

Using Economic Experiments to Test Electricity Policy

The industry's history of central generation, coordination, and regulation breeds a natural suspicion of whether or not decentralized coordination and a more market-based, decentralized regulatory approach can work. To see how people will behave in a decentralized environment with decentralized institutions, one must test the environment and institutions experimentally, with real people.

Lynne Kiesling

Lynne Kiesling is Director of the Center for Applied Energy Research at the International Foundation for Research in Experimental Economics (IFREE), as well as a Senior Lecturer in the Department of Economics at Northwestern University and a Senior Research Scholar at the Interdisciplinary Center for Economic Science at George Mason University. The author expresses her gratitude to Michael Giberson, Kent Lassman, and Tim Mount for their valuable contributions and comments.

Careful application of the scientific method can provide tremendous results to real-world problems. Scientists fight disease and inventors create new products. All over the world, students learn to observe, hypothesize, predict phenomena related to the hypothesis, and finally, to test the predictions and hypothesis with a controlled experiment. In the realm of public policy, experimental economics has applied this same formula successfully for decades. It is time to apply this tool to the discovery and testing of novel regulatory institutions. Specifically,

advances in technology and in economic theory have made decentralized regulatory institutions possible and more beneficial, and experimental economics provides the best set of tools for finding and investigating decentralized institutions that will be effective and valuable.

I. Experimental Test-Bedding of Policy

Experimental economics is a relatively young field, having grown following Vernon Smith's seminal work in the early 1960s.¹

Over the past decade, experimental economics has increasingly proven valuable as a methodology for test-bedding policy changes before implementing them. Particularly in complex network industries like electric power, the ability to create a “policy wind tunnel” can save ratepayers millions of dollars in costs, as well as reduce the stress of uncertainty for customers, firms, and regulators when new policies are implemented.

Experimental economics puts real human beings in a controlled, laboratory environment in which they must make decisions. Using this laboratory environment and profit-motivated human participants who get to keep what they earn, experimental economists evaluate market designs or institutions relative to a set of theoretical benchmarks. The most common benchmark is a measure of efficiency: Do the market participants generate all of the possible surplus (gains from trade) available in the environment? Another valuable measure for comparing two market designs is price volatility: Do prices move in a different range in the two designs? Evaluating efficiency and price volatility in conjunction can help policy-makers choose more efficient, and fair, designs.

One of the greatest advantages of experimental economics over other forms of testing market designs is this ability to construct a hypothetical environment with a theoretical benchmark and

observe how well a policy performs relative to that benchmark. We cannot observe theoretical benchmarks in real-world field testing. Experimental testing thus complements system-level computer simulations that are common in the industry by generating knowledge about how real humans with profit incentives will behave in the proposed market environment, and by catching design flaws and allow-

Experimental economics puts real human beings in a controlled laboratory environment in which they must make decisions.

ing correction before the market is implemented.

The environment created and the institutions tested in an experiment are carefully designed, and depend on the nature of the question. If the question is how well a particular mechanism performs, such as whether or not a uniform-price auction rule in a wholesale market performs better than a discriminatory, or pay-as-bid, auction rule, then the testing of those rules can start with a simple experiment and add complications (such as market power, or network congestion that leads to the divergence of nodal prices) until

the organization responsible for the rule (such as a regulator or an RTO/ISO) is comfortable enough with the performance of the mechanism to start the field implementation process.²

Another powerful benefit of experiments with human participants is their ability to capture some cognitive realities of the interaction of humans and physical infrastructure in complex networks. In general it is cognitively impossible for individuals to internalize and behave in accordance with the kind of complicated strategy spaces that game-theoretic economic theory relies on to drive its results. That does not mean, though, that individuals behave in irrational or random ways, but instead it means that we use heuristics and rules of thumb to govern our decision-making. Furthermore, in many instances we rely on tacit knowledge, or on decision rules that we do not even realize we are following. Despite these limitations, markets with robust, transparent institutions enable experiment participants to create substantial value, and thus the institutions can perform well, even in the face of human cognitive reality. In fact, in market design experiments participants often discover and exploit design flaws to their own benefit, even though they have limited information and cognitively limited strategy spaces. People also learn intuitively how to participate in markets. Economic experiments can incorporate the cognitive realities that humans bring to the

market, providing a valuable analysis of the performance of a mechanism or institution. Experiments do this precisely because they allow the expression of strategies that the experimenter (or computer simulation author) could not have predicted.

