that buys 100 MWh in the day-ahead market but only withdraws 90 MWh in real-time, sells the 10 MWh not consumed at the real-time price. Alternatively, if the load-serving entity consumes 110 MWh, then the additional 10 MWh not purchased in the day-ahead market must be purchased at the real-time price.

Wolak (2011b) quantifies the magnitude of the economic benefits associated with the transition to nodal pricing from a zonal-pricing market that was very similar to the standard market design in Europe and other industrialized countries. On April 1, 2009 the California market transitioned to a multi-settlement nodal-pricing market design from a multi-settlement zonal-pricing market. Wolak (2011b) compares the hourly conditional means of the total amount of input fossil fuel energy in millions of BTUs, the total hourly variable cost of production from fossil fuel generation units, and the total hourly number of starts from fossil fuel units before versus after the implementation of nodal pricing controlling non-parametrically for the total hourly output of the fossil fuel units in California and the daily prices of the major input fossil fuels. Total hourly BTUs of fossil fuel energy consumed to produce fossil fuel-fired electricity is 2.5 percent lower, the total hourly variable cost of production for fossil fuels units is 2.1 percent lower, and the total number of hourly starts is 0.17 higher after the implementation of nodal pricing. This 2.1 percent cost reduction implies a roughly $105 million reduction in the total annual variable cost of producing fossil fuel energy in California associated with the introduction of nodal pricing.

A multi-settlement LMP market design is also particularly well suited to managing a generation mix with a significant share of intermittent renewable resources. The additional operating constraints necessary for reliable system operation with an increased amount of renewable resources can easily be incorporated into the day-ahead and real-time market models. Therefore, the economic benefits from implementing a multi-settlement LMP market relative to market designs that do not model transmission and other operating constraint are likely to be greater the larger is the share of intermittent renewable resources. Consequently, any region with significant intermittent renewable energy goals should realize significant future economic and reliability benefits from implementing a multi-settlement LMP market.

Another strength of a multi-settlement LMP market is that it values the dispatchability of generation units even though it pays all resources at the same location in the grid the same price in the day-ahead and real-time markets. This property is particularly relevant to regions with intermittent renewable energy goals that have concerns that the thermal resources needed to provide energy when these renewable resources are not available now run at lower annual capacity factors and therefore may not receive sufficient revenues to remain financially viable.

The following examples illustrate this point and emphasize the fact that thermal resources earn revenue to recover their annual costs from three sources: (1) selling energy in the forward market that they don't subsequently produce, (2) selling energy in the forward market that they subsequently produce, and (3) selling incremental energy beyond their forward market position in short-term market. Suppose that a wind unit sells 50 MWh and a thermal resource sells 40 MWh in the day-ahead market at $30/MWh. If in real-time not as much wind energy is produced, the dispatchable thermal unit must make up the difference. Suppose that the wind unit produces only 30 MWh, so that the thermal unit must produce an additional 20 MWh. Because of this wind generation shortfall, the real-time price is now $60/MWh. Under this scenario, the wind unit is paid an average price of $10/MWh = (50 MWh x $30/MWh – 20 MWh x $60/MWh)/30 MWh for the 30 MWh it produces, whereas the dispatchable thermal unit is paid an average price of $40/MWh = (40 MWh x $30/MWh + 20 MWh x $60/MWh)/60 MWh for the 60 MWh it produces.

Similar logic applies to the case that the wind resource produces more than expected and the thermal resource reduces its output because the real-time price is lower than the day-ahead price because of the unexpectedly large amount of wind energy produced. For example, suppose the wind unit sells 30 MWh and the thermal resource sells 60 MWh in the day-ahead market at $30/MWh. However, in real-time there is significantly more wind, so the wind unit produces 50 MWh at a real-time price of $10/MWh. Because of this low real-time price the thermal resource decides that the least cost way to meet its 60 MWh forward market obligation is the produce 40 MWh from its own units and purchase the remaining 20 MWh from the real-time market. The average price received by the wind unit is $22/MWh = (30 MWh x 30/MWh + 20 MWh x 10 MWh)/50 MWh and the average price received the thermal unit is $40/MWh = (60 MWh x $30/MWh – 20 MWh x $10/MWh)/40 MWh. Despite paying the same prices for thermal and wind energy in the day-ahead and real-time markets, a multi-settlement market pays a higher average price to the dispatchable generation unit for the energy it provides during the same hour as the wind unit.

This example illustrates an important point about the financial viability of thermal resources in a wholesale market with a significant amount of hydroelectric or intermittent wind and solar resources. There are likely to be many hours of the year when a thermal resource owner that has sold a fixed-price, fixed-quantity forward contract to supply energy would find it more profitable to purchase this quantity of energy from either the day-ahead or real-time market rather that incur the expense of producing this energy from its own units. A straightforward way to ensure that the least-cost "make versus buy" decision is made by a thermal resource owner is for it to submit its marginal cost of producing energy as its offer into the day-ahead and real-time markets. The promise of full cost recovery if the unit is committed in the day-ahead market implies that this strategy is expected profit-maximizing if the unit owner has no ability to exercise unilateral market power.

One complaint often leveled against LMP markets is that they increase the likelihood of political backlash from consumers because prices paid for wholesale electricity can differ significantly across locations within the same geographic region. For example, customers in urban areas that primarily import electricity over congested transmission lines will pay more than customers located in generation-rich rural regions that export electricity to these regions. Because more customers live in urban areas than in rural regions, charging final consumers in urban areas a higher retail price to recover the LMP at their location may be politically challenging for the regulator to implement.

Most regions with LMP pricing have addressed this issue by charging all customers in a state, region, or utility service territory a weighted average of the LMPs at all load withdrawal points in the geographic region. In the above example, this implies charging the urban and rural customers the weighted average of the LMPs in the urban and rural areas, where the weight assigned to each price is the share of system load that is withdrawn at that location. Under this scheme, generation units continue to be paid the LMP at their location, but all loads pay a geographically aggregated hourly-price. For example, in Singapore all generation units are paid the LMP at their location, but all loads are charged the Uniform Singapore Electricity Price (USEP), which is the quantity-weighted average of the half-hourly LMPs for all generation nodes in Singapore. This approach to pricing load captures the reliability and operating efficiency benefits of an LMP market while addressing the equity concerns regulators often face with charging customers at different locations prices that reflect the configuration of the transmission network. Tangerås and Wolak (2018) present evidence that this market rule can result in more competitive behavior in the wholesale market by vertically-integrated suppliers with the ability exercise unilateral market power.

2.4. Managing and Mitigating Market Power

An important concern in any wholesale market design is the exercise of unilateral market power by large suppliers. Because most re-structuring processes started by either breaking-up large vertically-integrated state-owned utilities or a combining a number of vertically-integrated utility service territories into a single wholesale market, virtually all wholesale markets have a few large market participants which causes concern about the exercise of unilateral market power.

During the initial stages of industry re-structuring the typical approach to dealing with this concern was through market structure restrictions. Limitations would be placed on generation ownership shares within a geographic market. Concerns about buyer market power would be addressed by limitations on retail market shares.

Particularly in the US, these market structure approaches to limiting the exercise of unilateral market power have been largely abandoned because it is extremely difficult, if not impossible, to set market structure restrictions that limit the ability of suppliers or buyers to exercise unilateral market power. As discussed below, there are so many factors that go into the enabling the exercise of unilateral market power in a wholesale electricity and these factors can change on an hourly basis. Consequently, simple market concentraction measures are largely ineffective at limiting all but the most egregious instances of the exercise of unilateral market power.

This outcome has led to the adoption of two approaches to addressing the exercise of unilateral market power in wholesale electricity markets: (1) an automatic local market power mitigation mechanism in the short-term market and (2) sufficient coverage of final demand with fixed-price, fixed-quantity forward contracts purchased far enough in advance of delivery to allow new entrants to compete with existing firms to supply these contracts.

2.4.1. Solutions to the Local Market Power Problem

All US LMP markets have a local market power mitigation (LMPM) mechanism. Even the current single zone market design in Colombia would benefit from an automatic local market power mitigation mechanism. The configuration of the transmission network, the level and location of demand, as well as the level of output of other generation units can endow almost any generation unit or group of generation units with a significant ability to exercise unilateral market power. A prime example of this phenomenon is the constrained-on generation problem described earlier. The owner of a constrained-on generation unit knows that regardless of the unit's offer price, it must be accepted to supply energy. Without a local market power mitigation mechanism, there is no limit to what offer price that supplier could submit and be accepted to provide energy.

In all offer-based electricity markets, a local market power mitigation (LMPM) mechanism is necessary to limit the offers any supplier submits when it faces is insufficient competition to serve a local energy need because of combination of the configuration of the transmission network, the levels and geographic distribution of demands, and the concentration of ownership of generation units. One lesson from the experience of US markets (in particular) is that system conditions can arise when virtually any generation unit owner has a substantial ability and incentive to exercise unilateral market power. That is why market structure-based market power mitigation mechanisms, typically used in Europe and other industrialized countries and initially employed in the US, where certain generation units or suppliers are designated in advance as having the ability to exercise unilateral market power, miss many instances of the exercise of substantial unilateral market power.

A LMPM mechanism built into the market software that relies on actual system conditions to determine whether any supplier has a substantial ability and incentive to exercise unilateral is likely to be significantly more effective. This logic explains why all US markets currently have LMPM mechanism built into their market software that runs automatically each pricing interval.

A LMPM mechanism is a pre-specified administrative procedure (written into the market rules) that determines: (1) when a supplier has local market power worthy of mitigation, (2) what the mitigated supplier will be paid, and (3) how the amount the supplier is paid will impact the payments received by other market participants. Without a prospective market power mitigation mechanism system conditions are likely to arise in all wholesale markets when almost any supplier can exercise substantial unilateral market power. It is increasingly clear to regulators around the world, particularly those that operate markets with limited amounts of transmission capacity, that these automatic regulatory interventions are necessary to deal with the problem of insufficient competition to serve certain local energy needs.

