

Innovation and Market Design

Peter Cramton*

30 March 2008

Preliminary Draft

Abstract

Market design plays an essential role in promoting innovation. I examine emission allowance auctions, air slot auctions, spectrum auctions, and electricity markets, and demonstrate how the market design can encourage innovation. Improved pricing information is one source of innovation. Enhancing competition is another driver of innovation seen in all of the applications. Market design fosters innovation in other ways as well by addressing other potential market failures.

(JEL D44, C78, L96. Keywords: innovation, market design, auction design, spectrum auctions, electricity markets, air slot auctions.)

1 Introduction

Market design determines the rules under which market participants interact. In this chapter, I argue that market design can play an essential role in promoting innovation. Fundamentally, this is done by establishing rules that strengthen the incentives for innovation. Enhancing competition is one common way the market design encourages innovation. This is seen in all the examples I present. Improved price information is another. Better price information reduces innovation risk and improves decisions to innovate. Innovation is also encouraged by identifying other sources of market failure, and then mitigating these potential failures through the market rules.

I discuss four important applications of market design: 1) emission allowance auctions, 2) air slot auctions, 3) spectrum auctions, and 4) electricity markets. For each I describe how the market design can foster innovation.

Emission allowance auctions illustrate the important role of using prices to motivate innovation. In contrast to command and control approaches to managing pollution, emission allowance auctions—part of cap and trade programs—price the scarce allowances for the various pollutants. Firms can then respond to the prices and make efficient decisions, both short term and long term, on how best to reduce emissions. In this way, the environmental goal can be achieved at least cost. The objective of the market is to provide reliable price information for each of the pollutants and enable market participants to purchase the desired portfolio of allowances. The market design allows for both forward and spot purchase.

Air slot auctions are a preferred method of allocating scarce runway capacity at congested airports, such as those serving New York City. An open auction prices and assigns the scarce

* Department of Economics, University of Maryland. I am grateful to the National Science Foundation for funding.

runway capacity, which has desirable benefits for investment and operation, as was the case with emission allowances. In addition, an air slot auction can be designed to encourage competitive entry. The alternative of giving air slots to incumbent carriers and then allowing trade has proven to be ineffective. Trades other than barter transactions among incumbents have not taken place. In contrast, a well-designed auction of air slots provides a viable opportunity for entrants to secure the slots needed to provide an innovative service.

Spectrum auctions have promoted the efficient pricing and assignment of radio spectrum for wireless services. The market design has played a key role in enhancing competition and innovation in wireless services. State-of-the-art auction designs allow the band plan and technology choices to be made by the bidders rather than being set by the regulator before the auction. Innovative technologies and business models can then compete on par with the incumbent approaches. Technology neutral auction designs are currently being implemented in the United Kingdom.

Electricity market design demonstrates how the product design can play an important role in enhancing competition, reducing risks, and promoting innovation on the demand side as well as the supply side. Wholesale electricity markets are organized as a number of auctions. Long-term investment markets for capacity (or firm energy in the case of hydro systems), medium-term auctions for forward energy, and spot auctions for day-ahead and real-time electricity. These markets price a variety of products at different times and locations. Good designs mitigate market power problems that frequently can arise, especially during times of scarcity. Forward capacity auctions coordinate the efficient investment in new capacity. In addition, by bundling a call option to provide energy at prices above a strike price, the auction greatly mitigates incentives to exercise market power during times of spot scarcity. The forward energy market provides an additional hedge to customers at lower energy prices. The two markets combined go a long way to eliminating market power in the spot market, improving the dispatch of energy resources. The product definition in the forward energy market can also encourage innovation on the demand side. Customers with hourly meters can be hedged for their expected energy purchase, thereby reducing risk, yet exposed on the margin to the hourly price. This exposure motivates demand response and investment in innovative demand management systems.

Auction applications are rapidly expanding. Communication and computational advances have certainly played an important role, but the development of simple and powerful auction methods has been important too. Market designers now have a much richer set of tools to address more complex problems.

I now discuss each of the four applications—emission allowances, air slots, spectrum, and electricity—in greater detail. Each application uses auctions to facilitate the efficient allocation of scarce resources, promote competition, and foster innovation.

2 Emission allowance auctions

Controlling pollution is an essential role of government in our modern economy. Humans have an ever expanding ability to destroy or preserve our environment. In the past, pollution has been addressed through command-and-control regulation. Governments establish specific rules that describe what pollution abatement measures must be taken. The problem of course is that the government setting the rules has limited knowledge about the costs and benefits of various

approaches across emitters. Indeed, the government is not even aware of the full set of abatement methods. As a result inflexible and inefficient rules are established.

The alternative, which is fast growing in popularity, is for the government to use a market-based approach to pollution control. The lead example are cap and trade programs, which are now used for global pollutants like carbon dioxide as well as regional and local pollutants like sulfur dioxide and nitrogen oxides. The idea is simple. The government focuses on establishing the environmental goal—setting a cap on the maximum quantity of each pollutant. Then it relies on competitive auctions to price the scarce resource—emission allowances. There is no need to tell emitters whether and how each should reduce its emissions. Rather, the abatement decision-making is done by each emitter based on the price information determined in competitive auctions. In this way, the environmental goal is achieved at least cost. Each emitter reduces emissions to the point where the auction price of allowances is equal to the emitter's marginal cost of abatement. Reductions are done by those who can reduce emissions at least cost.

A key advantage of the approach is the enormous flexibility each emitter has to manage its emissions strategy. In addition, the approach is simple. Prices are used to motivate the emitter to adopt efficient approaches based on today's abatement measures. Prices also motivate the development of new abatement techniques—techniques that are both less expensive and more effective at reducing emissions. In contrast, with command and control there is no incentive for the development of new techniques, rather the incentive is to find ways to satisfy the command at least cost, which may actually undermine the environmental goal.

Experience with cap and trade programs, such as the U.S. acid rain program, have confirmed the economic insight that prices are effective at achieving environmental goals at least cost, and foster innovative techniques for abatement going forward (Ellerman et al. 2000).

