Prologue

Is it a fact–or have I dreamt it—that, by means of electricity, the world of matter has become a great nerve, vibrating thousands of miles in a breathless point of time? Rather, the round globe is a vast head, a brain, instinct with intelligence! Or, shall we say, it is itself a thought, nothing but thought, and no longer the substance that we dreamed it?

Nathaniel Hawthorne
The House Of Seven Gables
1851

POWER SYSTEM ECONOMICS **PROVIDES A PRACTICAL INTRODUCTION TO POWER-MARKET DESIGN**. To assist engineers, lawyers, regulators, and economists in crossing the boundaries between their fields, it provides the necessary background in economics and engineering. While Part 1 covers basics, it provides fresh insights ranging from a streamlined method for calculations, to the adaptation of economics to the quirks of generation models, to the distinction between the market structure and market architecture.

Part 2 focuses on the core structure of power markets which determines the basic character of supply and demand. It encompasses demand-side flaws, short-run reliability policy and the rigidities of supply. Together these determine the notorious price-spikes and unstable investment pattern of power markets. Because of its fundamental nature, this analysis can proceed without reference to locational pricing or unit commitment which gives less-technical readers access to the most important and fundamental economics of power markets.

Part 3 discusses the architecture of the day-ahead and real-time markets. This requires the introduction and analysis of the unit-commitment problem—the problem of starting and stopping generators economically. To avoid unnecessary complexity, the other primary problem of power-system economics, network congestion, is postponed until Part 5. This allows a clearer comparison of the three fundamental types of power trading: bilateral trading, exchange trading and pool-based trading.

Part 4 detours from the drive toward an increasingly detailed view of the market to examine market power. Although best understood in the context of Parts 2 and

Chapter 1-1

Why Deregulate?

The propensity to truck, barter, and exchange one thing for another . . . is common to all men.

Adam Smith
The Wealth of Nations
1776

IN THE BEGINNING THERE WAS COMPETITION—BRUTAL AND INEFFICIENT. Between 1887 and 1893, twenty-four central station power companies were established within Chicago alone. With overlapping distribution lines, competition for customers was fierce and costs were high. In 1898, the same year he was elected president of the National Electric Light Association, Samuel Insull solved these problems by acquiring a monopoly over all central-station production in Chicago. In his historic presidential address to NELA, Insull explained not only why the electricity business was a "natural monopoly" but why it should be regulated and why this regulation should be at the state level, not the local level. Insull argued that

> *exclusive franchises should be coupled with the conditions of public control, requiring all charges for services fixed by public bodies to be based on cost plus a reasonable profit.*

These ideas shocked his fellow utility executives but led fairly directly to regulatory laws passed by New York and Wisconsin in 1907 establishing the first two state utility commissions. Reformers of the Progressive era also lent support to regulation although they were about equally supportive of municipal power companies.[1] Their intention, to hold down monopoly profits, was at odds with Insull's desire to keep profits above the competitive level, but both sides agreed that competition was inefficient and that providing electricity was a natural monopoly.[2]

1. See Platt (1991) for information on central station companies and Samuel Insull. The quotation (p. 86) is from a contemporary account of Insull's address. Platt describes early competition as follows: "The Chicago experience of rate wars, distributor duplication, and torn-up streets presented an alternative that was attractive to virtually no one." For state commissions and progressives, see Rudolph & Ridley (1986).

2. Smith (1995) relies on Gregg Jarrell to conclude "regulation was a response to the utilities' desire to protect profits, not a consumerist response to monopoly pricing." But Knittel (1999) tests causation by utilities and consumers and finds no significant correlation between profit change and regulation after correcting Jarrell's endogeneity problem. This result would be expected from an analysis of profit when two equal forces have opposite motivations with respect to its level.

Chapter *1-2*

What to Deregulate

Nothing is more terrible than activity without insight.

Thomas Carlyle
(1795-1881)

DELIVERED POWER IS A BUNDLE OF MANY SERVICES. These include transmission, distribution, frequency control, and voltage support, as well as generation. The first two deliver the power while the second two maintain power quality; other services provide reliability.

Each service requires a separate market, and some require several markets. This raises many questions about which services *should* be deregulated and which *should not*. Even within a market for a single service, one side—either demand or supply—may need to be regulated while the other side of the market can be deregulated. For instance, the supply of transmission rights must be determined by the system operator, but the demand side of this market is competitive. In contrast, the demands for ancillary services are determined by the system operator while the supply sides of these markets can be competitive.

The most critical service in a regulated or a deregulated power market is that provided by the system operator. This is a coordination service. For a deregulated market it typically includes operation of the real-time markets and a day-ahead market. These provide scheduling and balancing services, but operating these markets is itself an entirely separate service. While the need for the system operator service is agreed to by all, the proper extent of that service is the subject of the central controversy in power market design.

Chapter Summary 1-2: Many services are required to bring high-quality reliable power to end users. Each might be provided by free markets, by the state, by regulated suppliers, or by some hybrid arrangement. Bulk power generation is the source of nearly half the cost of retail power and is one of the services most easily provided by a competitive market. Moreover, it seems to offer several possibilities

Chapter 1-3

Pricing Power, Energy, and Capacity

It is not too much to expect that our children will enjoy in their homes electricity too cheap to meter.

Lewis L. Strauss
Chairman, Atomic Energy Commission
1954

POWER IS THE RATE OF FLOW OF ENERGY. Similarly, generating capacity, the ability to produce power is itself a flow. A megawatt (MW) of capacity is worth little if it lasts only a minute just as a MW of power delivered for only a minute is worth little. But a MW of power or capacity that flows for a year is quite valuable.

The price of both power and energy can be measured in $/MWh, and since capacity is a flow like power and measured in MW, like power, it is priced like power, in $/MWh. Many find this confusing, but an examination of screening curves shows that this is traditional (as well as necessary). Since fixed costs are mainly the cost of capacity they are measured in $/MWh and can be added to variable costs to find total cost in $/MWh.

When generation cost data are presented, capacity cost is usually stated in $/kW. This is the cost of the flow of capacity produced by a generator over its lifetime, so the true (but unstated) units are $/kW-lifetime. This cost provides useful information but only for the purpose of finding fixed costs that can be expressed in $/MWh. No other useful economic computation can be performed with the "overnight" cost of capacity given in $/kW because they cannot be compared with other costs until "levelized." While the U.S. Department of Energy sometimes computes these economically useful (levelized) fixed costs, it never publishes them. Instead it combines them with variable costs and reports total levelized energy costs.[1] This is the result of a widespread lack of understanding of the nature of capacity costs.

Confusion over units causes too many different units to be used, and this requires unnecessary and sometimes impossible conversions. This chapter shows how to make almost all relevant economic calculations by expressing almost all prices and

1. In Tables 14 through 17 of one such report (DOE 1998a) the useful (amortized) fixed costs are not reported, and the fixed O&M costs are reported in $/kW which may be an amortized value reported with the wrong units or, if the units are correct, may represent a misguided conversion of an amortized cost to an "overnight" cost.

Chapter *1-4*

Power Supply and Demand

And when the Rain has wet the Kite and Twine, so that it can conduct the Electric Fire freely, you will find it stream out plentifully from the Key on the Approach of your Knuckle.

Benjamin Franklin
1752

THE PHYSICAL ASPECTS OF SUPPLY AND DEMAND PLAY A PROMINENT ROLE IN POWER MARKETS. Shifts in demand, not associated with price, play a role in all markets, but in power markets they often receive attention to the exclusion of price. This is not simply the result of regulatory pricing; even with market prices, demand shifts will play a key role in determining the mix of production technologies. In this way hourly demand fluctuations determine key long-run characteristics of supply.