One argument often made about LMP markets is that they are more susceptible to the exercise of unilateral market power than a zonal market or single-zone market. The reality from the experience in the US leads to the exact opposite conclusion for the following reasons. First, mechanism used to set prices and dispatch levels does change the extent of competition a supplier faces. That depends on the number of other firms in the industry, the configuration of the transmission network and the level and distribution of demand in the network. Second, arguments that state because a zonal market or single-zone market sets prices over a larger geographic area than a single node suppliers face greater competition ignores a basic fact explained earlier that the financial market used to set price and initial dispatch levels is just the first step in a supplier's expected profit-maximizing offer strategy, because the supplier knows that all final schedules must be made physically feasible. Third, because an LMP market does not allow any energy or any ancillary service schedule to be physically infeasible, suppliers no longer submit offers that attempt to exploit the difference between the financial market and the physical operation of the grid. They now submit offer curves closer to their marginal cost of production, which produces a more efficient dispatch. The evidence from Wolak (2011b) on the economic benefits of transitioning from a multi-settlment zonal market design to a multi-settltement LMP market design are consistent with this logic.

Another important component of any local and system-wide market power mitigation mechanism is the provision of information to market participants and public at large. This is often termed, "smart sunshine regulation." This means that the regulatory process gathers a comprehensive set of information about market outcomes, analyzes it, and make it available to the public in a manner and form that increases the likelihood of market participant compliance with all market rules and allows the regulatory and political process to detect and correct market design flaws in a timely manner. Smart sunshine regulation is the foundation for all of the tasks the regulatory process must undertake in the wholesale market regime. Wolak (2014) discusses the benefits of smart sunshine regulation and public data release on wholesale market performance.

The Colombian wholesale electricity market is transparent in terms of making market input and market outcome data available to market participants and third-parties. Transparency could be improved by having a formal market monitoring function withint XM or independent of XM that analyzes this data and presents it in a more informative and useful manner. Various indexes of market performance could be presented at regular intervals so that performance of the market could be compared over time and to other international markets. All US markets have formal market monitoring departments and/or independent market monitors that take this data and analyze it and make it available to stakeholders, the wholesale and retail market regulator, and the public in a more digestible manner. The transparency of the Colombian market could benefit from the existence of a formal independent market monitor.

Another tool a regulator has in managing local and system-wide market power in an offer-based market is determining the configuration of the transmission network. Because the configuration of the transmission network can often determine the extent of competition that individual suppliers face, the regulator must take a more active role in the transmission planning and expansion process to ensure that competition-enhancing upgrades that improve market efficiency are built. Wolak (2019b) described the distinctly different roles of transmission network planning in the vertically-integrated monopoly regime versus the wholesale market regime. Wolak (2015a) presents a framework for measuring the competitiveness benefits of transmission expansions in an offer-based wholesale market and applies it to the Alberta, Canada wholesale electricity market.

The Colombian transmission network planning process could benefit from incorporating the wholesale market "competitiveness benefits" described in the Wolak (2019b) into its transmission planning process. Explicitly recognizing the economic benefits associated with increased competition between suppliers associated with transmission upgrades would be ensure that Colombian consumers realize the full economic and reliability benefits of electricity industry restructuring.

2.4.2. Information-Based Solutions to System-Wide Market Power

As discussed in detail in Wolak (2000), fixed-price forward contract commitments sold by generation unit owners reduce their incentive to exercise unilateral market power in the short-term energy market because the supplier only earns the short-term price on any energy it sells in excess of its forward contract commitment and pays the short-term price for any production shortfall relative to these forward contract commitments.

This logic argues in favor of the regulator monitoring the forward contract positions of retailers as part of its regulatory oversight process to ensure that there is adequate fixed-price forward contract coverage of final demand. As discussed in Wolak (2003b) and reinforced by the simulation results of Bushnell, Mansur and Saravia (2008), the California electricity crisis is very unlikely to have occurred if there had been adequate coverage of California's retail electricity demand with fixed-price and fixed-quantity forward contracts. Consequently, in order to protect against periods when one or more suppliers has a strong incentive to exercise unilateral market power, the regulator should, at a minimum, monitor the forward contracting levels of the retailers they oversee as the primary mechanism to protect against the exercise of system-wide unilateral market power.

The recommended long-term resource adequacy mechanism proposed in Section 3 would significantly increase the transparency of forward market positions of suppliers and retailers in Colombia and significantly increase to liquidity in the forward market at horizons to delivery far enough in advance to allow new entrants to compete to supply the energy.

2.4. Recommended for Short-Term Market Design for Colombia

Colombia should begin the process of transitioning to a two-settlement LMP market where generation units are paid the LMP at their location in the day-ahead and real-time markets, but all loads in Colombia pay the quantity-weighted average of these LMPs, similar to the case of Singapore, if there are concerns about charging consumers at different locations different wholesale prices. Both the day-ahead and real-time markets should have an automatic local market power mitigation mechanism incorporated into the market-clearing mechanisms.

Generation unit owners would submit three-part offers for their start-up, no-load, and energy offer curves, along with offer curves to supply each ancillary service the unit is capable of providing. The market operator will choose the least mix of energy and each ancillary service to meet the demand energy and ancillary services at all locations in the transmission network subject to configuration of the transmission and other relevant operating constraints, such generation unit ramping constraints in both the day-ahead and real-time markets.

As noted above, an important feature of the LMP market design is that it creates strong incentives for suppliers with no ability to exercise unilateral market power to submit offers to the short-term market close to their true costs because they know that these offers will be used to schedule their units to provide energy and ancillary services in a physically feasible manner. Because the network model and operating constraints used in the day-ahead and real-time markets are representative of the actual network and generation unit operating constraints, suppliers have no incentive to submit offers that attempt to exploit this difference as is the case in the current Colombian market design.

Implementing an LMP market design will require at least two years from the decision to make this transition to the day the market begins operations. Because there are many LMP markets in the United States, Colombia should choose the specific design best suited to initial conditions in Colombia and adapt it to the Colombia context. Attempting to create the perfect market design for Colombia is an impossible task. The process of designing a wholesale market is a process of continuous improvement, so the best strategy for Colombia is to adopt an existing LMP market design that is likely to work best for Colombia and make adjustments over time to improve its performance as more experience is gained with an LMP market in Colombia.

Before "going live" with the LMP market in Colombia there are number of checkpoints that must be met. The experiences of all US markets with this transition provide many examples to choose from for the design of this market re-design and transition process. Finally, the two market designs should be run in parallel for a sustained period of time so that market participants can gain confidence with the transition to the new market design.

There is likely to be resistance from market participants about transitioning to this market design because of its perceived complexity. This argument was made by market participants in all of the US wholesale markets that transitioned to LMP. Many participants vociferously objected to the supposed complexity of an LMP market and fretted about their ability to be financially viable in such a market design. However, once the design was implemented, market participants quickly adapted and it is fair to say that none now regret making this transition. “Fear of the unknown” or “fear of change” from stakeholders is to be expected in any transition to a new market design, but the experience of all markets that have made this transition suggest that it is more than worth the effort and expense, as the experience of California described in Wolak (2011) demonstrates.

3. A Long-Term Resource Adequacy Mechanism for Colombia

Why do wholesale electricity markets require a regulatory intervention to ensure long-term resource adequacy? Consumers want to be able to withdraw electricity from the network when they need it, just like other goods and services. But it is unclear why electricity is so fundamentally different from other products that it requires paying suppliers for production capacity to exist. For example, consumers want cars, but they do not pay for automobile assembly plants. They want point-to-point air travel, but they do not pay for airplanes. They want a loaf of bread, but do not pay for the existence of a bakery. All of these industries are high fixed cost, relatively low marginal cost production processes, similar to electricity supply. Nevertheless, all of these firms earn their return on capital invested by selling the good that consumers want at a price above the variable cost of producing it. Clearly cars, air travel, and bread are in many way essential commodities, yet there is no regulatory invention that ensures that there is sufficient production capacity for these products to meet demand.

So what is different about electricity markets that necessitates the need for a long-term resource adequacy mechanism? The answer lies in how short-term markets for these products operate relative to that for wholesale electricity. This difference is the result of the regulatory history of the electricity supply industry and the technology of electricity metering. Limitations on the level of short-term prices and the way that real-time supply shortfalls are managed in wholesale electricity markets creates what Wolak (2013) has been termed a "reliability externality" that requires an explicit regulatory intervention to internalize.

Different from the case of wholesale electricity, in the market for automobiles, air travel, and even bread, there is no explicit prohibition on the short-term price of the good rising to the level necessary to clear the market. Take the example of air travel. Airlines adjust the prices for seats on a flight over time in an attempt to ensure that the number of customers traveling on that flight equals the number of seats flying. This can result in very different prices for a seat on the same flight, depending on when the customer purchases the seat. A customer that waits too long to purchase a seat faces the risk of an infinite price in the sense that all of the seats on the flight are sold out. This ability to use price to allocate the available seats is also what allows the airline the flexibility to recover its total production costs. Airlines can set low prices to fill flights with low demand and extremely high prices on other flights, or at other times for the same flight, when demand is high.

The ability to use the short-term price to manage the supply and demand balance in the electricity supply industry is limited by a finite upper bound on a supplier's offer price into the wholesale market and/or a price cap that limits the magnitude of the eventual market-clearing price. In addition, historically virtually all electricity supply industries did not have interval meters that can record a customer's 15-minute or hourly consumption throughout the month. Even today, most regions only have mechanical meters that compute the customer's consumption for the entire month as the difference between two consecutive meter readings. With monthly or bi-monthly reading of mechanical meters, it is impossible for the utility to know how much electricity a customer consumed within a given hour of the month.[3]

Although these offer caps and price caps can limit the ability of suppliers to exercise unilateral market power in the short-term energy market, they also reduce the revenues suppliers can receive during scarcity conditions. This is often referred to as the *missing money* problem for generation unit owners. However, this missing money problem is only a symptom of the existence of the "reliability externality."

This externality exists because offer caps limit the potential downside to electricity retailers and large consumers (able to purchase from the short-term market) delaying their purchases of electricity until real-time operation. Specifically, if the retailer or large consumer knows the price cap on the short-term market is $250/MWh, then it is unlikely to be willing to pay more than that for electricity in any earlier forward market. This creates the possibility that real-time system conditions can occur where the amount of electricity demanded at or below the offer cap is less than the amount suppliers are willing to offer at or below the offer cap. This outcome implies that the system operator must be forced to either abandon the market mechanism or curtail load until the available supply offered at or below the offer cap equals the reduced level of demand, as was the case a number of times during the period January 2001 and March 2001 in California and has occurred during past El Nino events in Colombia when involuntary curtailments were necessary.

---

[3] Many regions have eliminated or are eliminating this technological barrier to allowing price to manage the real-time supply and demand balance by installing interval meters for all customers and offering real-time meter reading as a regulated service.