Market designs that provide better price information both in the short term and in the long term will be more effective. When there are several pollutants, a simultaneous ascending clock auction can allow emitters to bid for an efficient portfolio of emission allowances (Ausubel and Cramton 2004). Forward auctions can allow emitters to lock in prices early as part of a risk reduction strategy. These forward prices are especially useful in motivating efficient investment decisions for longer-run abatement approaches.

Emission markets are relatively simple, especially for a global pollutant like carbon. Carbon allowances are a homogenous good that can be defined broadly. Complicating factors such as time and place are less important. Competitive factors and network factors are also less important. As we will see, the other applications I look at must deal with each of these complicating factors, so more complex market designs are required. Nonetheless, the foundation for each is the same: auctions are used to price and assign scarce resources efficiently.

3 Air slot auctions

Package auctions have been proposed for auctioning takeoff and landing rights at congested airports, such as the three New York City airports (Ball et al. 2007). The goal of the auction is to make the best use of the scarce runway capacity. Left to their own devices, airlines will overschedule flights during peak hours, creating congestion and costly delay. The package auction enables each airline to bid for its preferred package of slots. The resulting competitive prices motivate airlines to substitute away from expensive slots, either by shifting flights to less expensive times or using larger aircraft to carry the same passengers with less runway use.

Some airports, such as New York LaGuardia or Washington National, manage congestion by limiting the number of scheduled operations in any fifteen minute period to the airport's capacity. The air slots are assigned based on historical use and negotiation. Then the participants can trade slots as desired. What we have learned after many years with this approach is that there is almost no liquidity in the market for air slots. Despite large changes in the industry, there are few trades, and the few trades that do occur tend to be barter transactions between two airlines. Prices for air slots are not established, and an airline, such as a new entrant, is unable to buy any significant number of slots on the market. As a result, the air slot cap serves as an entry barrier, limiting competition, and discouraging the efficient use of the runway capacity. In addition, use-it-or-lose-it rules cause airlines with surplus slots to schedule more small-plane operations, rather than sell or lose the excess slots.

An air slot auction promises to improve the allocation of air slots as a result of transparent pricing and improved liquidity for slots. Liquidity is important in this application, since airlines require a critical mass of routes to and from the airport, and each route requires a minimum number of operations each day. Shifts in strategy are apt to involve many slots. An auction allows for entry and exit in a dynamic industry.

An alternative to an auction is congestion pricing. Both seek efficient pricing. Auctions can be designed with a multi-year lease, and thereby more stability in airline planning. Congestion pricing gives airlines more flexibility to change schedules. The difficulty with congestion pricing is establishing a workable process for setting prices. This process is apt to become politicized, and in any event, determining the market clearing prices without an auction is challenging.

A common critique of a slot auction is that it is just another tax that will raise prices to consumers and limit service. This is wrong in several ways. First, the auction is not a tax, but a method for efficiently pricing a scarce resource—runway capacity. Indeed, the revenues from an auction can displace the distortionary passenger fees and weight-based fees. Revenues can also be used to improve airport infrastructure, say to encourage more flexible use of airport facilities. Second, in airports that limit scheduling to capacity, the scarcity price is already reflected in the airline's fare. Finally, slot auctions do not limit service, but rather constrain scheduling to be consistent with runway capacity. When schedules are not constrained to physical limits, the result is congestion and delay, not increased throughput.

Consumers benefit most from a slot auction. The immediate benefit is reduced congestion and delay, and improved predictability. Travel costs are reduced, especially for the time-sensitive business traveler. In the longer term, consumers benefit from enhanced competition and innovation. Even just the threat of entry can cause airlines to reduce costs and improve services.

A slot auction also has benefits for airlines. With fewer delays and more predictable schedules, airlines save fuel, labor, and capital. Since business travelers put a premium on predictability and short trips, demand for business travel—the most profitable segment of the market—expands. Larger and fuller planes translate into more profit per flight.

I now describe the mechanics of one proposed air slot auction with illustrative parameters.

The product is the right to schedule a landing within a particular 15-minute time interval, and a takeoff within 90 minutes of landing. Bundling a landing and take-off makes sense, since every landing requires a take-off. The number of slots is determined by the FAA based on throughput.

Since schedules are done on roughly a yearly basis, and commitments to airports and routes are longer term, a five-year term for the slot is reasonable. The terms are staggered so that 20% of the slots expire each year. These slots are auctioned every year, providing annual liquidity and price information. Secondary market trades of both the primary product and derivatives are also allowed.

A single simultaneous auction is held for all congested airports. The annual auction is a package clock auction. All slots are up for auction at the same time. Bids are always for packages of slots: the bidder wins one of its desired packages or nothing. As a result, there is no risk that the bidder will win just part of what it needs.

The auction begins with an ascending clock stage. The auctioneer announces prices and then the bidders respond with the quantity demanded at these prices. Prices then increase on slots with excess demand, and bidders again respond with demands at the new prices. The process continues until there is no excess demand for any slot. The clock stage provides essential price discovery during the auction. The prices help focus each bidder's valuation efforts on the most relevant packages of slots.

A supplementary round follows the clock stage. Bidders can improve their clock bids and submit additional package bids.

Based on the set of all package bids, the auctioneer then determines the value-maximizing assignment of slots and the prices for slots in each time interval. To encourage truthful bidding, a second price rule may be used as described in greater detail in the next section.

This design has been well-tested in the two applications discussed next: spectrum auctions and electricity auctions. It has also been tested in the experimental lab and found to have excellent efficiency properties. In addition, in February 2005, a test auction with the FAA, airport, and airline participants was conducted to demonstrate the feasibility of the approach.

The package clock auction gives airlines a simple and effective way to express preferences for packages of slots. Given these preferences, the auction then determines the efficient assignment of slots and prices.

Efficient pricing of slots motivates airlines to adjust schedules. Flights are shifted to less congested hours. Larger planes and fewer flights are used to serve the same number of passengers. These adjustments are managed in a flexible way by each airline.

Importantly, the package clock auction facilitates efficient entry and exit. In part this comes from transparent pricing and liquidity in the market for slots, but more directly from the package auction, which enables a new entrant to bid for and win the package of slots it desires.