Laboratory experiments with human subjects differ from, and can be complementary to computer simulations and “wargaming” style scenario simulations. In computer simulations, agents behave in ways that they are programmed to behave; in other words, the model by which agents respond to each other is one that has been foreseen by the researcher and has been programmed into the environment. Human experiments allow individuals to follow strategies and behave opportunistically in ways that the researcher could not anticipate, and thus could not program into a simulation. It is precisely this type of strategic behavior that is of the most interest to policymakers when analyzing market design, and only human subject experiments will reveal that information in advance and allow revision of the design if needed. Scenario simulations that place, for example, policymakers in a trading environment can provide valuable learning opportunities and help us to understand better the situations in which people make particular decisions. One difference between experiments and scenario simulations is often the payoff to the

people involved. In experiments, the subjects get paid precisely what they earn, so the incentives remain pure; in scenario simulations the payoffs are not usually that clear, so individuals may not always “play for keeps” in the simulation. The three different methodologies can be useful in conjunction in corroborating economic theories and testing different market designs in different ways.

Lab experiments with human subjects can be complementary to computer simulations and “wargaming” style scenario simulations.

Suppose that instead of a simple comparison of two different market designs the question is the feasibility of a particular mechanism. Then the focus is on proof-of-concept, and some additional environmental complexity is valuable in the experiment, because the emphasis is on the ability to implement such a policy in a realistic environment and have it perform well. One example of the latter is the creation of new markets for transactions that have typically taken place within vertically integrated utilities, but may be valuable, both for system reliability and gains from trade, if the buyers and

sellers could transact through markets instead. Reactive power is one such resource; it is scarce, has value to buyers and sellers through their direct benefits and through its effect on system reliability, and thus could be a candidate for transacting through markets. Yet without some proof-of-concept experiments, the feasibility of reactive power markets remains unknown, and we perpetuate our use of regulatory and centralized ways of providing and paying for reactive power.

Experimental testing of market designs has been useful in a variety of policy applications ranging from simple tests of known mechanisms to more complex analyses of new mechanisms and institutions. Murphy *et al.* used computer-assisted “smart market” algorithms in testing different market designs for water allocation in the California water network.³ They found that the smart market algorithm coordinated the decisions of buyers, sellers, and transporters and generated highly efficient outcomes. Charles Plott reports on a test of a simple sealed-bid auction mechanism for allocating land at airports.⁴ More complicated applications of experiments to policy questions include the design of markets for the Federal Communications Commission’s auctions of spectrum licenses.⁵ Experiments have also been performed on unimplemented policies, such as the design of algorithms for allocating airport takeoff and landing slots.⁶

II. Experiments and Electricity Policy

Electric power provides a valuable example and an opportunity to use experiments to improve policy, as well as to improve and refine the economic theory that helps us understand complex network industries. The unique features of electricity and its non-storability make markets for electricity behave very differently from the typical markets analyzed by economists. These differences explain to some extent why it has been relatively difficult to design robust and efficient markets under the current move to deregulate electric utilities. Regulators have not followed the standard principles used by engineers of testing and verifying the design of a new product. In many cases, new untested market designs for electricity have been implemented and, when problems occurred, a series of patches were adopted to correct them. The initial market used in California is the obvious example. In contrast, the PJM market was implemented in a systematic series of steps, and the viability of each step was established before moving to the next one.

Experiments have clearly been an effective tool for evaluating and modifying policy changes since the first electricity experiments performed by Vernon Smith and his colleagues for the Arizona Corporation Commission more than two decades ago. In 1984 the Commission

requested an experimental analysis of alternatives to rate-of-return regulation of utilities. One aspect of this research involved the use of a smart market computer algorithm to perform economic dispatch while aggregating seller offers and buyer bids. This computerized dispatch would facilitate and decrease the cost of implementing double-sided wholesale and retail markets, and this pioneering study provided an

Experiments clearly have been an effective tool for evaluating and modifying policy changes since the first experiments performed by Vernon Smith.

important proof-of-concept exercise.