Because random curtailments of supply---also known as rolling blackouts---are used to make demand equal to the available supply at or below the bid cap under these system conditions, this mechanism creates a "reliability externality" because random curtailments imply that no retailer or large consumer bears the full cost of failing to procure adequate amounts of energy in advance of delivery. A retailer that has purchased sufficient supply in the forward market to meet its real-time energy demand is equally likely to be randomly curtailed as the same size retailer that has not procured adequate amounts of energy in the forward market. For this reason, all retailers and large loads have an incentive to under-procure their expected energy needs in the forward market.

Particularly for markets with very low offer caps, retailers have little incentive to engage in sufficient fixed-price forward contracts with generation unit owners to ensure a reliable supply of electricity for all possible realizations of real-time demand. For example, a 200 MW generation unit owner that expects to run 100 hours during the year with a variable cost of $80/MWh would be willing to sign a fixed-price forward contract to provide up to 200 MWh of energy for up to 100 hours of the year to a retailer. Because this generation unit owner is essentially selling its expected annual output to the retailer, it would want a $/MWh price that at least exceeds its average total cost of supplying energy during that year. This price can be significantly above the average price in the short-term wholesale market during the hours that this generation unit operates because of the offer cap on the short-term market and other market power mitigation mechanisms. This fact implies that the retailer would find it expected profit-maximizing not to sign the forward contract that allows the generation unit owner full cost recovery, but instead wait until the short-term market to purchase the necessary energy at prices that are limited by the offer cap.

Although this incentive for retailers to rely on a price-capped short-term market is most likely to impact generation units that run infrequently, if the level of demand relative to the amount of available supply is sufficiently large, it can also impact the revenues of other generation units. Because of the expectation of very low prices in the short-term market and the limited prospect of very high prices because of offer caps, retailers may decide not to sign fixed-price forward contracts with these generation unit owners and purchase their energy in the short-term market. By this logic, a short-term energy market with an offer cap always creates an incentive for retailers to delay purchasing some of their energy needs until real-time, when these caps can be used to obtain this energy at a lower price than the supplier would be willing to sell it in the forward market.

The lower the offer cap, the greater is the likelihood that the retailer will delay their electricity purchases to the short-term market. Delaying more purchases to the short-term market increases the likelihood of the event that insufficient supply will offer into the short-term market at or below the offer cap to meet demand. If a retailer knows that part of the cost of its failure to purchase sufficient fixed-price forward contracts will be borne by other retailers and large consumers because of random curtailment, then it has an incentive to engage in less fixed-price forward contracts than it would in a world where all customers had hourly meters and all customers could be charged hourly prices high enough to cause them to reduce their demand to equal the amount of supply available at that price.

Because externalities are typically the result of a missing market, another way of characterizing this reliability externality as a missing market for long-term contracts for energy. In this case, because retailers do not bear the full cost of failing to procure sufficient energy to meet their real-time needs in the future, there is a missing market for long-term contracts for long enough delivery horizons into the future to allow new generation units to be financed and constructed to serve demand under all possible future conditions in the short-term market. Therefore, a regulator-mandated long-term resource adequancy mechanism necessary to replace this missing market.

Specifically, unless the regulator is willing to eliminate or substantially increase the offer cap on the short-term market so that the short-term price can be used to equate available supply to demand under all possible future short-term market conditions, some form of regulatory intervention is necessary to internalize the resulting reliability externality. However, if customers do not have interval meters that can record their consumption on an hourly basis, they have a very limited ability to benefit from shifting their consumption away from high-priced hours, so raising or having no offer cap on the short-term market would not be advisable.

As the above discussion makes clear, relying on a capped short-term energy market price to ensure long-term resource adequacy does not address the reliability externality and leaves both a "missing money" and "missing market" problem. Capacity payment mechanisms are one approach to addressing this reliability externality designed primarily for thermal generation-dominated markets, where the major concern is insufficient generation capacity to meet future demand peaks. In a hydroelectric dominated market such as Colombia, the major reliability concern is more likely to be insufficient energy to meet a sustained demand for energy, which means that other approaches to addressing the reliability externality may dominate a capacity-based approach.

As the share of intermittent renewable generation in a wholesale electricity market increases, the magnitude of the reliability externality is also likely to increase. The uncertain availability of wind and solar resources increases magnitude and duration of potential future energy supply shortfalls that must be managed, which implies many more instances when a capped short-term energy market may not yield a sufficiently large energy supply increase or demand decrease to maintain real-time supply and demand balance.

3.1. Fixed-Price Forward Contract Approach to Long-Term Resource Adequacy

The fixed-price forward contract solution is the standard approach used to ensure a real-time supply and demand balance in markets for products with high fixed costs of production. The prospect of a high real-time price for the product provides incentives for customers to hedge this real-time price risk through a fixed-price forward contract. A supplier benefits from signing such a contract because it has greater quantity and revenue certainty as result.

The airline industry is familiar example of this phenomenon. There is a substantial fixed cost associated with operating a flight between a given origin and destination pair. Regardless of how many passengers board the flight, the airplane, pilot and co-pilot, flight attendants and fuel must be paid for. Moreover, there is a finite number of seats on the flight, so passengers wanting to travel face the risk that if they show up at the airport one hour before the flight and attempt to purchase a ticket, they may find that it is sold out or tickets are extremely expensive because of the high real-time demand for seats. Customers hedge this short-term price risk by purchasing their tickets in advance, which is a fixed-price, fixed-quantity (one seat) forward contract for travel on the flight. These forward market purchases allow the airline to better plan the types of aircraft and flight staff it will use to serve each route and how much fuel is needed for each the flight.

Similar arguments apply to wholesale electricity markets to the extent that real-time prices can rise to very high levels. For example, in Australia the offer cap on the short-term market is currently 14,500 Australia dollars ($AU) per megawatt-hour (MWh), yet annual average wholesale prices are less than $AU 100/MWh. The potential for short-term prices at or near the price cap provides a very strong incentive for electricity retailers and large customers to purchase their electricity through fixed-price forward contracts, rather than face the risk of these extreme short-term prices. However, even at this level of the offer cap on the short-term market in Australia the reliability externality still exists. Every few years there are small number of half-hour periods when supply shortfalls occur, consistent with the reliability externality argument.

Purchasing fixed-price and fixed-quantity forward contracts far enough in advance of delivery for new entrants to compete to provide this energy ensures that retailers will receive a competitive forward market price for their purchase. These forward market purchases far in advance of delivery also ensure that the seller of the contract has sufficient time to construct the new generation capacity needed to meet the demand to be served through the fixed-price forward contracts. Consequently, in the same sense that fixed-price forward contracts for air travel allow an airline to better match airplanes and flight staff to routes, fixed-price forward contracts for electricity allow electricity suppliers to choose the least cost mix of generation capacity to serve the demand that has purchased the fixed-price forward contracts for energy.

Many wholesale electricity markets outside of the US, particularly those in the developing countries, have offer caps far below $1,000/MWh. Low offer caps do not to create a strong enough incentive for load-serving entities to purchase enough fixed-price forward contracts far enough in advance of delivery to ensure sufficient generation capacity to meet future demand. Consequently, in a number of Latin American countries, there are regulator-mandated requirements for load-serving entities to purchase certain percentages of their final demand in fixed-price forward contracts in advance of delivery. For example, 90 percent of forecast demand one year in advance, 85 percent two years in advance and so forth. As I argue below, such a regulatory mandate can provide sufficient demand for long-term contracts far enough in advance of delivery to ensure enough generation capacity to meet future demand.

An important strength of a fixed-price, fixed-quantity forward contract approach to long-term resource adequacy is that these forward contract obligations significantly reduce the incentive of suppliers to exercise unilateral market power in the short-term market. To understand this logic, let PC equal the fixed price at which the supplier agrees to sell energy to an electricity retailer in a forward contract and QC equal to fixed quantity of energy sold. This contract is negotiated in advance of the date that the generation unit owner will supply the energy, so the value of PC and QC are predetermined from the perspective of the supplier's behavior in a short-term wholesale market.

Wolak (2000) demonstrates that the quantity of fixed-price forward contract obligations held by the supplier determines what short-term market price the firm finds ex post profit-maximizing given its marginal cost of producing energy, the supply offers of its competitors, and the level of aggregate demand. Incorporating the payment stream a generation unit owner receives from its forward contract obligations, its variable profit function for selling energy for a given hour of the day is:

$$\pi(PS) = (PC - C)QC + (QS - QC)(PS - C) \quad (1)$$

where QS is the quantity of energy sold in the short-term market and produced by the generation unit owner, PS is the price of energy sold in the short-term market and C is the supplier's marginal cost of producing electricity, which for simplicity is assumed to be constant.

The first term in (1) is the variable profit from the forward contract sales and the second term is the additional profit or loss from selling more or less energy in the short-term market than the supplier's forward contract quantity. Because the forward contract price and quantity are negotiated in advance of the delivery date, the first term, (PC – C)QC, is a fixed profit stream to the supplier before it offers into the short-term market. The second term depends on the price in the short-term market, but in a way that can significantly limit the incentive for the supplier to raise prices in the short-term market.

For example, if the supplier attempts to raise prices by withholding output, it could end up selling less in the short-term market than its forward contract quantity (QC > QS), and if the resulting market-clearing price is greater than the firm's marginal cost (PS > C), the second term in (1) will be negative. Consequently, only in the case that the supplier is confident it will produce more than its forward contract quantity in the short-term market does it have an incentive to withhold output in order to raise short-term prices.

The quantity of forward contract obligations held by a firm's competitors also limits incentive of that supplier to exercise unilateral market power in the short-term market. If a supplier knows that all of its competitors have substantial fixed-price forward contract obligations, then this supplier knows these firms will submit offer curves into the short-term market close to their marginal cost curves. Therefore, attempts by this supplier to raise prices in the short-term market by withholding output are likely to be unsuccessful because of the aggressiveness of the offers into the short-term market by its competitors with substantial fixed-price forward contract obligations limits the price increase a supplier can expect from these actions.

This logic implies an important role for transparency about the aggregate level of the fixed price forward contract obligations. If the regulator provides information on the aggregate quantity of a standardized fixed-price forward contracts held by all retailers and makes this information publicly available, then all suppliers can compute the magnitude of standardized fixed-price forward contract holdings of its competitors. If there an high degree of standardized fixed-price forward contract coverage by a supplier's competitors, then each supplier knows that it will face aggressive offer behavior (offer prices close to marginal cost) by its competitors and will find it uninlaterally profitable to submit more aggressive offers into the short-term market.