Although slot auctions have been discussed for at least twenty years, the application is still in the proposal stage. It may be some time before the benefits of slot auctions are seen in practice. In contrast, our next application, spectrum auctions, has been an active area of both design and implementation for fourteen years. The benefits of the approach are not speculative.

4 Spectrum auctions

Since 1994, spectrum auctions have been used to assign and price scarce radio spectrum for wireless communication services in the US and elsewhere. Before auctions, beauty contests, in which companies lobby the regulator for spectrum, were the primary method of assigning

spectrum. Beauty contests were both slow and costly. So slow in the US that, despite developing the cellular technology in the 1960s, by the 1980s the US had fallen behind both Europe and Japan in developing and deploying the technology. Innovation clearly was harmed by an ineffective method of assigning and pricing spectrum. Beginning in July 1994, the US switched to auctions. Auctions allowed the spectrum to quickly get in the hands of those best able to use it. Competition increased. Consumers enjoyed lower prices and improved services. Throughout this period, the market for wireless communications has been characterized by rapid growth and innovation.

The spectrum auction application involves assigning many items that are heterogeneous but similar. Often there are competing technologies that make use of the spectrum in different ways and therefore require a different organization of the spectrum. There is a complex structure of substitutes and complements, and this structure varies across bidders.

The government's primary objective is efficiency: make the best use of the scarce spectrum resource. To a large extent this involves assigning the spectrum to the companies that value it the most, although there are also important competition issues in the downstream market for wireless services that further complicate the market design problem.

4.1 Simultaneous ascending auction

In July 1994, the FCC began using the simultaneous ascending auction to award spectrum licenses. The approach has been replicated around the world with minor variations and has become the standard approach for assigning and pricing spectrum (Cramton 2002, 2006, Milgrom 2004). The FCC now has conducted over 67 auctions using the simultaneous ascending auction. Roughly 59 thousand licenses have been auctioned with winning bids exceeding \$79 billion. Over the last dozen years, enhancements to the design have appeared, but the enhancements have been evolutionary, not revolutionary.

The popularity and durability of the design is a reflection of its many desirable properties.

Simplicity is an important virtue. The simultaneous ascending auction is easily described and understood. It is a natural generalization of the English auction when selling many related lots. All the lots are auctioned at the same time. Each lot has a price and a high bidder associated with it. The bidders can bid on any of the lots by raising the high bid. The bidding continues until no bidder is willing to raise the bid on any of the lots. Then the auction ends with each bidder winning the lots on which it is the high bidder, and paying its bid for any lots won.

A transparent process of price discovery is another key virtue. As the auction progresses bidders see the tentative price information and condition subsequent bids on this information. Over the course of the auction, bidders are able to develop a sense of what the final prices are likely to be, and can adjust their purchases in response to this price information. To the extent price information is good and the bidder retains sufficient flexibility to shift towards its best package, the bidder is able to piece together a desirable package of lots. Moreover, the price information helps each bidder focus its valuation effort to relevant packages.

To encourage price discovery, an activity rule requires each bidder to maintain a level of bidding activity throughout the auction or have its eligibility reduced in future rounds. Each lot has a number of eligibility points, based on the size of the lot, typically measured as the product of the bandwidth (MHz) and the population coverage (pop) of the lot. A bidder's initial eligibility is determined by its deposit at qualification. The bidder must be active on a specified

percentage of its eligibility or its eligibility is reduced. A bidder is active on a lot either if it is the current high bidder or it places a bid on the lot. The bidder's total activity in the round is the sum of the eligibility points on all lots the bidder is active. Typically, the required activity percentage is less than 100% early in the auction to give the bidder greater flexibility, but is 100% or close to 100% later in the auction. With a 100% activity requirement, the quantity of spectrum a bidder bids for can only stay the same or fall as prices rise. This prevents the bidder from bid sniping—waiting until the end of the auction to reveal its true demand.

A main simplification is accepting bids only for individual lots, rather than packages. As a result, a bidder may attempt to win a synergistic package of lots, but may later find that the package is too expensive and yet remain a high bidder on some of the lots in the package. The auction allows bid withdrawals to enable the bidder to back out of a failed aggregation of lots. To maintain the credibility of bids, the withdrawing bidder remains on the hook for assuring that the seller receive at least the amount of the withdrawn bid in revenues—the withdrawing bidder must make up any shortfall between its bid and the eventual sale price. This ability to withdrawal bids is intended to mitigate the exposure the bidder faces of the possibility of winning just some of what it needs.

For reasons of transparency and also price discovery, the auctioneer reports all the bids made by each bidder at the end of each round.

Despite the simplicity of the rules, the simultaneous ascending auction is complex for bidders. The reason is that bidders often have strong incentives to engage in various gaming strategies, which undermine efficiency.

To summarize, the simultaneous ascending auction has many strengths. It is a simple price discovery process. Bidders can arbitrage across substitutes and piece together desirable packages of complementary licenses. The price discovery reduces common value uncertainty and the winner's curse.

At the same time, years of experience has demonstrated important weaknesses in the design, which can reduce efficiency. Large bidders engage in demand reduction strategies to reduce the prices paid for the spectrum (Ausubel and Cramton 2002). Tacit collusion has also been a problem with bidders proposing splits of the spectrum through their bids (Cramton and Schwartz 2002). The activity rule has led to various parking strategies, where eligibility is maintained by bidding on spectrum that the bidder does not desire. Since bids are for individual lots, a bidder is exposed to the risk of winning some of what it needs. Speculators engage in various hold up strategies. Finally, depending on how the spectrum is divided into lots by the regulator, the substitution across blocks may be limited.