Because electric power cannot be stored, production always equals consumption, so the difference between supply and demand cannot be indicated by flows of power. Neither is the instantaneous difference indicated by contracts since real-time demand is determined by customers physically taking power. The short-run supply–demand balance is indicated by voltage and, especially, frequency. This unusual market structure requires some elementary background in system physics. More detail is provided in Chapters 5-1 and 5-2.

Chapter Summary 1-4: Load duration curves are still relevant in unregulated markets, but their role in analysis is more subtle because their shape is affected by price and its correlation with load. They can still be used with screening curves to check an equilibrium, but to predict an equilibrium they must be used in combination with price elasticity.

Power production always equals consumption (counting losses as part of consumption) which makes it impossible to assess the supply–demand balance by observing quantities or quantity flows. Instead, frequency is the proper indicator of system-wide balance, and net unscheduled flows between regions are used to share the responsibility of maintaining this balance.

Chapter 1-5

What Is Competition?

The rich, . . . in spite of their natural selfishness and rapacity, . . . though the sole end which they propose . . . be the gratification of their own vain and insatiable desires, they divide with the poor the produce of all their improvements. They are led by an invisible hand to make nearly the same distribution of the necessaries of life, which would have been made, had the earth been divided into equal portions among all its inhabitants, and thus without intending it, without knowing it, advance the interest of the society.

Adam Smith
The Theory of Moral Sentiments
1759

COMPETITION IS LEAST POPULAR WITH THE COMPETITORS. Every supplier wants to raise the market price, just as every buyer wants to lower it. Perfect competition frustrates both intentions.

Some commodity markets provide almost perfect competition; eventually power markets may work almost as well. But designing such markets is difficult. Economic competition is not like competition in sports, which may be considered perfect when there are just two powerful and equal competitors. Economists consider competition to be **perfect** when every competitor is small enough (*atomistic* is the term used) to have no discernable influence against the "invisible hand" of the market.

Adam Smith guessed intuitively that a perfectly competitive market, in the economic sense, would produce an outcome that is in some way ideal. Many difficulties can cause a market to fall short of this ideal, but even a market that is only "workably competitive" can provide a powerful force for efficiency and innovation.

Power markets should be designed to be as competitive as possible but that requires an understanding of how competition works and what interferes with it. On its surface, competition is a simple process driven, as Adam Smith noted, by selfishness and rapacity; but the invisible hand works in subtle ways that are often misunderstood. Those unfamiliar with these subtleties often conclude that suppliers are either going broke or making a fortune. This chapter explains the mechanisms that keep supply and demand in balance while coordinating production and consumption to produce the promised efficient outcome.

Chapter 1-6

Marginal Cost in a Power Market

The trouble with the world is not that people know too little,
but that they know so many things that ain't so.

Mark Twain
(1835–1910)

SIMPLIFIED DIAGRAMS OF GENERATION SUPPLY CURVES HAVE CONFUSED THE DISCUSSION OF MARGINAL COST. Typically, these supply curves are diagrammed to show a constant marginal cost up to the point of maximum generation. Then marginal cost becomes infinite without taking on intermediate values. Typically it jumps from about $30 to infinity with only an infinitesimal increase in output. Mathematics calls such a jump a discontinuity. In fact, the curve would be discontinuous if it jumped only from $30 to $40.

The definition of marginal cost does not apply only to the points of discontinuity. Hence it does not apply to a right-angle supply curve at the point of full output, neither does it apply to the points of a market supply curve at which it jumps from one generators marginal cost to the next. Unfortunately market equilibria sometimes occur at such points, and concerns over market power often focus on them. Attempts to apply the standard definition at these points can produce confusing and erroneous results.

Fortunately, the definition is based on mathematics that generalizes naturally to discontinuous curves. Applying this generalization to the textbook definition clears up the confusion and restores the economic results that otherwise appear to fail in power markets. For example, in power markets, as in all other markets, the competitive price is never greater than the marginal cost of production.

Chapter Summary 1-6: Individual supply curves are often constructed with an abrupt end that causes the market supply curve to have abrupt steps. The standard marginal-cost definition does not apply at such points. Instead, left- and right-hand marginal costs should be used to define the marginal-cost range. Then the competitive price, which remains well defined, will always lie within that range. A market price exceeding the marginal-cost range indicates market power.

Chapter 1-7

Market Structure

The work I have set before me is this . . . how to get rid of the evils of competition while retaining its advantages.

Alfred Marshall
(1842-1924)

POOR MARKET STRUCTURE POSES THE GREATEST THREAT TO THE HEALTH OF POWER MARKETS. "Structure" refers to properties of the market closely tied to technology and ownership. The classic structural measure is a concentration index for the ownership of production capacity. The cost structure of an industry, another component of market structure, describes both the costs of generation and the costs of transmission.

Most aspects of market structure are difficult to alter and some, such as the high fixed costs of coal-fired generation, are impossible. But power markets contain some unusual technology-based arrangements that can easily be altered or that require administrative decisions regarding their operation. These arrangements are part of the market structure and require design just as do the architectural components described in the next chapter.

The notion of **market structure** developed as part of the "structure-conduct-performance" paradigm of industrial organization in the early 1950s. The present discussion, however, is based on the structure-architecture-rules classification of market-design problems presented by Chao and Wilson (1999a) and Wilson (1999). The present chapter extends their definition of structure, particularly in the direction of administered reliability policies.

Chapter Summary 1-7: Market structure has a decisive impact on market power and investment. The second demand-side flaw, the ability of users to take power from the grid in real time without a contract (see Section 1-1.5), makes structural intervention necessary. Regulators must trade-off price spikes against involuntary load shedding, thereby largely determining the incentives for investment in generation capacity.

Chapter 1-8

Market Architecture

The whole world may be looked upon as a vast general market made up of diverse special markets where social wealth is bought and sold. Our task then is to discover the laws to which these purchases and sales tend to conform automatically. To this end, we shall suppose that the market is perfectly competitive, just as in pure mechanics we suppose, to start with, that machines are perfectly frictionless.

Leon Walras
Elements of Pure Economics
1874

A MARKET'S ARCHITECTURE IS A MAP OF ITS COMPONENT "SUBMARKETS." This map includes the type of each market and the linkages between them.[1] The submarkets of a power market include the wholesale spot market, wholesale forward markets, and markets for ancillary services. "Market type" classifies markets as, for example, bilateral, private exchange, or pool. Linkages between submarkets may be implicit price relationships caused by arbitrage or explicit rules linking rights purchased in one market to activity in another.

Architecture should be specified before rules are written, but it is often necessary to test the architecture during the design process, and this requires a rough specification of the rules. Architectural design must also consider the market structure in which it is embedded, which may inhibit the proper function of some designs. Market design should not be rigidly compartmentalized, yet it is useful to consider the market's architecture apart from the details of the rules and the limitations of market structure.

Chapter Summary 1-8: A market design or analysis project concerns a collection of "submarkets" which are collectively referred to as the "entire market." (Both will often be referred to simply as markets.) Deciding which submarkets should be created for a power market is the first step in architectural design. Section 1-8.1 briefly discusses day-ahead and real-time energy markets and transmission-rights markets as a prelude to Part 3 which examines these choices in more depth.

Private submarkets range from disorganized to highly centralized, and each has its advantages. There is no simple rule for choosing between types of submarkets

1. This chapter owes a great debt to Wilson (1999) and Chao and Wilson (1999a), though it is not intended as a summary of their views.

Chapter 1-9

Designing and Testing Market Rules

Genius is one per cent inspiration, ninety-nine per cent perspiration.

Thomas Edison
c. 1903

If Edison had a needle to find in a haystack he would proceed . . . to examine straw after straw. A little theory and calculation would have saved him ninety percent of his labor.

Nikola Tesla
New York Times
1931

UNTESTED MARKET DESIGNS CAUSE REAL-WORLD MARKET FAILURES. Suppliers are quick to take advantage of design flaws, especially those that pay $9,999/MWh for a product that is worth less than $5/MWh.[1] Currently, many if not most, market designs are implemented without any explicit testing.