In more recent experimental work from Smith's group, Rassenti, Smith, and Wilson tested the effect of demand response and double-sided markets on the extent to which generators could exercise market power.⁷ They compared two bidding systems in a wholesale electricity market – one with only supply-side bidding, and one with both supply-side and demand-side bidding. Their experimental design built in the effect of retail customer demand inelasticity on the wholesale market. The buyer in the whole-

sale market could be an integrated utility or retail electricity provider. They also had three different cost levels for generators: base-load (low-cost), intermediate cost "load followers," and high-cost peaking units. Their results indicated that demand-side bidding decreased price volatility, the magnitude of price spikes, and the ability of suppliers to exercise market power. Average prices were lower when both demand-side and supply-side bidding existed, as opposed to the one-sided supply-side bidding that characterizes wholesale electricity markets in the U.S. This research provides the clear policy implication that active demand plays a valuable role in disciplining the pricing choices of generators in wholesale markets.

In earlier experimental work on pricing in natural gas pipeline networks, Rassenti, Reynolds, and Smith tested a "smart market" computer algorithm for allocating pipeline capacity.⁸ This research, sponsored by the Federal Energy Regulatory Commission, tested a market institution with decentralized ownership of pipelines and centralized coordination of its dispatch through a real-time computerized auction. Such research provides a further example of how experimental test-bedding of market institutions can inform policy decisions.

Other scholars also perform important experimental research in electricity, through the Power System Engineering Research Center (PSERC) collaboration of engineers and economists at

Cornell University. Much of their work is used to test market designs at the New York Independent System Operator (NYISO).⁹ This collaboration has focused on testing the performance of smart markets that incorporate the physical restrictions of a transmission network on the production and delivery of electricity. The testing platform used for this research is called PowerWeb.¹⁰

PowerWeb is an Internet-based platform, with a Web browser interface, that simulates the operations of a power system with real human decision-makers participating in a market based on complex underlying engineering models of the network. This platform brings together the human behavior aspect of economic decision making with consistent economic and engineering models of a physical AC power system. With this framework, it is possible to investigate the effects of complex features of electricity supply, such as congestion on the network and the demand for ancillary services.

PSERC researchers have performed auction design experiments both with Cornell University students and with staff of the New York Department of Public Service. Some sessions compared a uniform price auction, in which all sellers receive and all buyers pay a single market-clearing price, with a discriminatory auction, in which sellers receive exactly the offers they submit. Theoretical analyses of these auction designs suggest that

the discriminatory auction gives sellers an incentive to submit higher bids, especially for low-cost units. The experimental sessions bore out the theoretical prediction, although average prices did not differ greatly between the two designs. Then the experimenters introduced active demand response, which increased efficiency and decreased price volatility in the uniform auction. In the discrimi-

Although it is impossible to discover all of them, many problems with a new market design can be identified and corrected before the new design is implemented.

natory auction, though, active demand did not reduce price spikes because of the ability of suppliers to flatten the supply curve and speculate by submitting overstated offers. PSERC researchers have also used these experiments to show why FERC's soft-cap auction in California in the winter of 2001 did not reduce price spikes: a soft-cap design behaves similarly to the discriminatory auction.

Laboratory experiments can be a cost-effective tool for testing market design proposals, particularly relative to the possibly large and unknown costs that could arise in testing new designs

in practice in the field. Market design matters can rapidly become complex, and it can be hard even for experts in the industry to choose well. As a practical matter, market design decisions are often made by committees of experts, as constrained by regulatory guidance and the occasional attention of politicians. In such an environment it is sometimes difficult to distinguish what will work from what we want to work. The point is not that experiments can answer all market design questions. Instead, experimentalists offer a more value-focused proposition: Experiments in the lab can help answer some questions at far less expense than experiments in the world. Laboratory experiments also allow for control of variables that could confound and complicate observation of results in real-world field experiments.