This dynamic creates the following virtuous cycle of forward contracting. If a supplier knows all of its competitors have a substantial amount of their expected energy sales covered by fixed-price forward contracts, then it has an incentive to sign fixed-price forward to contracts for a substantial fraction of its expected energy sales. When all suppliers have a substantial fraction of their expected energy sales covered by fixed-price forward contracts, then all suppliers have a common interest in reducing the cost of meeting these fixed-price forward contract obligations.

The resulting reductions in the short term prices from more competitive behavior in the short-term market caused by high levels of fixed price forward contract coverage of final demand creates another virtuous cycle from fixed-price forward contracts. Suppliers that have sold forward contracts have a strong incentive to make the cost of supplying these contracts as low as possible, by making the least cost "make versus buy" decision about how to fulfil these forward contract obligations. These actions should result in low short-term prices for energy, which will then be factored into subsequent negotiations for next round of fixed-price forward contracts. This dynamic for reducing short-term wholesale prices results from a persistently high level of coverage of final demand by fixed-price forward contracts that creates the incentive for all suppliers to reduce the system-wide cost of meeting these standardized forward contract obligations.

3.2. The Capacity Payment Approach to Long-Term Resource Adequacy

Particularly in the US, capacity payment mechanisms appear to be a holdover from the vertically-integrated regulated regime with regional power pools where capacity payments compensated generation units for their capital costs, because the regulated power pool typically only paid unit owners their variable operating costs for the electricity they produced. Therefore, all fixed costs had to be recovered through other mechanisms besides the sale of wholesale electricity.

Capacity payments typically involve a dollar per kilowatt year ($/kW-year) payment to individual generation units based on some measure of the amount of their capacity that is available to produce electricity during stressed system conditions, what is often referred to as the unit's "firm capacity." For example, a base load coal-fired unit would have a firm capacity value very close to its nameplate capacity. Usually, the firm capacity of a thermal unit is equal to the unit's nameplate capacity in MWs times its availability factor.[4]

In hydroelectric-dominated markets, determining the firm capacity of a generation unit is an extremely challenging task. The firm capacity of a hydroelectric generation unit owner is typically based on the amount of energy the unit is capable of providing under the worst possible hydrological conditions. However, it is difficult, if not impossible, to determine the maximum amount of capacity or energy a hydroelectric supplier can provide under these conditions, so there is a significant degree of arbitrariness in setting hydroelectric unit's firm capacity value. Second, because every hydroelectric supplier would like a larger capacity value for their generation unit, in order to avoid accusations of arbitrary firm capacity values for individual generation units, the entity making this decision typically bases the figure on the amount of energy the unit produced during the historically worst hydrological conditions even though the system operator may have sound reasons for believing that this firm capacity value is set too high.

As consequence, particularly in Latin America, there are numerous examples of capacity payment mechanisms that failed to ensure an adequate supply of energy and rationing conditions have been declared. Virtually all of the restructured markets in Latin America that have capacity payment mechanisms—specifically, Brazil, Chile, and Colombia—have experienced supply shortfalls that have required rationing. McRae and Wolak (2019) present an analysis of the most recent supply shortfall period in Colombia and conclude that the capacity payment mechanism there was a major contributing factor to this outcome.

Wind and solar generation units have a firm capacity values significantly below their nameplate capacity, but substantially more than the amount of energy these units are able to produce during extreme system conditions, particularly those that occur when the sun is not shining or the wind is not blowing. For example, on an extremely hot night solar generation units are not likely to produce any energy or on a hot sunny day, very little wind energy will be produced. Consequently, the process of computing the firm capacity values for wind and solar units involves a significant number of unverifiable assumptions often aimed at increasing the resulting firm capacity values. These facts imply that the capacity market construct is poorly suited to an electricity supply industry with significant intermittent renewable generation capacity.

[4] Thermal generation units convert heat energy into electricity. These include coal-fired, natural gas-fired, oil-fired, nuclear, and geothermal.

Capacity payment mechanisms differ along a number of dimensions. In some regions, the payment is made to all generation unit owners regardless of how much total generation capacity is needed to operate the system. In all US capacity markets, the independent system operator (ISO) specifies a system-wide demand for capacity equal to peak system demand plus some planning reserve, typically between 15 to 20 percent, and only makes capacity payments to enough generation units to meet this aggregate demand for firm capacity.

In the US, there have been many attempts to use market mechanisms to set the value of the $/kW-year payment to the generation units needed to meet the total demand for capacity. However, these capacity markets have been subject to almost continuous revision because they are extremely susceptible to the exercise of unilateral market power. The nature of the product sold—installed generation capacity—and a publicly disclosed perfectly inelastic demand for the product creates extreme opportunities for suppliers to exercise unilateral market.

In early versions of eastern US capacity markets, there were instances of the exercise of enormous amounts of unilateral market power, largely because these markets were operated on a short-term basis. During the off-peak months of the year when no single supplier is pivotal in the capacity market, the price of paid for capacity was very close to zero, which is the marginal cost of a supplier providing an additional MW of available capacity from existing generation capacity.[5] During the peak and shoulder months when one or more suppliers are pivotal in the capacity market, there was no limit on the price a supplier could charge.

This market power problem leaves open the question of how to set the value of the $/kW-year price cap on the capacity payment. In all regions of the US with capacity payment mechanisms currently employ the following administrative process for determining this price. The value of the maximum capacity payment is based on the regulator's estimate of annual $/kW fixed cost of a peaking generation unit. This is maximum price is typically backed by the argument that because of the offer cap on the short-term market and other market power mitigation mechanisms this peaking unit could only set an energy price slightly higher than its variable operating costs. Because this generation unit and all other generation units are missing the hours when the market price would rise above this unit's variable operating costs, the annual $/kW cost of the peaking

[5] A supplier is said to be pivotal in a market with an inelastic demand if some of their supply is needed to meet this demand regardless of the offers of other market participants.

unit is needed to compensate all generation units for the revenues they do not receive because of the offer cap and other system-wide and local market power mitigation mechanisms.

Offer prices and quantities from existing generation units in the firm capacity auctions in the US are typically set by an administrative process, typically a zero offer price and the unit's firm capacity. Market participants are only allowed to submit offer prices for firm capacity from new generation units as long as these offer prices are above a minimum offer price floor. The market-clearing price of firm capacity is then determined from the intersection of the administratively-determined demand curve and partially administratively-determined aggregate offer curve, which yields a largely administratively determined price of firm capacity.

This logic for setting the price ceiling on the value of $/kW-year capacity payment at the net cost of entry of new peaking generation unit explicitly assumes that the real-time demand for electricity is completely price inelastic and that suppliers are unable to exercise significant amounts of unilateral market power in the short-term market. Both of these assumptions are clearly false. An increasing number of jurisdictions around the world are installing interval meters that allow dynamic pricing plans to be implemented. Wolak (2013) discusses these technologies and the pricing plans they enable.

There is ample empirical for evidence for the exercise of unilateral market power in short-term energy markets, Wolfram (1999) applies a structural model of imperfect competition to measure the extent of unilateral market power exercised in the early England and Wales electricity market. Patrick and Wolak (2001) document the strategic use of market rules to raise prices by two large suppliers in this same electricity market. Wolak (2003a) documents changes in the unilateral ability of the five largest suppliers in California electricity market to exercise unilateral market power from 1998 to 2000. Borenstein, Bushnell and Wolak (2002) document the magnitude of the rent transfers and economic inefficiencies caused by these actions. McRae and Wolak (2012) demonstrate that the four large suppliers New Zealand appear to exercise unilateral market power.

Different from the case of fixed-price, fixed-quantity forward contracts for energy, capacity payment mechanisms do little to limit the exercise of unilateral market power in the energy market. To the extent that suppliers are able to exercise unilateral market power in the short-term energy market, they can raise energy prices significantly above the variable cost of the highest cost unit operating within the hour for all hours of the year, on top of receiving a capacity payment set by the highest offer price needed to meet the system demand for capacity. This logic holds even in markets such as Colombia with capacity payments mechanism involving peak energy rent refunds where suppliers must refund the positive difference between the short-term energy price and a scarcity price times the amount of firm energy sold by the supplier. The existence of the peak energy rent refund does not impact the incentive of a supplier to exercise unilateral market power in the short-term energy market below this scarcity price. Moreover, McRae and Wolak (2019) demonstrate that the existence of the peak energy rent refund mechanism in the current Colombia capacity payment mechanism increased consumer costs, reduced average water levels, and reduced system reliability.

In all US markets, capacity payment mechanisms are typically accompanied by offer caps on the short-term energy market, or in the case of Colombia the market has a peak energy rent refund. These limitations on the short-term price of energy significantly limit the incentive for final consumers to become active participants in the short-term wholesale market. For example, if the maximum wholesale price in an hour is $250/MWh because of an offer cap at this level, then a 1 KWh reduction in demand for a residential customer (a very large demand reduction) during an hour only saves the customer 25 cents, which is likely to be insufficient to cause that consumer to reduce its demand. This lack of an active demand-side in the wholesale market impacts how generation unit owners offer their units into the market, because all suppliers know that system demand will be the same regardless of the hourly wholesale price.

There is also a natural bias in operation of capacity payment mechanisms toward setting a higher value of the firm capacity requirement. This can significantly reduce the incentive for active participation by final demand in the wholesale market or for the development of innovative new products and technologies such as energy shortage. Both the regulator and market operator would prefer a higher values of firm capacity because system operation is easier with more firm capacity, particularly during high demand conditions. However, a higher level of firm capacity will lead to lower average prices and less volatile short-term prices. This pattern of short-term prices significantly reduces the incentive for active demand-side participation in the wholesale market or investments in storage and other load-shifting technologies. Finally, because all the generation units providing firm capacity will likely require annual full cost recovery to provide this product, the annual average cost of wholesale energy and capacity is likely to increase.

Capacity markets are also poorly suited for regions with a significant share of renewables such as Colombia. In these markets it is rarely, if ever, the case that there is a capacity shortfall in the sense that there is insufficient installed generation capacity to meet peak demand. The more common problem is insufficient energy in the form water stored behind a dam to meet demand. With wind and solar photovoltaic generation units, capacity shortfalls are also extremely unlikely. It is more likely that the sun does not shine or the wind does not blow for a sustained period of time. In both of these cases, the problem is not a capacity shortfall, but an energy shortfall. Consequently, a capacity payment mechanism that focuses on ensuring adequate installed capacity is unlikely to deliver the most cost effective solution for consumers to the problem of long-term energy adequacy in regions with a significant amount of intermittent renewable resources.