Although the most serious market flaws typically arise from structural problems, while architectural problems rank second in importance, problems with rules are the most numerous and their cost can be impressive. The design of rules is more art than science, but economics offers two guiding principles: mimic the outcome of a classically competitive market, and design markets so competitors find it profitable to bid honestly. Simplicity is another virtue well worth pursuing but notoriously difficult to define.

Chapter Summary 1-9: In a pay-as-bid auction, a coal plant bidding its variable cost of $12/MWh would be paid $12/MWh, while in a single-price auction it would be paid the system marginal cost which might be $100/MWh. In this case many would object to paying the $100 competitive price to the "inexpensive" coal plant and seek to improve on the competitive model. Pay-as-bid is one suggestion. The result is gaming and, probably, a very modest decrease in price and a modest decrease in efficiency. Ironically, if pay-as-bid succeeded as its advocates hope, it would put an end to investment in baseload and midload plants. In the long run this would dramatically raise the cost of power. The pay-as-bid fallacy illustrates the topics of the first three sections: the danger in attempting to subvert competition, the benefits of "incentive compatible" design, and the relevance of auction theory.

1. One of several design flaws that produced this outcome was prohibiting the California ISO from substituting a cheaper better product for a more expensive poorer product (Wolak, 1999). Also see Brien (1999).

Chapter 2-1

Reliability and Investment Policy

When, by building theories upon theories, conclusions are derived which cease to be intelligible, it appears time to search into the foundations of the structure and to investigate how far the facts really warrant the conclusions.

Charles Proteus Steinmetz
The Education of Electrical Engineers
1902

RELIABILITY, PRICE SPIKES, AND INVESTMENT ARE DETERMINED BY REGULATORY POLICIES. Because these policies impinge on market structure rather than architecture, they have been overlooked too often as debates focused on "nodal pricing," "bilateral trading," or on market rules. The result has been a chaotic pricing policy and disaster in the Western U.S. markets. Part 2 assumes away two major problems, market power and transmission constraints, to focus exclusively on the structural core of a contemporary power market. The goal of Part 2 is to explain the major policy options and their implications. This requires an understanding of the causal links between policy controls and the key market outcomes—reliability, price spikes and investment. Both controls and outcomes are diagrammed in Figure 2-1.1.

Supply and demand characteristics comprise a market's core structure, but in a power market these are unusually complex. The supply side cannot store its output so real-time production characteristics are important, and two demand-side flaws interact detrimentally with this characteristic. Consequently, the market cannot operate satisfactorily on its own. It requires a regulatory demand for a combination of real-time energy, operating reserves, and installed capacity, and this demand must be backed by a regulatory pricing policy. Without this reliability policy, the power system would under-invest in generation because of the demand-side flaws. Reliability policy is the part of the structural core that can be affected immediately by design. The demand-side flaws can also be affected by policy, but these design changes take longer to implement.

Without the demand-side flaws and reliability policy, Figure 2-1.1 would represent a normal market; demand and supply conditions would feed into the market and determine prices. These would determine new investment which would

Chapter *2-2*

Price Spikes Recover Fixed Costs

Indebtedness to oxygen
The chemist may repay
But not the obligation
To electricity.

Emily Dickinson
(1830–86)
The Farthest Thunder That I Heard

WHAT PRICES WILL COVER A GENERATOR'S FIXED COSTS WITHOUT OVERCHARGING CONSUMERS? Short-run competitive prices perform this service and, in addition, induce the right level of investment in every type of generation technology. This does not solve the long-run problems of power markets because demand-side flaws prevent contemporary markets from determining competitive prices.[1] Subsequent chapters will discuss regulatory policies that compensate for these flaws, while this chapter focuses on how competitive power markets will work once the flaws are eliminated.

Although not completely accurate, competitive analysis is useful for dispelling two fallacies concerning fixed costs. The first asserts that prices equal to marginal cost (competitive prices) cannot cover fixed costs. The second asserts that although they can, they will do so only when the market is seriously short of capacity.

Chapter Summary 2-2: Short-run competitive prices would recover fixed costs for peakers and baseload plants alike. They would stimulate investment in the mix of technologies that produces the required power at least cost. Demand will occasionally push prices well above the *average* marginal cost of any supplier but not above the right-hand marginal cost. These price spikes can be summarized with a price-duration curve that facilitates computation of the competitive outcome.

Section 1: The Fixed-Cost Fallacy. This fallacy asserts that short-run competitive prices (marginal-cost prices) will prevent generators from recovering their fixed costs. These are covered *not* because short-run competitive forces set price equal to marginal cost, but because, if they were not covered, investors would stop building plants while demand continued to grow. This would cause shortages and

1. Power markets are termed "competitive" if their supply side is competitive even when demand-side flaws prevent the market from determining competitive prices.

Chapter *2-3*

Reliability and Generation

Edison's design was a brilliant adaptation of the simple electrical circuit: the electric company sends electricity through a wire to a customer, then immediately gets the electricity back through another wire, then (this is the brilliant part) sends it right back to the customer again.

Dave Barry

RELIABILITY IS AT THE HEART OF EVERY DEBATE ABOUT ENERGY PRICE SPIKES. If these are large enough, they induce the investment that provides the generating capacity necessary for a reliable system. In some markets capacity requirements also play an important role. A shortage of installed generating capacity is not the only cause of unreliable operation, but because it is the one most directly related to the operation of the wholesale markets, it is the only one considered in Part 2.[1]

Two aspects of reliability are always contrasted. *Security* is the system's ability to withstand sudden disturbances, while *adequacy* is the property of having enough capacity to remain secure almost all of the time. Part 2 focuses on adequacy and assumes that security requirements will be met if the system has adequate planning reserves. Requirements for operating reserves, which are intended to provide *security*, are of interest here mostly because of their role in raising price, stimulating investment, and thereby contributing to *adequacy*. This role is often overlooked because, under regulation, these requirements were unrelated to adequacy.

Chapter Summary 2-3: Operating reserves can be purchased directly by the system operator, but the market must be induced to provide adequate planning reserves. The first step in analyzing the market's effectiveness is to find what determines the optimal level of installed capacity. Under a simple but useful model of reliability, installed capacity is found to be optimal when the duration of load shedding is given by the fixed cost of a peaker divided by the value of lost load.

Section 1: Operating Reserves and Contingencies. A contingency is a possible or actual breakdown of some physical component of the power system. Typically, some operating generator becomes unavailable, leaving the system

1. Both distribution and transmission outages typically cause more loss of load.

Chapter *2-4*

Limiting the Price Spikes

We don't know what the hell it is, but it's very large and it has a purpose.

Dr. Heywood Floyd
2001, A Space Odyssey
1968

WHEN SUPPLY CANNOT EQUAL DEMAND, THE MARKET CANNOT DETERMINE A PRICE. When contemporary power markets have enough installed generation capacity (ICap) to prevent this market failure at all times, they have so much that generators cannot cover their costs. This contradiction occurs only for certain combinations of supply, demand curves and load-duration curves, but when it does occur, the market fails; there is no long-run equilibrium. Most current power markets may well satisfy the conditions for this failure.

The demand-side flaws are the cause of this market failure, but they need not be eliminated in order to remove it. Failure can be prevented simply by reducing the severity of the first demand-side flaw—by increasing demand elasticity. Until this is accomplished, the system operator must set the market price when the market fails to clear. At these times, some regulatory rule must be adopted for determining what price to set. If price is set only for these few hours, it must be set extremely high, so it may be better to set price more often and lower. Such details are considered in subsequent chapters. The purpose of this chapter is to explain why all current power markets need and have price limits.