Although it is impossible to discover all the problems with a market design through testing, many problems can be identified and corrected before the new design is implemented. It is inevitable that the rules in a deregulated market will evolve over time. In the long run, however, it is still far less expensive to use experimental economics to correct some of the problems than to experiment directly on the public with an untested design in a real market. The "soft-cap" auction implemented in California by FERC in the winter of 2001 provides an example, creating additional problems that could probably have been avoided.

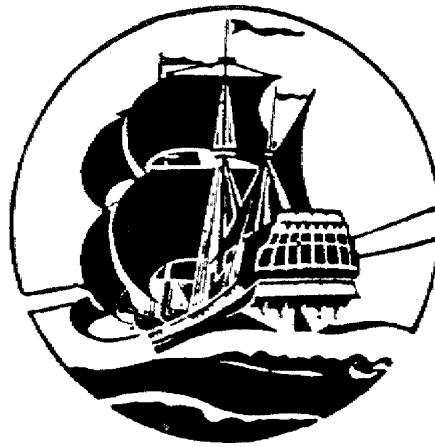
III. Experiments and Decentralized Regulatory Institutions

The value of experimental economics transcends the ability to test policy changes in existing environments. Experiments create a set of tools and an opportunity to test our thinking about policy as the world changes around us, in large part because in the laboratory we can create hypothetical environments that have not yet existed in reality. Experiments are a dynamic, forward-looking tool for understanding a dynamic, complex world.

Our current regulatory environment is a relic of the early Twentieth Century. Traditionally in the electric power industry we are used to the large-scale central generation paradigm and the centralized, top-down, rulemaking-based regulation that has accompanied it. The physical reality of a network of distributed customers with centralized sellers of power dictated strong centralized coordination of real-time operations and dispatch, of forward operations, and of capacity planning.

During the last several decades, technology has changed, potentially affecting the way we generate power, monitor power, and use power. The most obvious technological change thus far in electric power has been the combined-cycle gas turbine, which transformed the nature and scale of power generation. Information technology has changed how, and

how intensely, we use power, and in some instances it has enhanced our ability to monitor and understand our energy use, and to modify it in response to price changes. The reach of such technologies into the electricity industry is just beginning, though, notwithstanding the fact that digital information technology is over 50 years old. Innovation



in information technology has dramatically decreased monitoring costs, to the point where we can now shift control toward end users and away from the network. This shift is leading to the deployment of price-sensitive consumer appliances and complex, programmable systems, particularly for larger industrial and commercial customers.

Put more directly, technological change is inducing a paradigm shift in electric power, away from the central generation, top-down, rulemaking world and toward a more distributed, decentralized, yet still coordinated, network world. Digital technology for monitoring, for load control, and for voltage and frequency man-

agement are increasingly making distributed coordination and control feasible, and possibly even less costly than more centralized approaches have traditionally been.

Couple this technological change with the increasing role of markets and competition in electric power, and it becomes even more apparent that electric power is moving in a more distributed and decentralized direction. Borrowing from the economist Ronald Coase, in the past all transactions in the electric power supply chain occurred within vertically integrated firms. Increasingly, though, this value chain is changing, and includes both transactions within firms and transactions through markets. Again, technological change has contributed to the reduction in transaction costs that makes these market transactions more profitable than continuing the (regulated) within-firm transactions. As these economic and technological changes continue, we are likely to see an increasing number of transactions take place in markets, especially for aspects of electric power service such as power quality and reactive power. These changes naturally raise the question of whether a centralized regulatory approach still fits this decentralized, distributed, complex industry and environment. And if not, what kind of policies should replace it?