The argument for a capacity market is strongest in a region with all dispatchable thermal generation units and no potential for active participation of final consumers in the wholesale market, conditions which do not exist in Colombia. The generation fleet is dominated by hydroelectric generation units and there is desire by the government to scale the amount of intermittent wind and solar resources. Moreover, given the diversity of uses of electricity in Colombia, there is likely to be a significant amount final demand that would be willing to actively participate in the short-term market.

3.3. Recommended Long-Resource Adequacy Process for Colombia

The goal of this long-term resource adequacy mechanism is to internalize the reliability externality with the least amount of adverse market efficiency consequences. This mechanism requires all electricity retailers and large consumers to purchase a standardized fixed price forward contract for energy equal to various fractions of their expected demand at various horizons to delivery. For example, a retailer or large consumers could be required to purchase 95 percent of its actual annual demand in a fixed price forward contract one year in advance of delivery, 90 percent of its actual annual demand two-years in advance of delivery, 85 percent of its annual actual demand three years in advance of delivery, and 80 percent four years in advance of delivery, and all of these contracts would financially clear against the hourly short-term price during the delivery period. In the case of an multi-settlement LMP market, these contracts would clear against the day-ahead price of energy faced by all retailers and large consumers. Retailers and large consumers would be subject to financial penalties for under-procurement of these forward contracts relative to their actual demand.

For example, if the retailer's realized demand is 100 gigawatt-hours (GWh) and it purchased 96 GWh in the standardized forward contract one year in advance, 89 GWh two years in advance, 86 GWh three years in advance, and 81 GWh four years in advance, the retailer would only be subject to penalties for under-procurement two years in advance of delivery. These forward financial market energy purchases would provide retailers and large consumers with wholesale price certainty for the vast majority of their electricity demand. To the extent the regulator feels that these mandated financial contracting levels are insufficient to ensure a reliable supply of electricity at a reasonable price, higher levels of contract coverage can be mandated, say 98 percent one year in advance, 93 percent two years in advance, 90 percent three years in advance, and 85 percent four years in advance.

It is important to emphasize that mandating these contracting levels is unlikely to impose a financial hardship on retailers that lose customers to competing retailers. If a retailer purchased more fixed-price forward contract coverage than it ultimately needs because it lost customers to a competitor, it can sell this obligation in the secondary market. Unless the aggregate demand for energy in the future is unexpectedly low, this retailer is just as likely to make a profit on this sale as it is to make a loss, because one of the retailers that gained customers is going to need a standardized forward contract to meet its regulatory requirements for coverage of its final demand. Only in the very unlikely case that the aggregate amount of forward contracts purchased is greater than the realized final demand for the system, will there be a potential for stranded forward contracts held by retailers that lose customers.

Under this mandated standardized forward contract for energy resource adequacy regime, thermal resource owners would likely sell significantly more energy in the standardized fixed-price forward contract than they expect to produce and the renewable resource owners likely would sell significantly less than they expect to produce. When there are significant amounts of available renewable energy, the thermal resource owners will find it profitable not to produce electricity, but instead purchase from the short-term market to meet their forward financial contract obligations. In contrast, when little renewable energy available, the thermal resource owners will produce the difference between the market demand and renewable generation production. Because the renewable resource owners have sold less of the standardized fixed-price forward contract than their expected output, they are likely to have only a modest net short position in energy that they must cover through purchases from the short-term market.

Under this mechanism, the renewable supplier produces whenever the wind or solar resource is available and the thermal resource owner only produces when the wind and solar resources are unavailable. The thermal resources make the efficient, "make versus buy" decision by submitting an offer to supply energy in the short-term market at their variable cost. This offer price ensures that when it cheaper for the thermal resource owner to meet it standardized forward financial contract obligation from its generation unit, it will be accepted to supply energy in the short-term market and when it is cheaper to purchase this energy from the short-term market its unit will not operate.

The precise form of the standardized forward contract sold is an important aspect of this reliability mechanism. The most straightforward approach would be to make the contracts system-load-weighted. The values of Q_{hk}(Contract), the forward contract obligation of supplier k during hour h, would be computed as follows. Let QD_{hd} equal the system demand in hour h of day d. Define $w_{hd} = \frac{QD_{hd}}{\sum_{d=1}^{D}\sum_{h=1}^{24} QD_{hd}}$, as the share of system demand in the period, say a quarter, that the standardized forward contract clears. In this case, D is the number of days in the quarter. Suppose that supplier k sells QQ_k(Contract) MWh of energy for the quarter. The hourly value of the contract for both the buyer and seller of the contract is $QC_{hdk} = w_{hd}[QQ_k(\text{Contract})]$. Specifically, the quarterly total amount of energy sold is allocated to hours in the quarter according to the actual load shape during that quarter. The values of the w_{hd} would be higher during the hours of the day when the value of system load is higher, which would make QC_{hdk} higher during those hours.

The basic idea of this approach is to adjust the hourly values of the total amount of energy sold in a quarter to match the hourly load shape in order to ensure that suppliers in the aggregate have the strongest possible incentives to ensure that system demand is met a least cost. If all consumer purchase 100% of their quarterly demand in this standardized product, by matching the pattern of the hourly level of fixed-price forward contracts to the hourly pattern system demand, suppliers in the aggregate know that system demand in every hour is fully contracted. Therefore, suppliers jointly have an incentive to ensure this pattern of hourly demand is met at least cost.

This standardized product could be traded and cleared through the XM. The major challenge of enforcing this long-term reliability mechanism is managing the counterparty risk between buyers and sellers of this product. Specifically, XM would need to ensure that at any time until delivery of the contract is completed both the seller and buyer of this product would prefer to fulfill their contractual obligation rather than default. This implies that XM would need to operate a clearinghouse to manage margin requirements on this standardized contract similar to any other forward market for a commodity.

Different from the case for other commodities, the demand for these standardized forward contracts must come from a regulatory mandate because of the logic for the reliability externality described above. However, the prices paid for energy under these contracts will be determined through market mechanisms. Suppliers and loads are also free to re-trade these obligations. The only requirement is that once a retailer makes its forward contracting compliance filing with the regulator, it is no longer allowed to sell the contracts it has used for compliance. The contracts used for compliance purposes must be held to the clearing date of the contract.

For the regulated retail customers, the purchase prices of these forward contracts can also be used to set the wholesale price implicit in the regulated retail price over the time horizon that the forward contract clears. Specifically, the average price paid for standardized contracts that deliver during a given quarter by an electricity retailer can be used to set the average wholesale price of power during that quarter implicit in the quarterly regulated retail price. Consequently, one benefit of this reliability mechanism is that it provides a transparent market price to set the wholesale price component of a retailer's regulated retail price.

A final point to emphasize about this reliability mechanism is that there is no prohibition on generation unit owners or retailers engaging in other hedging arrangements outside of this mechanism. Specifically, a generation unit owner could enter into a bilateral contract for energy with another generation unit owner or a retailer. Because it mandates purchases of standardized contracts out to four years into the future, this reliability mechanism should stimulate the development of an active forward market for individually designed products to hedge the residual risks that suppliers, retailers and free consumers face that cannot be hedged with the standardized forward contract at these time horizons to delivery.

3.3.1. Advantages of Proposed Approach for Enhancing Long-Term Resource Adequacy

This mechanism has a number of advantages relative to a capacity payment mechanism. There is no up-front payment made to a generation unit to be available yet not operate as is case with the current capacity payment mechanism. This would avoid the situation described in McRae and Wolak (2016) where a large thermal generation received capacity payments for numerous years, but when it came time to operate during an El Nino event the unit owner refused to operate. All sellers of the long-term resource adequacy product are selling energy each hour of the day and make the efficient make versus buy decision by submitting an offer into the short-term market at the unit's marginal cost of producing energy.

XM would operate the standardized forward contract market and manage the counterparty risk of financial default of all buyers and sellers of the contracts through its existing margin requirements process in a manner consistent with how any standardized forward financial market operates. XM can tailor the margin requirement process to the financial health of individual market participants, if necessary. If a supplier has sold energy forward at a substantially lower price than the short-term price that it is likely to clear against, XM can require this market participant to post additional funds to ensure that the supplier will follow through with its forward market commitment. XM can manage a clearinghouse function among buyers and sellers of this standardized forward financial contract as part of its existing settlements system.

It is not unusual for standardized forward financial markets for energy to have outstanding volume trading at horizons to delivery up to four years in the future. The New York Mercantile Exchange (NYMEX) offers standardized forward contracts that clear up to 10 years in the future for natural gas and oil. There is typically significant outstanding volume in these contracts at delivery horizons at least four years into the future. Over the past 10 years, oil and natural gas markets have shown significant price volatility yet these standardized futures markets have continued to operate. Based on this evidence it appears feasible to manage the margin requirements on standardized electricity contracts in Colombia with delivery horizons up to four years in the future or even longer if the regulator determines it is necessary.

This long-term resource adequacy mechanism does not require the regulator to set capacity reserve margin. Generation unit owners collectively make this decision based on their standardized forward contract sales to retailers and large consumers. Moreover, as discussed earlier, collectively generation unit owner have a common financial interest in finding the least cost mix of generation capacity to meet demand in all hours of the year. In addition, if a retailer or large consumer is able to limit its demand during certain hours of the year, it can avoid purchasing as much energy in the forward market, which can reduce the amount of generation capacity necessary to serve demand on an annual basis.

This mechanism focuses on ensuring that there is adequate energy available to meet system demand at a reasonable price all hours of the year, whether it be from generation units providing energy or a final consumer reducing their demand in response to a higher short-term price. This mechanism significantly limits the incentive suppliers have to exercise unilateral market power in the short-term market, because all suppliers have substantial fixed-price forward contract obligations to supply energy during all hours of the year and all market participants know that the hourly aggregate magnitude of these obligations follows the shape of system demand.

Before each compliance period for the standardized-forward-contract reliability mechanism the following sequence of events would occur. Each retailer and free consumer would show the regulator the quantity of contracts that it has purchased at various horizons to delivery. Specifically, each retailer would show the total energy purchased in each quarterly standardized forward contract for each quarter over the next four years. This showing could be validated by the XM. These contracts would then be placed in that retailer's compliance account and held until the clearing date. The retailer would be required to hold these standardized long-term resource adequacy contracts until the delivery.