FERC approved price limits of $750, $500 and $250/MWh for California between 1998 and August 2001. In the summer of 2000 it reduced the NY ISO's limit from its previously approved level of $10,000 to $1,000/MWh bringing it in line with PJM's limit. A year later it limited prices indirectly in the West to roughly $100/MWh. In between it suggested that what Western markets really needed was no price limit at all. The Australians tell us prices must be capped at the value of lost load, which they put at between $15,000 and $25,000/MWh AU. The new electricity trading arrangement in England allows much higher prices and promptly set a record of over $50,000/MWh. Settling on a reasonable policy will require understanding the nature of failure in power markets.

Chapter *2-5*

Value-of-Lost-Load Pricing

Fifteen years ago I used charred paper and card in the construction of an electric lamp on the incandescent principle. I used it too in the shape of a horse-shoe precisely as you say Mr. Edison is now using it.

Joseph Swan
in a letter to *Nature*, January 1, 1880

There you have it. No sooner does a fellow succeed in making a good thing than some other fellows pop and tell you they did it years ago.

Thomas Edison
in reply

SHEDDING LOAD IS AN EXPENSIVE WAY TO CURB DEMAND. It makes no distinction between those who need the power most and those who need it least. Because most customers' usage is not metered in real time, and because most do not know the price, contemporary markets have little ability to ration demand with price. Instead, when it is necessary, the system operator must ration demand by shedding load. In this case, the value of another megawatt of power equals the cost imposed by involuntary load curtailment. This value is called the *value of lost load*, VOLL.

Basic economic theory says it is efficient to pay suppliers the value of supplying another unit of output. Because VOLL is very high, perhaps above $10,000/MWh, this implies a very high price whenever load must be shed. Implementing this policy causes extreme price spikes, but these will be brief and lead to optimal investment in generating capacity and optimal reliability.[1] Although basic theory ignores risk and market power, it provides valuable insights and a basis for discussing more subtle theories of setting energy prices.

Chapter Summary 2-5: Because of the two demand-side market flaws, power markets are not yet able to use market forces to determine an appropriate reliability level. Some authority must estimate VOLL or some other determinant of optimal reliability. All such approaches are based directly or indirectly on VOLL, so the consequences of error in the estimated value of VOLL cannot yet be avoided. These consequences are not dramatic and can be reduced by overestimating VOLL. Price risk and market power are negative side effects of VOLL pricing.

1. For a discussion of generation adequacy and VOLL pricing in the U.S. context, see Hirst and Hadley (1999).

Chapter *2-6*

Operating-Reserve Pricing

By September [1879] a little building at Fourth and Market was completed and two tiny Brush arc-light dynamos were installed. Together they could supply 21 lights. Customers were lured by the unabashed offer of service from sundown to midnight (Sundays and holidays excluded) for $10 per lamp per week. Yet in light-hungry San Francisco, customers came clamoring. By the first of the next year, four more generators with capacity of more than 100 lights had been added. Electricity had come to the West.[1]

Pacific Gas and Electric
Our History

OPERATING RESERVES ARE NEEDED TO SOLVE VERY SHORT-RUN RELIABILITY PROBLEMS, BUT THEIR PRICING CONTROLS THE LONG RUN AS WELL. Engineering suggests appropriate levels for operating reserves, but it cannot, on its own, determine what price to pay for them. Surprisingly, their price should depend on the value of lost load (VOLL) and on long-run, more than short-run, reliability considerations. By setting prices to a relatively modest level when the system is short of operating reserves, rather than to the extremely high value of VOLL when the system is short of capacity, operating-reserve (OpRes) pricing can substitute for VOLL pricing. This opens up a wide range of policy options which can be used to solve some of the most pressing problems of today's power markets.

Chapter Summary 2-6: A market with random shifts in the annual load-duration curve is examined to compare the side effects of high and low price spikes. High price spikes are found to cause investment risk and to encourage the exercise of market power. Low spikes are just as effective as VOLL pricing at encouraging optimal investment in generation capacity. High price spikes are more useful on the demand side than on the supply side, so different price limits should be used for the two sides of the market.

Section 1: Less Risk, Less Market Power. In a market with two load-duration curves, one for "hot" years and one for normal years, short-run profits are found to fluctuate between zero and 400% of normal under VOLL pricing but only

1. This plant preceded Edison's "invention of the electric light" and opened three years before Edison's Pearl Street power station—the "first" central station. Charles Brush lit Broadway, New York from a central station in 1880. (See http://www.pge.com/009_about/past/&e=921.)

Chapter 2-7

Market Dynamics and the Profit Function

An electrick body can by friction emit an exhalation so subtile, and yet so potent, as by its emission to cause no sensible diminution of the weight of the electrick body, and to be expanded through a sphere, whose diameter is above two feet, and yet to be able to carry up lead, copper, or leaf-gold, at the distance of about a foot from the electrick body.

Sir Isaac Newton
Samuel Johnson's Dictionary of the English Language, 1755

ECONOMICS FOCUSES ON EQUILIBRIA BUT HAS LITTLE TO SAY ABOUT THE DYNAMICS OF A MARKET. Once economics shows that a system has a negative feedback loop so that there is a point of balance, it considers its job done. Engineers move beyond this stage of analysis to consider whether a system will sustain oscillations and, if not, whether it is over- or under-damped. Economics understands that investment dynamics can produce "cycles" but has faith that rationality will generally prevent this. It also ignores the noise sources (randomly fluctuating inputs) that keep economic systems excited.

Usually these oversights do not offer much cause for concern. In power markets, however, a 4 or 5% fluctuation in either load or capacity, coupled with the wrong pricing policy, can cause the average annual spot price to triple. Such dynamics cannot be ignored. If they are not corrected at the time of market design, they will be reported later by the press.

Many profit functions determine the same optimal equilibrium value of installed capacity, but this means only that they agree at one point. At other points, they may differ dramatically, and these differences imply different market dynamics. Although the profit function falls far short of providing a theory of those dynamics, it does provide some basic insights which make possible a discussion of the topic.

Chapter Summary 2-7: Profit functions can be calculated from the load duration curve and two policy variables: the price cap and the required level of operating reserves. Once calculated they reveal the equilibrium level of installed capacity and give some indication of the market's riskiness and conduciveness to the exercise of market power. Thus the first step in assessing a pricing policy should be the calculation of the associated profit function.

Chapter 2-8

Requirements for Installed Capacity

The shaft is 20 feet long and 6 ½ inches in diameter. The wheel, which is 56 feet in diameter [is] provided with 144 blades twisted like those of screw propellers. The sail surface is about 1,800 square feet. The speed of the dynamo at full load is 500 revolutions per minute, and its normal capacity at full load is 12,000 watts. The working circuit is arranged to automatically close at 75 volts and open at 70 volts. The amount of attention required to keep it in working condition is practically nothing. It has been in constant operation more than two years.[1]

Scientific American
December 20, 1890

PRICE SPIKES ENCOURAGE INVESTMENT INDIRECTLY; A CAPACITY REQUIREMENT GETS RIGHT TO THE POINT. The capacity approach is defined by two regulatory parameters but, as has been demonstrated, so is the price-spike approach. Both can induce any desired level of reliability while preserving the correct mix of technology, so the choice between them should be based on their side effects.

Chapter Summary 2-8: A capacity requirement produces an easily controlled, low-risk profit function. It can be combined with a price-spike approach to produce a profit function that is still relatively low risk while providing high prices at a few crucial times to tap existing high-priced resources. Combining price-spike and capacity-market profit functions does not increase equilibrium profits, but it does increase the equilibrium value of installed capacity unless the two policies are properly adjusted. Adjustment requires taking account of random fluctuations in the level of installed capacity.

Section 1: The Capacity-Requirement Approach. All load-serving entities are required to own, or to have under contract, a certain required capacity. The sum of these is the market's installed capacity (ICap) requirement and is typically about 18% greater than annual peak load. A load-serving entity is penalized if it fails to meet its requirement.