An open question that would benefit from experimental testing is the kind of regulatory institutions that would be most useful

and best suited to the increasingly decentralized electric network. The experimental research discussed above indicates that the central (physical and economic) coordination that remains necessary in the electric power network should be facilitated using a smart market algorithm. Beyond that centralized coordination, though, regulatory institutions that focus on reducing transaction costs that prevent private parties from negotiating are likely to be better suited to the technology-enabled distributed complexity of the electric power system. In other words, a regulatory approach that is more contractual and forward-looking and less oriented toward top-down control and management may protect customers and provide incentives for investment and innovation by creating a transparent legal platform from which distributed, heterogeneous private parties can mutually benefit. Modeling such a contractual regulatory institution and testing it experimentally against the existing institution would be a fruitful and valuable direction for experimental public policy research.

Experiments are a natural tool to use when considering policy changes in the midst of transitions such as the one currently happening in electric power. More than any other analytical tool available to us, experiments enable us to test what is not. Experiments put human beings in situations and policy environments that do not exist in reality and evaluate the

performance of the policy institution. Such a tool has enormous value in evaluating policy changes during the evolution from a centralized environment to a decentralized and distributed one. Experiments enable us to measure and get comfortable with the types of decentralized coordination and regulation approaches that are feasible and



beneficial in this period of change. No other analytical evaluation tool delivers those kinds of results.

IV. Conclusion

Experimental economics is a viable way to evaluate the performance of a new market design, and in complex markets, it is hard to predict the outcome of a test from economic theory. Electricity markets are a prime example of complex markets, and the results of tests are usually surprising in the sense that outcomes are different from expectations. In some cases, students have been able to find and exploit weaknesses in a

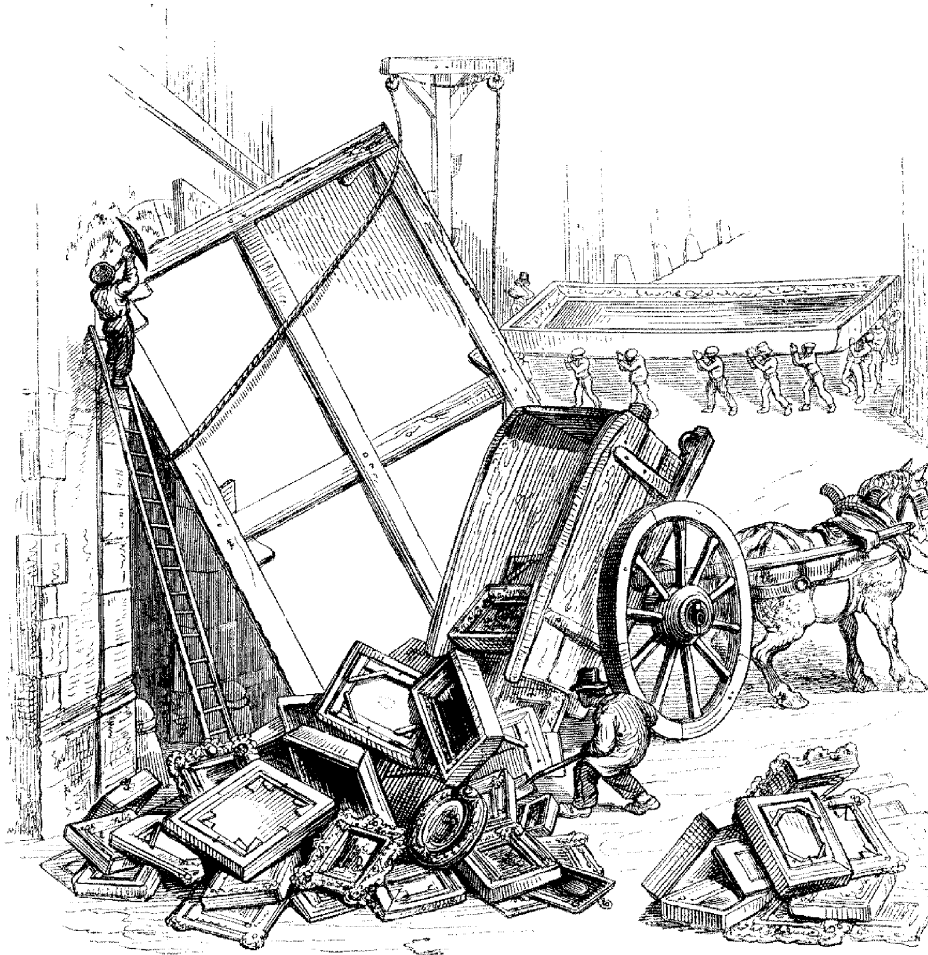
market design that the “expert” researchers had overlooked completely.