Limited substitution has been especially problematic in the two most recent spectrum auctions in the US: the AWS auction and the 700 MHz auction. In both auctions, the FCC, in order to accommodate the needs of a diversity of bidders, split the bandwidth into blocks of various sizes, and then used a number of different geographic partitions for licensing. Lot sizes varied from large regional license with 6 covering the continental US, to small submarkets, requiring 734 lots to cover the nation. This heterogeneity in lot sizes undermines substitution. Despite the simultaneous sale, bidding in the AWS auction was largely sequential. Blocks with large regional lots were bid on first, followed in turn by the blocks with smaller lots. A bidder seeking large regions would start out bidding for the large lots. If prices got too high, it could switch down to the small lots, but it would find it almost impossible to switch back up to the

large lots, since it would likely remain the high bidder on many of the small lots. This one-way substitution means that the bidder must guess the right time to drop down to the smaller licenses. If the bidder waits too long it may overpay for the large lots; if the bidder drops down too early it may overpay for the small lots. In the AWS auction, concluding in September 2006, the limited substitution led to large price differences across substitute blocks throughout the auction. At the end, the blocks with small lots sold for a 40% discount relative to the blocks with large lots. In the 700 MHz auction, concluding in March 2008, the difference in prices across blocks was even more extreme, but this time in the other direction—the block with the small lots sold at a price several times higher than the price of the block with the largest lots.

Given the packaging limitations of the simultaneous ascending auction, there is a strong tendency for the regulator to address these limitations with a band plan that offers something for everyone. The basic problem with this approach is that it tends to resolve matters that should be settled in the auction. The more the band plan is tailored to fit the needs of particular bidders, the less competition in the auction, and the less efficient the outcome. The auction has little to do if the band plan makes clear who should win what.

4.2 Package clock auction

With a package auction, bidders can express preferences for complementary items without running the risk that they will win just some of what they need. This is important, for example, in spectrum auctions in which different technologies require that the spectrum be organized in different ways. In the past, the regulator has been forced to decide how the spectrum is organized with a specific band plan—effectively deciding how much spectrum is available for each technology. A package auction enables the regulator to conduct a technology-neutral auction, which lets the bidders determine the band plan through their competitive bids. A good example is the United Kingdom’s 2.6 GHz auction of 2008: the quantity of paired vs. unpaired spectrum is determined in the auction, not by the regulator. Some technologies, such as LTE, require paired spectrum; others, such as WiMax, require unpaired spectrum.

One of the challenges of package auctions is finding an effective way for bidders to convey preferences. There are simply too many packages to ask for preferences for all possible packages. A common approach is to begin with a clock auction (Ausubel and Cramton 2004, Ausubel et al. 2006, Porter et al. 2003). The auctioneer names a price for each product and each bidder responds with its most preferred package. The price is then raised on all products with excess demand, and the bidding continues. This price discovery process focuses the bidders’ attention on packages that are most relevant. Once this price discovery is over the bidders are in a much better position to submit any additional bids, as well as improve the bids already submitted. An optimization is then done to determine the value maximizing assignment, as well as competitive prices that satisfy the stability constraints. Typically, there are many such prices, so a further optimization is done to find the prices that provide the best incentives for truthful bidding. This is the basis for the package clock auction, introduced in the last section. It addresses the deficiencies of the simultaneous ascending auction.

This basic design is being used for a series of spectrum auctions in the UK and has been proposed for an auction in the Netherlands. The design is readily tailored to alternative settings.

Here are some of the important design choices of the UK’s 2.6 GHz auction to assign 190 MHz of spectrum.

- *Generic 5MHz lots.* The use of generic, rather than specific, lots is a huge simplification. This improves substitution, enhances competition, and greatly simplifies bidding. Further, a winner’s lots are guaranteed to be contiguous, which makes the spectrum more valuable. Generic lots are appropriate in this auction, since the value differences of specific assignments are likely second order. The bidder’s main determinant of value is the quantity of paired or unpaired spectrum won.
- *Package bids.* There is no exposure problem, since the auction uses package bids. A bid specifies the number of paired and unpaired lots from 2500-2690 MHz. The number of packages is modest, thanks to the generic lots.
- *Clock stage.* The auction begins with a clock stage. Typically, only a single price is needed: the price of a generic 5MHz lot from 2500-2690. The clock stage provides excellent and simple price discovery. The simplicity of the process is enhanced by generic lots.
- *Activity rule.* A simple yet powerful activity rule improves price discovery. As prices increase, a bidder cannot increase the size of the package.
- *Supplementary bids.* At the end of the clock stage, each bidder can improve its clock bids, as well as bid on any other packages that appear desirable given the information revealed in the clock stage. For packages that are larger than the bidder’s final clock bid, the supplementary bid has an upper bound. In particular, the supplementary bid can be no more than the package price at the clock prices from the round in which the bidder first shifted to a package that is smaller than the supplementary package.
- *Assignment stage.* The assignment stage translates the generic winnings of the clock stage into a specific assignment for each winning bidder. This is a sealed-bid combinatorial auction in which each bidder may submit top-up bids for each feasible package consistent with the bidder’s winnings. The value-maximizing assignment is determined.
- *Pricing rule.* Incentives for truthful bidding are enhanced through the use of “second pricing” (bidder-optimal core pricing). This rule is used both to determine base prices for the generic assignments and for the additional payment for the specific assignment.

The design is both highly practical and theoretically sound. Although the design is complex, my view is that it is as simple as possible, given the complex problem that the auction is asked to solve. Any further simplifications would compromise one or more of the objectives of the auction. In addition, much of the complexity of the rules has the effect of simplifying bidder decision-making. The design allows each bidder to focus on its demand for paired and unpaired spectrum, rather than complex gaming strategies.

I have conducted a series of mock auctions for Ofcom, the UK regulator, to test the auction design. The mock auctions demonstrated the high efficiency of the package clock auction. Twelve auctions were conducted with realistic scenarios and well-motivated and experienced bidders. In eleven out of twelve mock auctions, full efficiency was achieved. In the single inefficient mock auction, the inefficiency was the result of bidder error—the failure of a single bidder to submit her supplementary bids. Both the assignment and prices were identical to the theoretical benchmark. Base prices were at or close to opportunity cost (Vickrey prices) in all cases, despite some complementarities.

The scenarios were constructed to span a wide range of outcomes. In some, the paired bidders were so strong that the unpaired bidders won the minimum quantity of 9. In others, the unpaired bidders were so strong that all the spectrum went to unpaired bidders. In still others, there was a good mix of paired and unpaired winners.

I evaluated the mock auctions with respect to six measures.