1. This wind generator was constructed by Charles Brush for his personal use. The article begins by noting that "After Mr. Brush successfully accomplished practical electric illumination by means of arc lights, incandescent lighting was quickly brought forward and rapidly perfected."

Chapter *2-9*

Inter-System Competition for Reliability

In the same open market, at any moment,
there cannot be two prices for the same kind of article.

W. Stanley Jevons
The Theory of Political Economy
1879

There are two fools in every market; one asks too little, one asks too much.

Russian proverb

MARKETS WITH LOW PRICE CAPS HAVE LITTLE PROTECTION FROM COMPETING MARKETS WITH HIGH PRICE CAPS. At crucial times, the high-cap market will buy up the reserves of the low-cap market. This can cause a group of competing markets to evolve toward a risky high-priced regulatory approach. Competing system operators are not "led by an invisible hand" to the optimal policy.

This chapter uses previously developed models and tools to investigate what happens when pairs of markets, operating under different pricing rules, compete for energy and capacity.[1]

Chapter Summary 2-9: Competition between markets with different price caps will favor the market with the higher cap. The lower-price-cap market will spend as much on inducing investment but will find its reserves bought out from under it at crucial times by the high-price-spike market. The result will be competition between system operators for the higher price cap unless a regional regulator prevents this. Capacity-requirement markets can solve a similar problem by requiring capacity rights to be sold on an annual basis.

Section 1: Price-Cap Competition. A VOLL market and OpRes market, identical except for their pricing policies, are modeled as able to trade energy and operating reserves. The result is that when load is expected to be high, more generators sell power and reserves to the VOLL market. This equalizes the expected short-run profits in the two markets; as a result the OpRes market suffers reduced reliability. This will force it to adopt a higher price cap to protect itself from inter-system competition.

1. Cournot anticipated Jevons, Menger, and Walras in discovering "the law of one price" and other neoclassical results. Jevons (1835-1882), with his publication of *The Coal Question* in 1865, was the first to draw attention to the imminent exhaustion of energy supplies. In 1870 he invented the "logical piano", the first machine to solve problems faster than a human.

Chapter *2-10*

Unsolved Problems

The day when we shall know exactly what "electricity" is, will chronicle an event probably greater, more important than any other recorded in the history of the human race. The time will come when the comfort, the very existence, perhaps, of man will depend upon that wonderful agent.

Nikola Tesla
c. 1893

LOW PRICES, EVEN THOSE ONLY 200% ABOVE NORMAL, CAN INDUCE INVESTMENT, BUT THEY CANNOT INDUCE CURTAILMENT OF HIGH-VALUE LOAD. Neither can they induce supply from the odd high-marginal-cost generator nor from a generator that is producing in its emergency operating range. The problems of low price caps have nothing to do with adequate investment in generating capacity. They concern only the short-run responses of high-value load and high-marginal-cost supply. These problems are small compared with the problem of long-term market stability and generation adequacy, but they are worth solving. This chapter defines them and suggests steps toward solutions.

Chapter Summary 2-10: Genuinely high-marginal-cost power should be purchased in a market separate from the regular supply market and with its own high price cap. Load reductions, most of which have high marginal values, should be purchased at a price that is allowed to exceed the price cap in the supply market. The (relatively) low price cap of the supply market should be set regionally and enforced strictly to prevent "out-of-market" purchases.

Section 1: High Marginal Costs and Low Price Caps. Several small sources of high-marginal-cost power make high prices worthwhile. To avoid disruptive side effects, these should be restricted to an emergency power market that caters only to high-cost sources. The problem is to select only these sources and limit market power under a high price cap. An adequate solution appears to be within reach.

Section 2: Pricing Supply and Demand Separately. Allowing demand-side prices to rise above the supply-side cap makes use of existing demand responsiveness and stimulates the development of greater elasticity. This requires a method of setting price higher than any supply bid and of refunding the extra revenue collected.

Chapter *3-1*

Introduction

The conclusion seventeen years later, is essentially the same . . . industries differ one from the other, and the optimal mix of institutional arrangements for any one of them cannot be decided on the basis of ideology alone. The "central institutional issue of public utility regulation" remains . . . finding the best possible mix of inevitably imperfect regulation and inevitably imperfect competition.

Alfred E. Kahn
The Economics of Regulation
1995

REAL-TIME TRANSACTIONS REQUIRE CENTRAL COORDINATION; WEEK-AHEAD TRADES DO NOT. Somewhere in between are dividing lines that describe the system operator's diminishing role in forward markets. Where to draw those lines is the central controversy of power-market design. A related controversy, not considered in Part 3, is how finely the system operator should define locational prices. Those who favor a large role for the system operator in one sphere tend to favor it in others. Thus the controversies of market architecture have a certain consistency. Although the rhetoric focuses on how centralized a design is, the litmus test in most of the controversies is the extent of the system operator's role. This too may be a distraction. A larger role for the system operator implies a smaller role for profitable enterprises. One side fears the inefficiency and market-power abuses of private parties playing social roles. The other side fears the inefficiency of nonprofit organizations but also covets the central market roles played by the system operator.

Power markets present unusually acute coordination problems. They are the only markets that can suffer a catastrophic instability that develops in less than a second and involves hundreds of private parties interacting through a shared facility. The extent and speed of the required coordination are unparalleled. Generators 2000 miles apart must be kept synchronized to within a hundredth of a second. Such considerations require a market that in some respects is tightly controlled in real time. Historically, this control has extended to areas far from the precarious real-time interactions. As deregulation brings markets into new areas, it is not surprising to find the proponents of markets reaching beyond their ability and to find the traditional system-control structure attempting to perpetuate now unnecessary roles for itself. This clash of interests has produced much heat and shed little light.

Chapter *3-2*

The Two-Settlement System

We can scarcely avoid the inference that light is the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.

James Clerk Maxwell
1861

This velocity is so nearly that of light, that it seems we have strong reasons to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws.[1]

1864

THE REAL-TIME PRICE ALWAYS DIFFERS FROM THE DAY-AHEAD PRICE. WHICH IS IN CONTROL? Day-ahead (DA) prices, and especially earlier prices, differ significantly from the corresponding real-time (RT) price. The differences are due to misestimations made before traders know all the details of the RT conditions. In a competitive market the RT prices are true marginal cost prices, and the forward prices are just estimates, sometimes very rough estimates. With most trade occurring in the forward markets, does this imply that only a small proportion of generation is subject to the correct incentives? Not under a proper two-settlement system. The purpose of the RT market is to correct the prediction errors of the past. If the transaction costs in this market are minimized so that profitable trade is maximized, the RT price will be accurate and will control actual production. Past mistakes have financial impacts but will not cause inefficiency which is a purely physical phenomenon.

Contracts for differences (CFDs) insulate bilateral trades from all risks of spot price fluctuations while allowing the inevitable inefficiencies of forward trading to be corrected by accurate RT price signals. Both the two-settlement system and CFDs allow efficient re-contracting—a standard economic solution to the problems of decentralized forward trading. Advocates of bilateral trading have often failed to recognize this point and have opposed the very mechanisms that make decentral-

1. Maxwell developed the mathematics of electromagnetic fields, later used to design AC motors, transformers, and power lines. He predicted the possibility of electromagnetic waves and calculated their theoretical velocity from laboratory measurements on electric and magnetic fields. At first his suggestion that light is electromagnetic was dismissed as a "not wholly tenable hypothesis."

Chapter *3-3*

Day-Ahead Market Designs

From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics.

Richard Feynman

One scientific epoch ended and another began with James Clerk Maxwell.[1]

Albert Einstein

CENTRAL DAY-AHEAD MARKETS CAN BE DESCRIBED AS AUCTIONS. The most obvious design sets energy prices based on simple energy-price bids. A different approach turns the system operator into a transportation-service provider who knows nothing about the price of energy but sells point-to-point transmission services to energy traders.