In the middle of a regulatory paradigm shift and complex markets, the ability to study what is not becomes even more important. Our history of central generation, coordination, and regulation in this industry breeds a natural suspicion of whether or not decentralized coordination and a more market-based, decentralized regulatory approach can work. If you want to see how people will behave in a decentralized environment with decentralized institutions, you have to test the environment and institutions experimentally, with real people. Only by doing that will you have the data to assess the feasibility and performance of decentralized, distributed networks and companion regulatory institutions. That claim is particularly true for evaluating the feasibility of new markets, such as differentiated power quality and reactive power. In other words, proposals for new decentralized regulatory and market institutions need the proof-of-concept testing that experiments provide. Without such testing, we are likely to retain our traditional top-down, centralized approach past its usefulness as the world dynamically changes and innovates around the electric power industry. We are understandably risk-averse regarding new regulations that may be worse than those we currently have. Experiments can help relieve our

uncertainty and provide policy-makers with a guide to effective action. ■

Endnotes:

1. Vernon L. Smith, *An Experimental Study of Competitive Market Behavior*, J. POLIT. ECON., Vol. 70 (1962): 111–137.
2. For further discussion of this point see Charles Plott, *Market Architectures, Institutional Landscapes and Testbed Experiments*, ECON. THEORY, Vol. 4 (1994): 3–10.
3. J.J. Murphy, A. Dinar, R.E. Howitt, S.J. Rassenti, and V.L. Smith, *The Design of "Smart" Water Market Institutions Using Laboratory Experiments*, ENV'L. & RESOUR. ECON., Vol. 17 (2000): 375–394.
4. Plott, *supra* note 2, at 7.
5. See Jeffrey Banks, Mark Olson, David Porter, Stephen Rassenti, and Vernon Smith, *Theory, Experiment, and the Federal Communications Commission Spectrum Auctions*, J. ECON. BEHAV. & ORG., Vol. 51 (2003): 303–350; for an experimental analysis of a United Kingdom spectrum license auction, see Charles Plott and Timothy Salmon, *The Simultaneous, Ascending Auction: Dynamics of Price Adjustment in Experiments and in the UK3G Spectrum Auction*, J. ECON. BEHAV. & ORG., Vol. 53 (2004): 353–383.
6. See David M. Grether, R. Mark Isaac, and Charles R. Plott, *The Allocation of Landing Rights by Unanimity Among Competitors*, AMER. ECON. REV., Vol. 71(2) (May, 1981): 166–171; see also S.J. Rassenti, Vernon L. Smith, and R.L. Bulfin, *A Combinatorial Auction Mechanism for Airport Time Slot Allocation*, BELL J. ECON., Vol. 13(2), (Autumn 1982): 402–417.
7. Stephen J. Rassenti, Vernon L. Smith, and Bart J. Wilson. *Controlling Market Power and Price Spikes in Electricity Markets: Demand-Side Bidding*, PROC. NATL. ACAD. SCI., Vol. 100 (2003): 2998–3003.
8. Stephen J. Rassenti, Stanley S. Reynolds, and Vernon L. Smith, *Cotenancy and Competition in an Experimental Market for Natural Gas Pipeline Networks*, ECON. THEORY, Vol. 4 (1994): 41–64.
9. For a recent example of experimental work from this group, see R.J. Thomas, T.D. Mount, R. Zimmerman, W.D. Schulze, R.E. Schuler, and L.D. Chapman, *Testing the Effects of Price Responsive Demand on Uniform Price And Soft-Cap Electricity Auctions*, PROCEEDINGS OF 35TH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (2002).
10. See www.e3rg.pserc.cornell.edu.



In complex markets, it is hard to predict the outcome of a test from economic theory.