- *Efficiency.* The outcomes were highly efficient in a full-scale mock auction with realistic preferences and well-motivated subjects.
- *Robustness.* The desirable properties of the package clock auction were robust to different levels of competition and different demand structures.
- *Risk.* Variation in outcomes was the result of different valuation models. High efficiency and other desirable properties were observed in all mock auctions.
- *Simplicity for bidders.* Subjects were able to understand the auction format and participate in the full-scale mock auction after only a few hours of training. Strategic considerations were easy to manage.
- *Revenues.* The mock auctions achieved competitive revenues. There was little tendency to overbid or underbid.
- *Simplicity for the auctioneer.* The full-scale mock auctions were readily conducted. The mock auctions demonstrated the feasibility of implementation.

The success of these mock auctions leads me to conclude that the package clock auction should perform well in the 2.6GHz environment, as well as related environments. As a result, I expect it to be adopted more broadly.

4.3 Competition issues

Assigning the spectrum to those companies with the highest value goes a long way to fostering innovation. This has been readily apparent over the last fourteen years of spectrum auctions. Nonetheless, there is good reason to think that there may be a divergence between the bidder's private value and its social value of the spectrum, as a result of the bidder's current position in the market. Incumbents typically have an incentive to limit competition. To avoid foreclosure of new entry, the market design needs to address this important asymmetry between incumbents and new entrants (Jehiel and Moldovanu 2000, Cramton et al. 2007).

Since oligopoly rents fall as competition increases, an incumbent has an incentive to foreclose new entry. The incumbent's value consists of its economic value of incremental spectrum plus the value of deterring entry. Since an entrant's value is just its economic value of the spectrum, it is certainly possible for the entrant to have a higher economic value for the spectrum, yet the incumbent is willing to pay more, as a result of its foreclosure value. This potential inefficiency is strongest when there are only a few strong incumbents. Then the benefits of foreclosing entry are greater and the ability of the incumbents to coordinate on an entry blocking strategy are greater.

The first US broadband auction provides an excellent illustration. At the time of that auction at the end of 1994, there were two cellular operators in each US market. The auction sold two additional large licenses in each market. The FCC understood the importance of competition in

the downstream market for wireless services, and took essential steps to enhance competition. Specifically, the auction rules forbid either cellular incumbent from bidding on the licenses, and an entrant could win at most one in any geographic market. As a result, the auction resulted in two new entrants in each market.

If instead, had the FCC ignored the competition issue, then I suspect that each incumbent would win one large license in each of the markets. Doing so would maintain the substantial duopoly profits, and “each incumbent wins one” is an obvious strategy for coordinating the foreclosure of entry by the duopolists. Indeed, it is unlikely that the incumbents would face much competition in the auction, since potential entrants would see the futility of bidding against the incumbents.

This restriction on incumbents, implemented as a spectrum cap, was arguably the most important feature of the initial market design in fostering competition and innovation in the downstream market. Moving from a duopoly to four competitors had a dramatic impact in reducing prices and enhancing services.

Interestingly, in the 700 MHz auction, concluding March 2008, no restrictions were placed on how much any incumbent could win in any market. The result was largely predictable (Cramton et al. 2007). The two strongest incumbents, Verizon and AT&T, bid for and won the lion’s share of the spectrum (85% by value). This spectrum was particularly valuable to Verizon and AT&T because of their incumbent position of holding nearly all the cellular spectrum located at 850 MHz. In contrast, the other operators such as Sprint and T-Mobile rely on PCS and AWS spectrum, which is at higher frequencies. Thus, Verizon and AT&T, who held a duopoly position in low-frequency spectrum before the auction, retained their duopoly position after the auction. This position gives Verizon and AT&T a competitive advantage in providing nationwide coverage. Customers value coverage, and the low-frequency spectrum allows for much less expensive geographic coverage. Dramatically fewer cell sites are needed, since the low-frequency signals travel farther. Importantly, Verizon and AT&T did not need the new spectrum to provide nationwide coverage with low-frequency spectrum—both already had that capability with their current cellular licenses. Rather, winning the 700 MHz spectrum *prevented* any other competitor from acquiring this same capability. My fear is that this outcome will further cement the dominant position of Verizon and AT&T in the market for wireless communications. Competition and innovation may suffer as a result.

One countervailing force is the “open access” provision included in the C Block won by Verizon. At least in theory, this provision opens the Verizon network to any device and any application. In the past, each operator tightly restricted the devices that could operate on its network and what applications would be allowed and how they could be used. This level of control enables the operator to extract additional economic rents from the network, but limits innovation. Any party with a good idea for a device or an application has to negotiate with the network operator first. In contrast, the open access provision pushes the industry closer to the Internet model, where the network operator is not able to control devices or applications. Any nondestructive device or application is allowed. Including the provision on just one block may suffice, provided competitive pressures cause AT&T and others to adopt Verizon’s open access policy.

Regulators have used a variety of tools to address competition issues in spectrum auctions. The most direct are spectrum caps, set-asides, bidding credits, and installment payments. In

many instances these instruments have been effective at promoting entry and enhancing competition. At other times, the instrument resulted in undesirable and unintended consequences. For example, in the second US broadband auction, which was a set-aside auction for small businesses, the FCC used overly attractive installment payments. The result was speculation, default, and delayed use of the spectrum. Installment payments were subsequently eliminated. In other instances, bidding credits led to the formation of fronts closely aligned with incumbents.

Establishing desirable competitive interventions is challenging. Still it is a challenge that the regulator must face with careful analysis and judgment. The interventions are often of first-order importance to successful market outcomes.

5 Electricity markets

Our final example of market design is electricity markets. Modern electricity markets are organized as a number of auction markets. The markets taken together are designed to provide reliable electricity at the least cost to consumers. Spot markets determine how much each supplier is generating on a minute-by-minute basis; forward energy markets enable customers and suppliers to lock in medium-term (1 to 3-years) prices for electricity; and long-run investment markets coordinate new entry to cover any expansion in electricity demand. These auction markets must be carefully designed to work together in achieving the goal of least-cost reliable supply. Design failures can be quite costly as the California electricity crisis of 2000-01 demonstrated. When the stakes are high, an important step in market design is building prototypes and then testing those prototypes in the experimental lab or in the field before full-scale implementation.