Either of these approaches presents generators with a difficult problem. Some generators must engage in a costly startup process (commitment) in order to produce at all. Consequently, when offering to sell power a day in advance, a generator needs to know if it will sell enough power at a price high enough to make commitment worthwhile. Some day-ahead (DA) auctions require complex bids which describe all of a generator's costs and constraints and solve this problem for the generators. If the system operator determines that a unit should commit, it ensures that all its costs will be covered provided the unit commits and produces according to the accepted bid. Such insurance payments are called "side payments," and their effect on long-run investment decisions is considered in Sections 3-9.3 and 3-9.3.

Chapter Summary 3-3: Day-ahead markets run by system operators are run as auctions. Although some trade energy, some sell transmission, and some solve the unit-commitment problem, they all use the same philosophy for choosing which bids to accept and for setting prices. Four archetypical markets are summarized: (1) a power exchange, (2) a transmission-rights market, (3) a power pool, and (4) PJM's DA market which mixes all three.

1. The universe is governed by four forces and matter is made of their associated particles. The electron and photon are the carriers of the electromagnetic force, and the other three forces are gravity, the weak force and the nuclear force. Maxwell discovered the mathematics of the electromagnetic force, Einstein the mathematics of the gravitational force, and Feynman the laws of the weak force which he unified with the electromagnetic force.

Chapter 3-4
Ancillary Services

Electric energy thus is the most useful form of energy—and at the same time it is the most useless. [It] is never used as such . . . [but] is the intermediary.

Charles Proteus Steinmetz
Harper's
January 1922

POWER IS THE PRIMARY SERVICE, BUT SIX ANCILLARY SERVICES ARE NEEDED TO ENSURE RELIABLE, HIGH-QUALITY POWER, EFFICIENTLY PRODUCED. Usually ancillary services are defined by how they are provided rather than by the service rendered. This results in a plethora of services and little insight into their relationship to market design. Defining services by the benefit they provide and defining them broadly produces a short but comprehensive list. Services directed at long-term investment are not counted as ancillary to real-time (RT) power delivery. The six listed services require planning by the system operator, but economic dispatch is jointly provided by the system operator and the market. How that task is shared should be the subject of intense debate.

Chapter Summary 3-4: Of the six ancillary services, the system operator or its agent must *directly* provide transmission security and trade enforcement, and to some extent economic dispatch. The other services, balancing, voltage stability, and black-start capability can be purchased from a competitive market, but the system operator must demand and pay for these services.

Section 1: The List of Ancillary Services. Services are defined by the benefit they provide to the market and its participants, not by their method of provision. An accurate frequency is required by some motors and particularly by large generators. Many appliances need a fairly accurate voltage, and together these two services define the provision of power from the customer's perspective. Transmission security and occasionally black-start capability are indirect services needed by the market to provide the first two. Economic dispatch can include the solution of the unit commitment problem and often includes efficiently dispatching around congestion constraints. Trade enforcement is required to provide property rights essential for bilateral trading.

Chapter *3-5*

The Day-Ahead Market in Theory

Economic questions involve thousands of complicated factors which contribute to a certain result. It takes a lot of brain power and a lot of scientific data to solve these questions.

Thomas Edison
1914

THE DAY-AHEAD MARKET IS THE FORWARD MARKET WITH THE GREATEST PHYSICAL IMPLICATIONS. By providing financial certainty, it can remove the risk of incurring startup expenses. The more efficient the market, the more accurate the startup decisions and the lower the cost of power. Even without the unit-commitment problem, reducing financial risk would reduce the cost of capital.

As explained in Chapter 3-3, the day-ahead (DA) market can utilize one of three basic architectures or a combination. Bilateral markets, exchanges and pools can each provide hedging and unit commitment. The controversy over the choice of architecture is driven by concerns over the shortcomings of private markets and nonprofit system operators in performing the coordination functions associated with unit commitment. Hedging is also an issue as pools claim to provide it more completely than exchanges.

Some theory of market clearing—when it is possible and when not—helps to provide a framework for evaluating the various designs. “Nonconvex” production costs are the key to this theory, and while conceptually arcane, the focus of current controversy and volumes of market rules attest to their impact on market design.

Chapter Summary 3-5: Nonconvex generation costs violate an assumption of perfect competition, but the magnitude of the resulting problems is unknown. The pool approach is designed to minimize these problems, but its pricing ignores investment incentives. The bilateral approach faces formidable coordination problems in the DA market. It may be less efficient and provide less reliability than a centralized approach. Side-payments made by a DA pool do not increase reliability.

Chapter *3-6*

The Real-Time Market in Theory

It must be done like lighting.

Ben Johnson
Every Man in his Humour
1598

UNLIKE A DAY-AHEAD EXCHANGE, A REAL-TIME EXCHANGE CANNOT USE BIDS. The real-time (RT) market consists of trades that are not under contract—power that just shows up, or is taken, in real time and accepts the spot price. An RT exchange works like a classical Walrasian auction. A price is announced and suppliers and customers respond. If the market does not clear, a new price is announced. The difference is that in a power market trade takes place all the time; there is no waiting to trade until the right price is discovered. Like a Walrasian auction and unlike a day-ahead (DA) exchange, an RT exchange may find there is no price that balances supply and demand.[1] Consequently, if an exchange is used, it must be supplemented with another exchange or perhaps an operating-reserve market in the form of a pool. There are many possibilities, and little is known about their relative merits.[2]

Chapter Summary 3-6: Pure bilateral markets are too slow to handle RT balancing and transmission security. A centralized market is needed which can take the pool approach, the exchange approach, or something in between. Pools have an easier time achieving a supply-demand balance than do exchanges because they utilize different prices for different generators. This allows them to offer an option in a forward market that depends on the real-time market price. An RT exchange can achieve a similar effect but only by employing one or more additional exchanges. These could be for detrimental generation or operating reserves. Alternatively an operating-reserve pool could supplement an RT exchange.

1. Although the lack of a market clearing price is indicated by a static analysis, the dynamics of a power market which are limited by ramping constraints may reverse this conclusion.

2. See Rothwell and Gomez (2002) for descriptions of Norwegian, Spanish, and Argentinian spot markets.

Chapter 3-7

The Day-Ahead Market in Practice

Predicting is pretty risky business, especially about the future.

Mark Twain
(1835–1910)

IS THE DAY-AHEAD POOL PRICE DETERMINED BY ARBITRAGE OR COMPUTATION? Forward prices are usually determined by arbitrage between the forward market and the real-time (RT) market. Day-ahead (DA) markets are forward markets, but DA pools with multipart bids have been promoted for their ability to determine through computation the optimal dispatch and the efficient price. Both theories might prove true, or half true, but more likely one is essentially right and the other wrong.

There is no question that the computation takes place accurately and in a mechanical sense determines the DA pool price. The result of the computation is also determined by its inputs. Because the computation itself is a fixed procedure, while the input changes daily, it may be best to view the pool price as determined by inputs. Then the question becomes: Do the pool's input data accurately reflect the producer's reality, or do they deliberately misrepresent that data in order to take advantage of arbitrage opportunities. In the first case, the pool calculation makes use of good input data to produce a price that reflects the true details of generation costs. In the second case, the calculation is not a sensible unit-commitment calculation because its inputs are false. The bidders have manipulated the pool's computation, and the outcome may be thought of as being determined by arbitrage.