There are a number of reasons why this mechanism will achieve long-term resource adequacy in Colombia in a more cost-effective manner than a capacity market. First, the sale of fixed price forward contract for energy three to four years starting delivery two years in the future, provides a revenue stream that will significantly increase investor confidence in recovering the cost of any investment in new generation capacity. To the extent that the regulator would like to provide potential new investors with revenue certainty for a longer period time, the terms of these standardized forward contracts could be increased and fraction of realized demand that must be purchased in advance could be increased. This mechanism will provide strong incentives for demand to be served under all possible system conditions. Each generation unit owner will have fixed-price forward contract obligations by hour of the day that matches the pattern of system demand throughout the quarter of the year. Moreover, the final demand of all retailers and large loads are hedged against short-term prices through these standardized fixed-price forward financial contracts. Therefore, generation unit owners bear all of the short-term price risk associated with supply shortfalls and have the means to manage this risk because they own and operate generation units. However, this does not mean that final consumers cannot become active participants in the short-term market to help maintain the real-time supply and demand balance if this is a more cost-effective solution.

Because retailers and loads are protected from high short-term prices by these standardized forward contracts, the offer cap on the short-term market can be increased which will increase the incentive for suppliers to produce during stressed system conditions. The higher price cap and potentially greater short-term price volatility will support investments in storage and load-shifting technologies and encourage active participation of final demand in the wholesale market, further enhancing system reliability.

If the regulator is concerned that adequate new generation capacity is being built to meet future demand, milestones could be set for completing the proposed generation project that is the basis for a forward contract energy sale. For example, if energy was sold four years in advance based on a proposed generation unit, then XM could require that construction of the new unit to begin within a pre-specified number of months after the signing date of the contract or require posting a substantial larger amount of money in the seller's margin account, otherwise the contracted quantity of energy would be automatically liquidated. This means the supplier would have to buy back the energy that it sold based on the expected energy output of the proposed unit.

Other completion milestones would also have to be met at future dates to ensure the unit will be ready to provide energy on its original initial operating date, and if any of these milestones were not met, the contract would be liquidated. On this point, it is important to emphasize that in order for a supplier to liquidate contracted energy, the supplier must find a willing seller of the energy that it previously sold in a standardized forward contract. Because the retailers and free consumers that purchased these contracts must use them to meet their compliance obligations, buying back this energy is likely to require the supplier to pay an extremely high price, which provides the supplier with a strong incentive to deliver the forward energy sold from a proposed generation unit. Because of these incentives, our proposed mechanism is unlikely to leave nearly as many, if any, proposed projects sold as standardized-contract energy to be canceled or completed far behind schedule as McRae and Wolak (2016) document for current capacity payment mechanisn in the Colombia.

One question likely to be raised about an approach that focuses on the development of an active forward market for energy is whether sufficient generation resources will be built to meet demand if consumers only purchase standardized forward financial contracts that clear against the short-term price. On this point, it is important to bear in mind the incentives faced by a seller of the forward financial contract once this contract has been sold. The supplier has an obligation to ensure that the forward contract quantity of energy can be purchased in the short-term market at the lowest possible short-term price.

The seller of the contract bears all of the risk associated with higher short-term prices. In order to prudently hedge this risk, the seller has a very strong incentive to ensure that sufficient energy is available from its generation units or generation units owned by other market participants to set the lowest possible price in the short-term market for the quantity of energy it sold in the fixed-price forward contract.

This logic implies that if a supplier signs a forward contract guaranteeing the price for 500 MWh of energy for 24 hours a day and 7 days per week, it will construct or contract for more than

500 MW of generation capacity to hedge this short-term price risk. Building only a 500 MW facility to hedge this risk would be extremely imprudent and expose the supplier to significant risk, because if this 500 MW facility is unavailable to provide electricity, the supplier must purchase the energy from the short-term market at the price that prevails at the time. Moreover, if this generation unit is unavailable, then the short-term price is likely to be extremely high.

The standardized energy contracting approach to long-term resource adequacy has a number of features that are likely make it significantly more affordable than a capacity-based approach. First, the procurement process focuses precisely and only on what consumers want--energy. Second, the precise mix of generation capacity and active demand-side participation required to achieve this goal is left up to market participants rather than a regulatory mandate. Third, the potential to raise the offer cap on the short-term market creates greater opportunities for active demand-side participation, storage and load-shifting technologies to eliminate the need to build new generation capacity because of the lower load factor enabled by active demand-side participation.

The energy contracting approach is substantially less complex to administer than a capacity-based approach. There is no need to define the firm capacity of a generation unit. There is no need to set an aggregate firm capacity requirement. There is no need to set the parameters of an administrative demand curve for firm capacity. The only aspect that does involve some level of complexity is the verifying compliance of retailers and large loads with the contracting requirement. However, this administrative and regulatory burden is likely to be no greater than that necessary to verify compliance with the capacity obligations associated with administering a capacity market.

This mechanism will also encourage cross-hedging between intermittent renewable generation resource owners and dispatchable generation unit owners, which is likely to become even more important to ensuring long-term resource adequacy in an electricity supply industry with a significant amount of wind and solar resources. A wind or solar resource owner that sells the standardized fixed-price, fixed-quantity forward contract for energy long-term resource adequacy product to a retailer would like to reinsure the quantity risk associated with selling a fixed-quantity product.

The wind or solar resource owner could sign a contract with a thermal resource owner that provides insurance against the quantity risk it faces from selling a fixed-price and fixed-quantity contract. For example, the wind resource owner could purchase a cap contract for the quantity of energy sold in this contract at a certain strike price and in this way have insurance against having to purchase energy from the short-term market at an extremely high price to fulfil this contractual obligation when the wind or solar resource is not producing energy. The up-front payment to the thermal resource for the price spike insurance would help to finance the fixed costs of thermal resource that operates significantly less frequently because of the large amount of intermittent renewable generation capacity. The wind or solar resource owner would then factor in the cost of this quantity risk insurance in the price it is willing to sell the standardized fixed-price, fixed quantity forward contract for energy long-term resource adequacy product.

3.4.2. Transition to an Standardized-Energy Contract Approach to Resource Adequacy

The most straightforward way to transition to a standardized-energy contract approach to long-term resource adequacy is to assign all generation units standardized forward contracts equal to their firm capacity value for the term of the unit's capacity contract. The contract price would be set equal to an average per MWh Bolsa market revenue plus total capacity payments less refunds over a number of historical years that is sufficient to account for the relative frequency that normal water years versus El Nino event years occur. This price could then be adjusted up or down by the regulator on a case-by-case basis to ensure no suppliers are harmed or excessively benefited by this transition.

The process uses the average of past market outcomes to set the price at which firm energy contracts are converted into standardized long-term resource adequancy energy contracts. Suppliers that believe this price is insuffient to recover their costs net of any ancillary services revenues, could make a cost-of-service showing with the CREG to raise their price. The resulting converted contracts would then be allocated to all retailers and load according to their annual load share.

4. Policies to Support NCRES, New Products, and Services

A multi-settlement LMP market design with an mandated standardized forward contracts for energy long-term resource adequacy mechanism is ideally suited to support the expansion of the NCRES energy in Colombia. In addition, this market design will also foster investments in storage and other load-shifting technologies, as well as the entry to purely financial participants that can improve the performance of the competitive retailing sector and wholesale market outcomes.

4.1. Policies to Support NCRES

As discussed in Section 2.2, a multi-settlement LMP market values the dispatchability of generation units by paying higher average short-term prices to these units relative to intermittent wind and solar resources. This does not mean that intermittent NCRES units are disadvantaged relative to dispatchable resources. If the levelized cost of an NCRES generation unit is less than the market value of the energy this resource produces, then investment in the NCRES generation unit is financially viable. The levelized cost of a generation resource is equal to $\text{LCOE} = \frac{\sum_{t=0}^{T} \frac{C_t}{(1+r)^t}}{\sum_{t=1}^{T} \frac{E_t}{(1+r)^t}}$, where C_t is the net cost of the project in year t=0,1,2,...,T, E_t is the net energy in MWh produced in year t=1,2,…,T and r is the discount rate for future cash flows. The average market value of energy produced is equal to $\text{MVAL} = \frac{\sum_{t=0}^{T} \frac{p_t E_t}{(1+r)^t}}{\sum_{t=1}^{T} \frac{E_t}{(1+r)^t}}$, where p_t is the market price paid to the energy produced in period t. If MVAL is greater than the LCOE for a unit then the unit owner's investment has a positive net present value. If this is not the case then an explicit subsidy must be paid to resource owner to make the unit financially viable.

Any subsidies for NCRES investments should be provided in the form of lump sum payments needed to make the project financial viable rather in the form of a feed-in tarff or fixed-price contract for all of the energy the unit produces. Providing support to a NCRES resource in the form of a fixed-price for all of the energy produced by the unit, effectively subsidizes the NCRES resource and places the quantity risk associated with this contract on the buyer and other market participants. Under the contract form the unit owner has no incentive to manage its energy production risk. Even for a dispatchable generation unit, paying a fixed-price for all the energy produced by the unit would effectively subsidize this generation unit as well. This unit owner would be significantly less interesting producing when short-term prices are high.

As discussed earlier, a NCRES resource could sell a standardized fixed price, fixed quantity long-term resource adequacy forward contract that is shaped to the pattern of system load, but this would require the unit owner to purchase quantity risk insurance from a dispatchable generation unit. If an NCRES resource sells a forward contract that pays a fixed-price for all of the energy produced by the unit regardless when the energy is produced, this contract cannot be used for compliance with the long-term resource adequacy mechanism. However, a dispatchable generation unit could purchase this contract and use the dispatchability of its generation unit to firm up this contract and sell a standardized long-term resource adequacy forward contract.

The NCRES unit owner can factor in how these imbalances will be settled in making offers to supply fixed-price and fixed-quantity long-term contracts for energy. Shifting renewable resource owners to fixed-price and fixed-quantity forward contract from fixed-price and quantity-produced contracts will also provide financial incentives for renewable resource owners manage the intermittency of their production through storage investments and financial contracts that finances fast-ramping dispatchable generation resources to provide insurance against renewable energy shortfalls. Transitioning forward contracts for renewable energy to require the seller to manage the quantity risk associated with the energy they sell is an important step in the process of increasing the amount intermittent renewable energy produced in a region in a cost-effective manner.