Electricity markets have a number of complicating features, which makes their design especially challenging. It is not possible to address all of these complications here. Rather, I will focus on some particular features of designs for the medium-term and long-term markets, and argue how these markets can work together to better achieve the goal of least-cost reliable supply.

5.1 Long-term: forward reliability market

In current electricity markets, the demand side has no way to express its preferences for reliability. For this reason, in most markets the regulator has taken on the role of assuring that there are adequate generating resources. One recent approach with many desirable properties is the forward reliability market, adopted in New England and Colombia (Cramton 2006, Cramton and Stoft 2007, 2008). In this market generating resources are procured on an annual basis well in advance (three years or more) of the eventual need. By lining up capacity well in advance of the need, new entry can compete to provide capacity. This has two important benefits. First, the market coordinates new entry and avoids the boom/bust cycles that are common with uncoordinated entry. Second, it makes the market for capacity contestable, and lets competitive new entry set the price for capacity.

A key element of the design is the definition of the capacity product, which I call a reliability option. It is physical capacity bundled with a financial call option to supply energy above a strike price. The physical capacity assures that there are adequate generating resources; whereas, the financial call option provides efficient performance incentives. Capacity resources face a financial obligation to supply energy whenever the spot price is above the strike price.

This obligation follows load in that it is stated as a share of the actual demand in the hour. In this way, the capacity payment fully hedges load from high spot prices, and reduces supplier risk as well. Market power is reduced in the spot market, since suppliers enter the spot market with a nearly balanced position in times of scarcity. Market power in the reliability market is addressed by not allowing existing supply to impact the capacity price. The approach is readily adapted to either a thermal system or a hydro system.

Importantly, the hedge in the form of a reliability option does not distort a supplier's marginal incentives in the spot market. Since any deviation from the obligation is priced at the spot price, the supplier receives the spot price on the margin and is motivated by the spot price. An important variation for load with hourly meters is to hedge load for its expected demand, rather than its actual demand, thus exposing load to the spot price on the margin to preserve incentives for demand response during scarcity. The impact of the option is to hedge both load and suppliers from price volatility above the strike price. This approach greatly reduces the risk from weather related price fluctuations. It does not however reduce performance risk. Although all risk is costly and hence undesirable, performance risk cannot be eliminated without eliminating the performance incentive.

The second advantage of the option is that it greatly reduces market power in the spot market during times of scarcity. Whenever the price exceeds the strike price of say \$300/MWh, the supplier obligation puts the suppliers in roughly a balanced position. A supplier in a balanced position has no incentive to distort its offer away from marginal cost. This reduction in market power improves the efficiency of the real-time dispatch as well as the spot price signal.

A third advantage of the hedge is that it makes high scarcity prices politically acceptable. High prices during scarcity are essential to motivate performance. A pervasive problem in nearly all electricity markets has been prices that are too low in times of true scarcity. A hedge enables the spot price to go much higher during scarcity without political backlash.

The capacity auction sets the payments to generators for providing reliability options just high enough to induce optimal investment and adequate capacity. An annual auction is used to purchase new capacity up to the level required for reliability. These auctions determine the price of reliability options that is just sufficient to induce the required entry. If the cost of constructing new capacity increases or decreases, due to environmental restrictions or new technology, new entrants will bid just enough higher or lower to maintain a normal rate of return, assuming there are no barriers to entry.

The result is that the regulator controls the level of capacity, but the market controls the price of capacity and the type and quality of capacity built. Hence the regulatory intervention is limited to determining the one factor about which the market has little information—the adequate level of capacity.

Although the auction design requires care to address the potential exercise of market power, the following simple procedure would work well. Each September an auction is held for reliability options, ROs, which take effect on January 1, just over three years in the future. Existing generators may choose either to enter the auction as a price taker (that is, with a zero bid), or not to sell ROs. New projects are allowed to bid without restriction. The regulator bids a demand curve that intersects the target adequacy level at the most recent RO price. The auction is held using a descending clock procedure. All accepted bids are paid the clearing price, but existing generation receives one-year contracts while new generation may choose any contract

length up to ten years. Once a new generator's initial contract expires it becomes an "existing" generator. If no new generation is purchased in a given year, all existing generators that bid, have their contracts extended for one year.

All generation, new and existing, will want to sell reliability options for their full capacity because these options fetch a high price relative to the financial cost of the option.

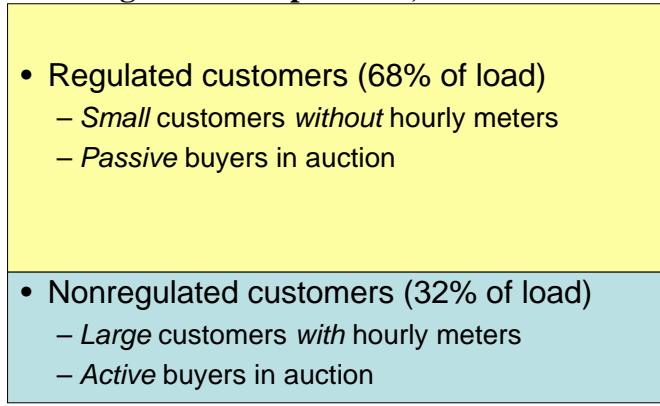
The benefits of this design are significant. The design minimizes risk and market power, while coordinating efficient entry.

5.2 Medium-term: forward energy market

Forward contracts for energy play an important role in reducing risk and market power in electricity markets. The reliability options, discussed above, provide price coverage above the strike price, but forward energy contracts are still needed to provide price coverage below the strike price. In most markets these forward contracts are negotiated bilateral contracts. These informal markets often are fragmented with little liquidity, nonstandard contracts, and high transaction costs. An organized market with standard contracts can increase competition and liquidity, and reduce transaction costs. Such an approach has recently been proposed for Colombia (Cramton 2007). The design summarized here, illustrates again the importance of the product design to achieve the objectives of the market.