Unfortunately, it takes only a small percentage of arbitragers to dominate the outcome. In PJM's DA market, 11% of the generators that are needed, and eventually produce power, in the RT market are rejected by the unit commitment calculation (PJM, 2001, 30).[1] This alone proves the calculation is not highly accurate. Morever, while 3,260 MW of multipart bids representing needed generation are rejected, 6,169 MW of supply-side arbitrage bids (one-part bids) are accepted on

1. Only 26,771 MW submit DA multipart bids, while 30,031 MW proves to be necessary in real time.

Chapter *3-8*

The Real-Time Market in Practice

Where there is much light, the shadows are deepest.

Goethe
Wilhelm Meister's Apprenticeship
1771

WHEN MARGINAL-COST PRICES WON'T CLEAR THE MARKET, ARE THEY STILL THE RIGHT PRICES? When production costs are "nonconvex," as are startup and no-load costs, competitive market theory predicts the market may not clear, marginal-cost prices may not be optimal, and the market may not be efficient. The sole purpose of the multipart bids used by power pools is to overcome problems caused by nonconvex costs. The pool approach recommends setting the market price equal to marginal cost exactly as if there were no problem and then making side payments to generators who are needed for the optimal dispatch. These payments cover only the costs that marginal-cost prices fail to cover.

The pool approach recognizes the first failure of marginal-cost pricing and corrects it with side payments. But true competitive prices do more than minimize production costs; they send the right signals (1) to consumers and (2) to investors in new generation. Can a power pool's combination of marginal-cost prices and side payments replicate these benefits of competitive prices? Using an example from a dispute over NYISO's pricing, this chapter shows that pool pricing fails both of these tests. Surprisingly, it was NYISO's position that standard marginal-cost pricing—pool pricing—was inefficient for at least five reasons.

Pool prices are neither the prices of Adam Smith, nor those of competitive economics. They are not right for the demand side and they are not right for long-run investment. Pool prices are right for the centralized solution of the problem of minimizing short-run production-costs, given an output level that is incorrectly determined when demand is elastic.

This does not mean the pool approach is a bad idea; it simply means that adopting it because it gets the prices right would be naive. The problem of nonconvex costs is difficult and a complex market design, such as a pool, could

Chapter 3-9

The New Unit-Commitment Problem

Let us not go over old ground, let us rather prepare for what is to come.

Marcus Tullius Cicero
(106–43 B.C.)

THE OLD PROBLEM ASKS WHICH UNITS SHOULD BE COMMITTED; THE NEW PROBLEM ASKS WHAT MARKET DESIGN WILL BEST SOLVE THE OLD PROBLEM. The old problem was solved by collecting data on all the generators and applying the techniques of mathematical programming. The new problem might be solved by a market designed to induce generators to voluntarily and accurately provide this same data. The market coordinator could then purchase power from the generators identified by the old algorithm. This is the power-pool approach. A power exchange is an alternative approach which pretends the old problem does not exist.

It seems impossible that ignoring the old problem could be the best way to solve it, but most market architectures ignore just such complex commitment problems. When the market coordinator ignores the problem, the suppliers take it up, and they may do a remarkably good job. Although the programming techniques used to solve the old problem are astoundingly complex, most generators can get the right answer on most days simply by looking at the calendar. If you have a baseload plant and it's summer, keep your plant committed. If you have a peaker, don't look at the calendar, just watch the real (RT) price—day-ahead (DA) forecasts are not needed. When it really matters, on the hottest days, every supplier knows to commit. But those who commit units for individual suppliers will do much better. They will have years of experience and the necessary resources. Moreover, they may have access to an exchange that uses two-part prices or multiple rounds of bidding. To beat a good exchange market, a pool must be very good indeed.

Two central concerns have motivated the power-pool approach: efficiency and reliability. Committing the wrong units costs more, but this problem is limited by the magnitude of the cost involved and the efficiency of markets without centralized unit commitment. Reliability is more of a wild card. Perhaps a decentralized market

Chapter *3-10*

The Market for Operating Reserves

A Telegrapher's Valentine, by James Clerk Maxwell, 1860.

The tendrils of my soul are twined
With thine, though many a mile apart.
And thine in close coiled circuits wind
Around the needle of my heart.
Constant as Daniel, strong as Grove.
Ebullient throughout its depths like Smee,
My heart puts forth its tide of love,
And all its circuits close in thee.

O tell me, when along the line
From my full heart the message flows,
What currents are induced in thine?
One click from thee will end my woes.
Through many a volt the weber flew,
And clicked this answer back to me;
I am thy farad staunch and true,
Charged to a volt with love for thee.

A MARKET FOR OPERATING RESERVES PAYS GENERATORS TO BEHAVE DIFFERENTLY FROM HOW THE ENERGY MARKET SAYS THEY SHOULD. If generators are cheap and will produce at full output, the market might tell them to produce less. If they are too expensive to produce at all, it may tell them to start "spinning," and this may require them to produce at a substantial level. Its purpose is to increase reliability and moderate price spikes.

Chapter Summary 3-10: Not maximizing profits has an opportunity cost, and generators must be paid for this to secure their cooperation. There are two philosophies: (1) have the system operator calculate this value from the real-time price and pay them accordingly; (2) have the generators guess this value and include it in their bids. The first approach may be quite susceptible to gaming while the second is optimal in theory but risky for generators in practice and may increase the randomness of the outcome.

Section 1: Types of Operating Reserve. Operating reserves come in several qualities classified by how quickly the generator can respond. "Regulation" keeps the system in balance minute by minute. Ten-minute spinning reserve can start responding almost instantly and deliver its full response within ten minutes. This type of reserve will serve as a model for considering market designs. The problem of linking the different reserve markets is not considered.

Section 2: Scoring by Expected Cost. One approach to conducting a market for spin is to have suppliers submit two-part bids, a capacity price, CC_{bid}, and an energy price, VC_{bid}. An obvious way to evaluate such bids is to score them by their

Chapter *4-1*

Defining Market Power

Sixty minutes of thinking of any kind is bound to lead to confusion and unhappiness.

James Thurber
(1894-1961)

MARKET POWER, A CENTRAL TOPIC IN ECONOMICS, HAS BEEN DEFINED CAREFULLY. The standard economic definition is a central concept of industrial organization. Market power is the ability to alter profitably prices away from competitive levels. This definition, with slight variations, has probably been in use for more than a hundred years and is supported by a large body of empirical and theoretical work. It is terse and carefully worded as a good technical definition should be. Frequently, regulators ignore it and attempt their own definition. FERC announced a new one in its report, State of the Markets 2000. By the end of 2001 it had been discarded and a new one was under design.

> *Market power is defined as the ability to withhold capacity or services, to foreclose input markets, or to raise rival firms' costs in order to increase prices to consumers on a sustained basis without related increases in cost or value.* (FERC 2000a)

Market power is a three step process: (1) an exercise, (2) an effect on price and quantity, and (3) an impact on market participants. The first step can take on many forms and appearances. For clarity, the economic definition considers only price (step 2) and profit (step 3). FERC's definition focuses on step 1 and misses some methods of exercise such as raising the offer price of a supply bid. Because it omits "profitably," all baseload plants would have market power whenever they are needed even though they would lose money if they exercised it. It ignores the concept of *competitive price*. It adds the clause "without related increases in cost"—any time a supplier exercises market power it is "trying to recover" some "related increase in cost." The notion of "sustained basis" may exclude peakers and is vague. No authority is cited for this definition.

Both the FERC definition just cited and the revisions to it under consideration in January 2002 include the concept of raising competitors' costs. In a nontechnical,

Chapter 4-2
Exercising Market Power

People of the same trade seldom meet together, even for a merriment or diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices.

Adam Smith
The Wealth of Nations
1776

MONOPOLY POWER ALWAYS RESULTS IN COMPETITIVE SUPPLY BEING GREATER THAN DEMAND AT THE MARKET PRICE—A POSITIVE QUANTITY WITHHELD. This signature of market power must be checked to determine the nature of any observed price increase and to determine its significance. A market that cannot tolerate a few hundred megawatts of withholding cannot withstand a hot afternoon or a generator outage. Such a market should be shut down and redesigned.