4.2. Active Demand-Side Participation and Storage Investments

The active involvement of final consumers in the wholesale market can reduce the amount of installed generation capacity needed to serve them and can reduce the cost of integrating an increasing amount of intermittent renewable generation. A multi-settlement market with a day-ahead forward market and real-time market allows loads to purchase energy in the day-ahead market that they can subsequently sell in the real-time market. Without the ability to purchase demand in the day-ahead market that is not consumed in real-time, demand reduction programs require the regulator to set an administrative baseline, which can significantly reduce the system-wide benefits of active demand-side participation. This issue is discussed in Bushnell, Hobbs, and Wolak (2009).

There are three necessary conditions for active involvement of final consumers. First, customers must have the necessary technology to record their consumption on an hourly basis. Second, they must receive actionable information that tells them when to alter their consumption.[6] Third, they must pay according to a price that provides an economic incentive consistent with the actionable information to alter their consumption. A major challenge to active involvement of final consumers in the wholesale market in a developing country is the availability of the technology to record the customer's consumption on an hourly basis.

---

[6] McRae and Meeks (2016) resents the results of a field experiment in Central Asia that demonstrates the importance of actionable information for facilitating active demand-side participation. Kahn and Wolak (2013) find that once customers understand nonlinear pricing, they subsequently make energy consuming decisions consistent with responding to the marginal price. Wolak (2015b) presents evidence consistent with real-time consumption feedback producing energy conservation efforts by households in Singapore.

Mechanical meters that record a customer's consumption during a monthly or bi-monthly billing cycle as difference of two consecutive meter readings cannot be used in dynamic pricing plan because the retailer has no way of knowing when during the billing cycle customer consumed electricity. Even if the customer is billed for wholesale electricity using a standardized load shape, the customer's monthly bill falls by the same amount regardless of when the customer reduces their consumption within the month. For example, if w(h) is the load shape weight for hour h in the month ($\sum_{h=1}^{H} \mathrm{w(h)} = 1$), p(h) is the hourly wholesale price, and Q(m) the household's monthly consumption, the wholesale electricity cost for the monthly billing cycle is $\sum_{h=1}^{H} \mathrm{p(h)w(h)Q(m)}$, where H is the total number of hours in billing cycle.

The above equation makes it clear that with mechanical meters read on a monthly or bi-monthly basis, dynamic pricing is impossible to implement because the customer's monthly bill falls by same amount regardless of when consumption is reduced within the month. Therefore, a household faced with a higher average monthly price would reduce consumption when it is least costly to do so, not when the hourly price is highest.

There is growing empirical evidence that all classes of customers can respond to short-term wholesale price signals if they have the metering technology to do so. Patrick and Wolak (1999) estimate the price-responsiveness of large industrial and commercial customers in the United Kingdom to half-hourly wholesale prices and find significant differences in the average half-hourly demand elasticities across types of customers and half-hours of the day. Wolak (2006) estimates the price-responsiveness of residential customers to a form of real-time pricing that shares the risk of responding to hourly prices between the retailer and the final customer. The California Statewide Pricing Pilot (SPP) selected samples of residential, commercial, and industrial customers and subjected them to various forms of real-time pricing plans in order to estimate their price responsiveness. Charles River Associates (2004) analyzed the results of the SPP experiments and found precisely estimated price responses for all three types of customers. More recently, Wolak (2011a) reports on the results of a field experiment comparing the price-responsiveness of households on a variety of dynamic pricing plans. For all of pricing plans, Wolak found large demand reductions in response to increases in hourly retail electricity prices across all income classes. Although most dynamic pricing experiments and programs have relied on day-ahead price signals, recent work by Andersen, Hansen, Jensen and Wolak (2019) has shown that customers can respond to within-day price signals.

The business case for investments in storage and other load-shifting technologies is built on exploiting prices differences across hours of the day, week, month or even the year. Consequently, price volatility is a major driver of these investments. A multi-settlement LMP market with an energy contract-based long-term resource adequacy mechanism can allow a significantly higher offer cap on the short-term market which is likely to result in greater short-term price volatility, which will make more investments in storage cost-effective. A capacity-based long-term resource adequacy mechanism that eliminates short-term price volatility provides virtually no economic incentive for investments in storage.

Storage investments also earn revenues from selling ancillary services. An LMP market where energy and ancillary services procurement are co-optimized, combined with an energy contracting long-term resource adequacy mechanism is also likely to increase the level of ancillary services prices during stressed system conditions because the opportunity cost of generation units providing ancillary services is higher. Consequently, the expected ancillary services revenues earned by storage investments are also likely to be higher than those under a capacity-based long-term resource adequacy mechanism.

The above logic suggests that if the recommendations of Sections 2 and 3 are adopted, Colombia would most productively spend government funds to deploy interval meters for as many customers as possible. This will allow as many customers as possible to benefit from dynamic retail pricing and allow as many as customers as possible to benefit from storage and other automated load-shifting investments. In most jurisdictions, investments in interval meters have been undertaken by the regulated distribution network owner with meter reading services provided at regulated prices to all electricity retailers. This approach takes advantage of the economies to scale in the installation and operation of interval meter reading systems. The most important reason to deploy interval meters is to provide the capability to capture and record a customer's hourly consumption for each hour of the year. This logic argues in favor of investments in low-cost interval meters that simply read and broadcast a customer's 15-minute consumption to the distribution utility. All value-added services can be competitively provided by retailers.

4.3. New Products, Services and Market Participants

A multi-settlement LMP market allows purely financial participation in the wholesale market mechanism. As shown in Jha and Wolak (2019), purely financial participants can reduce the cost of supplying thermal energy to the wholesale market. A liquid forward market for energy can allow the entry of purely financial participants in electricity retailing, which Wolak (2019b) has shown can increase the competitiveness of retail market outcomes and wholesale market outcomes.

4.3.1. Virtual or Convergence Bidding

All US LMP markets currently allow for purely financial participation through explicit virtual bidding. With explicit virtual bidding (EVB), every market participant has access to the following purely financial instrument: buy (sell) one MWh of electricity at a given location and hour-of-the day in the day-ahead market if the day-ahead price is below (above) the offer price, with the obligation to sell (buy) back one MWh at the same location and hour-of-the-day in the real-time market as a price-taker (i.e.: accepting the prevailing real-time price for closing out this purely financial position in the day-ahead market). These financial offer curves are termed ``virtual bids'' or ``convergence bids'' because an expected profit-maximizing purely financial trader will typically take positions at a location in the day-ahead market that reduces the magnitude of the difference between day-ahead and real-time prices at that location.

Although virtual bids and physical bids are separately identified to the independent system operator (ISO), it treats physical and virtual bids the same when running the day-ahead LMP pricing and dispatch process. However, the ISO knows that any day-ahead ``virtual energy'' sale or ``virtual energy'' purchase must be reversed in the real-time market. Recall that in the real-time market only physical generation units that submit offers into the real-time market are dispatched to meet the actual demand at each location in the transmission network. Recall that generation units with energy schedules sold in the day-ahead market can either produce this energy from their own unit or purchase the energy from the real-time market.

If a purely financial player sells 10 MWh of virtual energy at a given location in the day-ahead market, she must purchase this 10 MWh back at the real-time price for that location because she cannot actually supply any energy in real-time. Similarly, if a purely financial player buys 10 MWh of virtual energy in the day-ahead market, she must sell 10 MWh at the real-time price at that location because she cannot consume any energy in real-time.

This logic implies that the actions of virtual bidders directly influence day-ahead and real-time market outcomes, typically by closing the gap between day-ahead and real-time prices. For example, submitting a virtual bid to sell (buy) one MWh in the day-ahead (real-time) market earns positive revenues if and only if the day-ahead price is higher than the real-time price. However, submitting this virtual bid increases supply (demand) in the day-ahead market (real-time market), making it less likely that day-ahead prices will be higher than real-time prices.

EVB allows both physical market participants and purely financial players to submit virtual bids in the day-ahead market at any location for any hour of the day. The profit-maximizing actions of market participants seeking to exploit day-ahead/real-time price spreads at any location on the grid creates incentives for the day-ahead schedules of generation units to be as close as possible to their real-time output. This outcome means that the California ISO has to accept fewer supply and demand offers from generation units in the real-time market. These actions also create incentives for day-ahead generation schedules to be as close as possible to the real-time output levels that minimize the cost of operating the transmission network in real-time.

As noted earlier, LMPs and generation unit dispatch levels in the day-ahead market are the solution to a mixed-integer programming problem, which can be challenging to solve during high demand hours when transmission constraints and generation unit run-time and ramping constraints are likely to be binding. The LMP optimization problem determines which generation units to start up and how much each of these generation units are scheduled to produce. This optimization problem has thousands of choice variables and thousands of constraints and thus has many potential local optima. If virtual bidders are able to figure out lower-cost solutions to meeting demand across the transmission network, they will likely also earn profits from the resulting difference between day-ahead and real-time prices for the reasons described below.

Consider the example of a transmission link between nodes A and B where the virtual bidder believes that 3 MWs more transmission capacity will be made available in real-time than is available in the day-ahead market, so that more energy can flow from A to B in real-time. A virtual bidder can submit a virtual demand bid for 3 MWhs at node A. This demand bid increases the amount of generation scheduled at node A by 3 MWhs. Similarly, she can submit a virtual supply bid at node B for 3 MWhs that reduces the amount of energy supplied at node B by this amount. These two virtual bids create 3 MWs of virtual transmission capacity between nodes A and B. When the real-time market clears and it turns out that 3 MWs more transmission capacity between A and B becomes available, the additional 3 MWhs of virtual supply provided in real-time at node A will flow to node B, replacing the 3 MWhs of virtual demand provided in real-time at node B. In this case, the virtual bidder is likely to earn a profit on both transactions because she correctly anticipated the existence of the additional transmission capacity in real-time.

It is important to emphasize that the purely financial participant does not have to know that there is a difference between the amount of available transmission capacity between node A and node B in the day-ahead versus real-time markets to find this profitable strategy. She would only need to notice that day-ahead prices at node A are lower than real-time prices and day-ahead prices at node B are higher than real-time prices. This alone would cause her to submit a demand bid in the day-ahead market at node A and submit a supply bid in the day-ahead market at node B.

Another example concerns the case of whether to start a fossil fuel-fired generation unit in anticipation of producing the next day. Because of the runtime constraints, if this unit is not committed in the day-ahead market it will not be available to operate in real-time. This energy must instead be produced by units that are more responsive but have higher operating costs. Suppose a purely financial player believes that committing this long-start unit in the day-ahead market provides a lower cost solution to meeting demand in the peak hours of the day than turning on fast-responding but more expensive units. This financial player can submit virtual demand bids at that unit's location in the day-ahead market for the peak hours of the day in order to commit the unit in the day-ahead market, which in turn enables lower cost real-time dispatch. In real-time, the financial player would then sell the virtual energy back at a higher average price for these hours.