The forward energy market is an organized market to procure energy for electricity customers on a forward basis. It includes both the regulated market (residential and other small customers) and the nonregulated market (large customers). Currently, regulated customers represent 68% of the total electricity demand and nonregulated customers represent the remaining 32%. The proposed design is novel in that it integrates both the regulated and nonregulated customers into a single organized market. Although the regulated and nonregulated energy products remain distinct, their integration into a single market facilitates arbitrage between the products, improves liquidity, and reduces transaction costs. Both regulated and nonregulated customers benefit from this unified approach.

Figure 1. Two products, one market



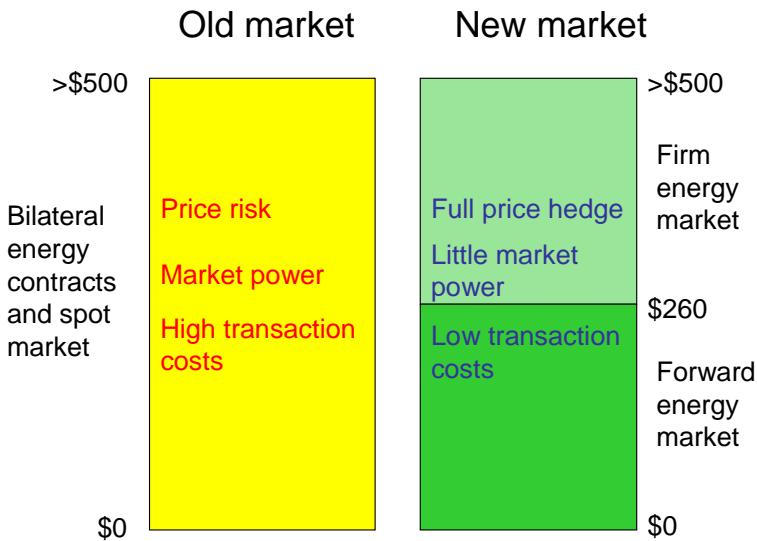
- Regulated customers (68% of load)
 - Small customers *without* hourly meters
 - Passive buyers in auction
- Nonregulated customers (32% of load)
 - Large customers *with* hourly meters
 - Active buyers in auction

As shown in Figure 1, the two customer groups, regulated and nonregulated, are integrated into a single market. Regulated customers are small customers without hourly meters; nonregulated customers are large customers with hourly meters. The nonregulated product will make use of the hourly meters to encourage demand response. In addition due to their large size, nonregulated customers will be active buyers in the forward energy market, submitting demand

bids. In contrast, the regulated customers will have a more limited demand response capability and will not be active buyers in the forward energy auction—their demands will be set administratively.

The proposed market is based on two load-following products, a regulated product and a nonregulated product. For the regulated product, each supplier bids to serve its desired share of Colombia's regulated load. A supplier that wins a 10% share at auction has an obligation to serve 10% of the actual regulated load in every hour of the commitment period. The supplier is paid the clearing price for every MWh of energy supplied. Deviations between the supplier's hourly supply and obligation are settled at the spot energy price or the scarcity price, whichever is lower. The spot settlement price is capped at the reliability options strike price, since the firm energy market provides price coverage for prices above the strike price (about \$260/kWh in January 2007 Colombian pesos; or US\$120/MWh); see Cramton and Stoft (2007). The nonregulated product is essentially the same, except each supplier bids to serve its desired share of the nonregulated load.

Figure 2. Price coverage of regulated customer



One-hundred percent of regulated load is purchased on behalf of the regulated customers in a sequence of auctions. Thus, the forward energy market together with the firm energy market provides 100% price coverage for all regulated customers, as shown in Figure 2. The forward energy market provides price coverage from zero to the scarcity price, and the firm energy market provides price coverage above the scarcity price. This accomplishes two things: 1) it provides rate stability for regulated customers, and 2) it provides revenue stability for suppliers. The result is reduced risk for both sides of the market.

The prior approach of bilateral contracts and the spot market suffers from three problems. Price risk is greater since the contract cover is incomplete. Market power in both the spot market and the bilateral market is more of a concern, since supplier positions are more apt to be out of balance entering the spot market and competition for bilaterals is weaker with specialized local products. The absence of a standard contract also results in high transaction costs in addition to weaker competition. In contrast, the new market provides full price coverage, suppliers enter the spot market with nearly balanced positions, and the single product minimizes transaction costs. In addition, the problem of self-dealing between the utility and its supplier affiliate is eliminated.

The market is mandatory for regulated customers, but voluntary for suppliers. Mandatory participation on the demand side motivates robust participation on the supply side.

The nonregulated customers enjoy similar benefits from the forward energy market as the regulated customers. The main difference is that the nonregulated customers actively bid in the forward energy market, and thus decide how much contract cover to purchase and in which auction. Although nonregulated customers participate voluntarily, I believe most will decide to participate fully and will adopt procurement strategies that do not differ too much from those of the regulated customers. Since the regulated and nonregulated products are close substitutes, the active participation of the nonregulated customers should yield improved pricing for both products.

An enormous simplification is having only a single product for each customer group, regulated and nonregulated. Although differences in load shapes across customers means that the cost of serving each customer is different, these differences are relatively minor, and can be accounted for with a customer-specific load-shape adjustment to avoid any cross subsidies that otherwise would result. Notice that the adjustment only impacts the load side and not the suppliers, since each supplier is serving a share of the aggregate load for the particular customer group.

The energy-share product enables load to be fully covered with a single product. For a supplier, the load-following product is natural, since in aggregate suppliers *must* follow load. A supplier is able to manage its exposure to the spot energy price through its portfolio of resources and its portfolio of nonregulated energy contracts. Even for a small supplier without a portfolio of resources or energy contracts, the risk from spot-price exposure is modest.

The proposed product does an excellent job of rate stability. Regulated load is fully hedged from the spot price. This makes sense for customers without hourly meters and demand management systems. However, for large nonregulated customers hourly meters are required and demand response is encouraged. This is done by basing the nonregulated product on *expected* load, rather than *actual* load. The actual load contract (pay as demand) is based on the customer's actual load in each and every hour of the commitment period. In contrast, the expected load contract is based on the customer's expected (forecasted) load in each and every hour of the commitment period as specified by the nonregulated customer or, if not specified, as estimated from its historical load shape and estimated growth over the period.

The expected-load contract hedges price risk, yet still exposes the customer to the spot price on the margin, motivating demand response, and innovation in demand management systems.

There are a number of possible choices for the timing and frequency of auctions, and the duration of contracts. These three elements can be adjusted to manage price and credit risk, while minimizing transaction costs. One sensible approach is quarterly auctions of 2-year contracts, which are rolling on an annual basis. The use of 2-year contracts is consistent with the most common contract in Colombia's bilateral market. The approach is simple and yet provides broad time diversification, shielding customers from transient events. One-eighth of regulated load is purchased in each auction. At any one time, two products are active and the customer rate reflects the average of eight auctions equally spaced over a two-year period. Even the auction with the shortest planning period occurs five months before the start of the contract. This means that the auction price will be set before there is much resolution of how severe conditions will be

in the following dry season. I believe that this structure strikes the right balance between risk reduction and the cost of guarantees to assure performance.

Efficient price formation is one of the most important objectives of the forward energy market. The simultaneous descending clock auction is used to promote efficient price formation. The descending clock auction provides excellent price discovery and enables suppliers to freely arbitrage across the regulated and nonregulated products. This assures that any price difference between the two products is a reflection of cost differences.

The integration of the regulated and nonregulated markets will lead to greater liquidity, improved price formation, and lower transaction costs. My view is that the forward energy market as proposed here will dramatically improve the energy contract market for both regulated and nonregulated customers, and improve the spot market as well, since suppliers typically will enter the spot market with a nearly balanced position, eliminating incentives to exercise market power.

At first glance, it might appear that Colombia's earlier design of informal bilateral contracts would enhance innovation. It did lead to hundreds of different contract types. But this heterogeneity fragmented markets and reduced competition to the detriment of the consumer. Rather, by standardizing products and relying on one central market, we enhance competition and liquidity, and improve price signals. This fosters the right kind of innovation—innovation for cost reduction on the supply side and demand response on the load side.

6 Conclusion

I presented four examples of how market design fosters innovation. In all cases, a basic ingredient is harnessing the power of markets and prices to motivate decentralized decision-making. Effective pricing of scarce resources drives market participants to make good decisions in both the short and long term. Innovation flows naturally from the incentives of prices, allowing flexible and creative responses to managing resource use.

In the case of air slots, spectrum, and electricity, a second ingredient in fostering innovation is a market design that enhances competition. Competition can be especially difficult to maintain in network industries, since network constraints tend to limit substitution and introduce scale economies and other nonconvexities in production. A good market design promotes competition in a number of ways, such as transparency of prices, improved liquidity, enhanced substitution, and reduced transaction costs. A fragmented market can be unified with standard products and a centralized auction. Sometimes competition is directly managed in the design with quantity caps that limit the size of any party, or set asides or bidding credits that favor new entrants. Competition inspires innovation, and good market design enhances competition.

A final ingredient in fostering innovation with market design is addressing other potential market failures. The free-rider problem must be addressed in markets for pollution and runway use. Coordination failures of new entry are seen in electricity markets. Another potential failure is overcoming economies of scale in providing demand management systems for electricity.

One exciting aspect of market design is working on the forefront of theory and yet bringing that theory to practice. In the auction applications discussed here, solving real problems has proved to be an excellent way to develop new theory. Nonetheless, the goal of market design is

not theory but practice: making markets work better through designs that promote the efficient use of scarce resources, enhance competition, and encourage innovation.

References

Ausubel, Lawrence M. and Peter Cramton (2002), "Demand Reduction and Inefficiency in Multi-Unit Auctions," University of Maryland Working Paper 9607, revised July 2002.

Ausubel, Lawrence M. and Peter Cramton (2004), "Auctioning Many Divisible Goods," *Journal of the European Economic Association*, 2, 480-493, April-May.

Ausubel, Lawrence M., Peter Cramton, and Paul Milgrom (2006), "The Clock-Proxy Auction: A Practical Combinatorial Auction Design," in Peter Cramton, Yoav Shoham, and Richard Steinberg (eds.), *Combinatorial Auctions*, Chapter 5, 115-138, MIT Press.

Ball, Michael O., Lawrence M. Ausubel, Frank Berardino, Peter Cramton, George Donohue, Mark Hansen, and Karla Hoffman (2007), "Market-Based Alternatives for Managing Congestion at New York's LaGuardia Airport," in *Optimal Use of Scarce Airport Capacity, Proceedings of AirNeth Annual Conference*, The Hague.

Cramton, Peter (2006), "New England's Forward Capacity Auction," Working Paper, University of Maryland.

Cramton, Peter (2007), "Colombia's Forward Energy Market," Working Paper, University of Maryland.

Cramton, Peter and Jesse Schwartz (2002), "Collusive Bidding in the FCC Spectrum Auctions," *Contributions to Economic Analysis & Policy*, 1:1, www.bepress.com/bejep/contributions/vol1/iss1/art11.

Cramton, Peter, Andrzej Skrzypacz and Robert Wilson (2007), "The 700 MHz Spectrum Auction: An Opportunity to Protect Competition In a Consolidating Industry" submitted to the U.S. Department of Justice, Antitrust Division.

Cramton, Peter and Steven Stoft (2008), "Forward Reliability Markets: Less Risk, Less Market Power, More Efficiency," *Utilities Policy*, forthcoming.

Cramton, Peter and Steven Stoft (2007), "Colombia Firm Energy Market," *Proceedings of the Hawaii International Conference on System Sciences*.

Ellerman, A. Denny, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M. Bailey (2000), *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge: Cambridge University Press.

Jehiel, Philippe and Benny Moldovanu (2000), "Auctions with Downstream Interaction Among Buyers," *Rand Journal of Economics*, 31, 768-791.

Milgrom, Paul (2004), *Putting Auction Theory to Work*, Cambridge: Cambridge University Press.

Porter, David, Stephen Rassenti, Anil Roopnarine, and Vernon Smith (2003), "Combinatorial Auction Design," *Proceedings of the National Academy of Sciences*, 100, 11153-11157.