Again, the purely financial player would not need to understand why day-ahead prices are lower than real-time prices at the location of the runtime-constrained fossil fuel generation unit during these hours. Purely financial players would only need to attempt to exploit this profitable arbitrage opportunity by submitting virtual demand bids during peak hours of the day at this location. The implied sale in the real-time market improves the efficiency of the real-time market outcome by dispatching the run-time constrained unit in the day-ahead market.

With the introduction of explicit virtual bidding (EVB), all physical market participants--generation unit owners and load-serving entities--as well as all purely financial participants can submit virtual supply and demand bids at any of the thousands of nodes where virtual bidding is allowed by the California ISO. As the above examples illustrate, the process of exploiting profitable day-ahead/real-time arbitrage opportunities at nodes throughout the transmission network can yield market efficiency gains.

4.3.1. Purely Financial Retailers

Purchasing a fixed-price forward contract for wholesale electricity is a lower cost and less risky approach to entering electricity retailing that can increase the extent of competition faced by incumbents in electricity retailing and wholesale supply. A new entrant purchases a fixed-price forward contract for wholesale energy for the term of the retail contract for the total amount of energy the customer is likely to consume. This forward contract hedges the vast majority of the short-term wholesale price risk faced by the entrant and also sets a lower bound on the fixed-price retail contract it can offer. Armed with a portfolio of fixed-price forward contracts at various delivery horizons in the future, the new entrant can compete against incumbent suppliers that own generation assets to sell retail contracts that deliver energy during these future periods.

If the new entrant is successful in obtaining retail customers and this forward contract is held to the clearing date, the generation unit owners that sold these forward contracts now have larger total fixed-price forward contract obligations when they submit their offers into the short-term market. By the logic in Wolak (2000), these generation unit owners have an incentive to submit offers into the short-term wholesale market that are closer to their marginal cost of production, which should result in lower wholesale prices.

Wolak (2019a) examines the empirical validity of the above logic that the introduction of a futures market for wholesale electricity increased competition in both electricity retailing and wholesale electricity supply in Singapore. These standardized fixed-price forward contracts are traded on the Singapore Exchange Limited (SGX) and clear against the half-hourly Uniform Singapore Electricity Price (USEP) at a rate of 0.5 megawatt-hour (MWh) each half-hour of the quarter.

Using data on the prices and other observable characteristics of all fixed-price retail contracts signed between large consumers and electricity retailers from October 2014 to March 2016, Wolak (2019a) find lower prices for these retail contracts signed after the introduction of purely financial retailers enabled by the start of the standardized futures market. Higher average volumes of open futures contracts that clear during the term of the retail contract during the month before the retail contract started delivering energy also predict a lower retail contract price. Both of these empirical results are consistent with the presence of purely financial retailers increasing the competitiveness of retail market outcomes.

Wolak (2019b) finds that the total savings in energy purchased in retail contracts from May 1, 2015 through March 31, 2016 from the introduction of the futures market and average monthly volume of open positions in the market during that period is, depending on the econometric model specification, between 18 and 24 percent of the total spending on retail energy contracts during this same time period.

Wholesale market performance in Singapore also appears to have improved as a result of the introduction of the futures market. Wolak (2019b) find that higher volumes of futures contracts clearing against the USEP during a half-hour predicts a lower wholesale price during that half-hour. This result is consistent with the forward market purchases of purely financial retailers increasing the aggregate fixed-price forward market obligations of generation unit owners in Singapore, which, by the logic described in Wolak (2000), should cause these generation unit owners to submit offer prices closer to their marginal cost of producing electricity, which then results in lower wholesale electricity prices.

Short-term energy wholesale energy cost savings from July 1, 2015 through April 30, 2016 from the introduction of the futures market and the open positions in the futures market during that period are, depending on the econometric specification, between 5 and 19 percent of the total spending on wholesale energy over that same time period. These retail and wholesale market price results support the argument that establishing a formal futures market can increase the competitiveness of both retail and wholesale market outcomes. Independent retailers can initially serve final consumers incurring minimal sunk costs of entry through purchases of wholesale energy from the futures market. As these retailers increase the number of customers served and system demand grows, they can invest in new generation capacity. This lower cost and less risky entry strategy is particularly relevant for electricity supply industries where generation asset ownership is concentrated, a common initial condition in many restructured electricity supply industries.

The experience of Singapore provides a strong argument for the case that a forward contract-based approach to long-term resource adequacy in Colombia will facilitate the entry of purely financial participants, which will increase the competitiveness of both retail and wholesale market outcomes in Colombia.

References

Andersen, L.M., Hansen, L.G., Jensen, C.L., and Wolak, F.A. (2019) "Can Incentives to Increase Electricity Use Reduce the Cost of Integrating Renewable Resources," available at http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/Into_versus_Away_January_2019.pdf

Bohn, Roger E., Caramanis, Michael C. and Schweppe, Fred C. (1984) "Optimal Pricing in Electrical Networks over Space and Time," *RAND Journal of Economics*, Vol. 15, No. 5, 360-376,

Borenstein, Severin, Bushnell, James, and Wolak, Frank A. (2002) "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market," *American Economic Review*, December, 1367-1405.

Bushnell, J. B., Mansur, E. T., & Saravia, C. (2008). Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets. *The American Economic Review*, 98(1), 237-266.

Bushnell, J., Hobbs, B. F., & Wolak, F. A. (2009). When it comes to demand response, is FERC its own worst enemy?. *The Electricity Journal*, 22(8), 9-18.

Bushnell, James, Hobbs, Benjamin, and Wolak, Frank (2008) "Final Opinion on 'The DEC Bidding Activity Rule under MRTU'," available at https://www.caiso.com/Documents/MSCFinalOpiniononDECBiddingActivityRuleunderMRTU.pdf

ENTSO-E (2018) Bidding Zone Configuration Technical Report 2018, October, available at https://docstore.entsoe.eu/Documents/Events/2018/BZ_report/20181015_BZ_TR_FINAL.pdf

Charles River Associates (2004) "Statewide Pricing Pilot Summer 2003 Impact Analysis," Charles River Associates, 5335 College Avenue, Suite 26, Oakland, California 94618.

Kahn, M.E. and Wolak, F.A. (2013) Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment, available at http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/files/kahn_wolak_July_2_2013.pdf

Mansur, Erin T. and White, Matthew W. (2012) Market Organization and Efficiency in Electricity Markets," available at http://www.dartmouth.edu/~mansur/papers/mansur_white_pjmaep.pdf

McRae, Shaun D. and Meeks, Robyn (2016) "Price Perceptions Price Perceptions and Electricity Demand with Nonlinear Prices," available at http://www.sdmcrae.com/research/

McRae, Shaun D. and Wolak, Frank A. (2012) "How Do Firms Exercise Unilateral Market Power? Evidence from a Bid-Based Wholesale Electricity Market," *in Manufacturing Markets: legal, political and economic dynamics*," Jean-Michel Glachant and Eric Brousseau (editors), Cambridge University Press.

McRae, Shaun D. and Wolak, Frank A. (2016) "Diagnosing the Causes of the Recent El Niño Event and Recommendations for Reform," available at http://web.stanford.edu/group/fwolak/cgi-bin/

McRae, Shaun D. and Wolak, Frank A. (2019) "Market Power and Incentive-Based Capacity Payment Mechanisms," available at http://web.stanford.edu/group/fwolak/cgi-bin/

Patrick, Robert H. and Wolak, Frank A. (1999) "Customer Response to Real-Time Prices in the England and Wales Electricity Market: Implications for Demand-Side Bidding and Pricing Options Design Under Competition," in *Regulation Under Increasing Competition*, Michael A. Crew (editor) Kluwer Academic Publishers, 155-182.

Tangerås, Thomas P. and Wolak, Frank A. (2018) "The Competitive Effects of Linking Electricity Markets Across Space and Time," available at at http://web.stanford.edu/group/fwolak/cgi-bin/

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

Wolak, F. A., & Patrick, R. H. (2001). The impact of market rules and market structure on the price determination process in the England and Wales electricity market (No. w8248). National Bureau of Economic Research.

Wolak, F. A. (2003a). Measuring unilateral market power in wholesale electricity markets: the California market, 1998-2000. *American Economic Review*, *93*(2), 425-430.

Wolak, Frank A. (2003b) "Diagnosing the California Electricity Crisis," *The Electricity Journal,* August, 11-37.

Wolak, Frank A. (2006) "Residential Customer Response to Real-Time Pricing: The Anaheim Critical-Peak Pricing Experiment," available from http://www.stanford.edu/~wolak.

Wolak, Frank A. (2009) "Report on Market Performance and Market Monitoring in the

Colombian Electricity Supply Industry," available at http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/files/sspd_report_wolak_july_30.pdf

Wolak, Frank A. (2011a) "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment," *American Economic Review*, May, 83-87.

Wolak, Frank A. (2011b) "Measuring the Benefits of Greater Spatial Granularity in Short-Term Pricing in Wholesale Electricity Markets," *American Economic Review*, May, 247-252.

Wolak, Frank A., (2013) "Economic and Political Constraints on the Demand-Side of Electricity Industry Re-structuring Processes," *Review of Economics and Institutions*, Vol. 4, No. 1, Winter, Article 1.

Wolak, F. A. (2014). Regulating competition in wholesale electricity supply. In *Economic Regulation and Its Reform: What Have We Learned*? (pp. 195-289). University of Chicago Press.

Wolak, F. A. (2015a). Measuring the competitiveness benefits of a transmission investment policy: The case of the Alberta electricity market. *Energy Policy*, 85, 426-444.

Wolak, F.A. (2015b) Do Customers Respond to Real-Time Usage Feedback? Evidence from Singapore, available at http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/Singapore%20residential_nov_2015.pdf

Wolak, F. A. (2019a). The benefits of purely financial participants for wholesale and retail market performance: lessons for long-term resource adequacy mechanism design. *Oxford Review of Economic Policy*, *35*(2), 260-290.

Wolak, F.A. (2019b) "Transmission Planning and Operation in the Wholesale Market Regime available at http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/transmission_wolak_july_27.pdf

Wolfram, C. D. (1999). Measuring duopoly power in the British electricity spot market. *American Economic Review*, 89(4), 805-826.