Chapter Summary 4-2: Market power should be looked for only in the real-time (RT) markets. It should be looked for among inframarginal as well as marginal generators. The amount of withholding should always be examined along with the price increase.

Section 1: Market Power and Forward Markets. If prices in a forward market are too high, customers can wait for the next forward market or for real time. In the RT market they can wait no longer. Consequently, market power cannot be exercised in forward markets; this includes day-ahead (DA) markets. Market power exercised in real time is reflected into forward-market prices through arbitrage.

Section 2: Long-Run Reactions to Market Power. The threat of new supply entering can discipline the exercise of market power because those exercising it fear high prices will attract too many future competitors. Suppliers may also choose to exercise less market power for fear of retaliation by regulators.

Section 3: Marginal and Inframarginal Market Power. Frequently the generator that sets the market price is not the one exercising market power. Often it is some inframarginal generator, possibly one that would be marginal had it not priced itself out of the market.

Chapter *4-3*
Modeling Market Power

Merchants are occupied solely with crushing each other:
such is the effect of free competition.

Charles Fourier
(1772-1837)

Like many businessmen of genius he learned that free competition was wasteful, monopoly efficient.

Mario Puzo
The Godfather
1969

MODELING MARKET POWER HELPS EXPLAIN THE FACTORS THAT CONTROL THE EXERCISE OF MARKET POWER. The models, however, are not accurate predictors of market power. Except in the case of pure monopoly, market power is not well understood because it involves strategic behavior by several competitors—an **oligopoly**. Game theory is subtle, and the game of oligopolistic competition has complex rules. By abstracting from much of its complexity, models of market power explain its main features.

Elasticity of demand is the most important factor in present power markets. The distribution of the size of competitors is also a key factor. Even this is extremely hard to compute in a power market because it depends on barriers to trade which are very complex and vary continuously, often dramatically. The style of competition is also crucial. Economics can model competition based on price or quantity, but in power markets suppliers compete using "supply curves" that combine the two. The theory of supply-curve competition (discussed in Chapter 4-4), while promising, is not yet well developed.

Chapter Summary 4-3: The market power of a monopolist is limited by demand elasticity. An oligopoly's market power is also limited by the number and relative size of competitors. Market share is a supplier's sales divided by total trade in the market. The sum of the squares of market shares is called the Herfindahl–Hirschman Index (HHI) and is one of three factors determining market power in the Cournot model.

Section 1: Monopoly and the Lerner Index. A monopolist will raise price to maximize profit but could raise it even higher. The more customers respond to high prices by curtailing demand, the less a monopolist will raise the price. The

Chapter 4-4

Designing to Reduce Market Power

Monopoly, in all its forms, is the taxation of the industrious for the support of indolence, if not of plunder.

John Stuart Mill
Principles of Political Economy
1848

MARKET POWER CAN BE CONTROLLED BY COMPETITION OR BY MONITORING AND ENFORCEMENT. Competition is preferable but does not automatically reach satisfactory levels. The four key determinants of the competitiveness of a power market are: (1) demand elasticity, (2) supplier concentration, (3) the extent of long-term contracting, and (4) the extent of supply-curve bidding.

All four of these are partially susceptible to regulatory policy and/or market design. None are particularly easy to affect, with the exception of the system operator's demand elasticity for operating reserves which translates directly into elasticity of demand for power. Since lack of demand elasticity is the primary source of market power in the industry, this simple change is long overdue. Fortunately, demand elasticity of large customers is also relatively easy to increase, once policy makers understand its importance and how to minimize the risks of real-time (RT) pricing.

Supplier concentration is quite difficult to decrease, but merger policy should prevent its increase at least until the demand elasticity problem has been fixed. There is little theory on how to increase long-term contracting except at the time of divestiture when vesting contracts can be required. Supply-curve bidding is naturally encouraged by uncertainty in the demand level. A bid that must span the entire day effectively increases this uncertainty. Supply-curve bids tend to be very elastic at lower output levels and very inelastic at high output levels, which explains why power markets have relatively little trouble with market power off peak. Research is needed to find a policy to harness this effect.

Chapter Summary 4-4: An ounce of prevention is worth a pound of cure and strengthening the normal forms of market-power mitigation may be the best form of prevention. Demand elasticity tops the list. Forward contracting should be

Chapter *4-5*

Predicting Market Power

A friend in the market is better than money in the chest.

Proverb
collected in Thomas Fuller, *Gnomologia*
1732

STANDARD "WISDOM" HOLDS THAT HHIS BELOW 1000 ARE CERTAINLY SAFE—THEY ARE NOT. The HHI accounts for only one factor, concentration, out of five key economic factors that determine the extent of market power. The other four, demand elasticity, style of competition, forward contracting, and geographical extent of the market, can each affect market power by an order of magnitude.

Chapter Summary 4-5: The HHI misses most of the action in power markets. Cournot models can capture much more but still miss the mark widely. Results are often reported in terms of the Lerner index, which frequently reports a decline in marginal cost as if it were an increase in price. A combination of estimating absolute market power and predicting relative market power may answer some questions a little more accurately. These include the impact of mergers, transmission upgrades and transmission pricing.

Section 1: Four Factors which HHI Ignores. The Herfindahl-Hirschman index is computed from the market shares of suppliers. It takes no account of demand inelasticity which, other things being equal, makes market power at least 10 times worse in power markets than in most other markets. It predicts only the Lerner index, which is loosely connected to market power. It takes no account of the style of competition, the extent of forward contracting, or the geographical extent of the market.

Section 2: Difficulties Interpreting the Lerner Index. While the definition of market power compares market price with the competitive price, the Lerner index compares market price with marginal cost in the uncompetitive market. The Lerner index combines the effects of reduced marginal cost and increased price. In a power

Chapter 4-6

Monitoring Market Power

Why is there only one monopolies commission?

Anonymous

DISCOVERING MARKET POWER OFTEN REVEALS MARKET FLAWS. While fixing these flaws, temporary restraints may need to be placed on market participants, but the goal of fixing the market should be kept in focus. If the rules are flawed, repairs can be made quickly, but when the architecture or structure is flawed, the required changes can take years. Then market monitors are forced to spend too much effort controlling an unruly market.

Chapter Summary 4-6: The trick to market monitoring is to ignore vague definitions and rigorously apply the economic definition of market power. If the market price is above the competitive level, then necessarily, some supplier is not acting as a price taker—somewhere there is a gap between the profit-maximizing supply of a price taker and the actual supply. That gap is termed the "quantity withheld," and it is proof of market power. Observing a high price is not proof, nor is observing a price higher than left-hand marginal cost. Observing the real-time (RT) price above right-hand marginal cost demonstrates short-run withholding. In the RT market, this proves market power, except for unusual cases of opportunity cost such as exhibited by hydrogenerators.

Section 1: FERC's Ambiguous Standard. A popular misinterpretation of the DOJ/FTC definition of market power, abetted by DOJ's Guidelines, requires the price increase to be significant in magnitude as well as duration. FERC's requirement that "no market power" be exercised makes sense only under this misinterpretation; otherwise, the requirement would be an impossibility. As a consequence, FERC's policies regarding market power are highly ambiguous, and its pronouncements on the need to eliminate all market power cause confusion.

Section 2: Market Monitoring. Power markets need monitoring because of two structural problems: (1) electric energy cannot be stored, and (2) RT demand

Chapter *5-1*

Power Transmission and Losses

ELECTRICITY. n.s. [from electrick. See ELECTRE.] A property in some bodies, whereby, when rubbed so as to grow warm, they draw little bits of paper, or such like substances, to them.

Samuel Johnson
The Dictionary of the English Language, 1755

ELECTRICITY, n. The power that causes all natural phenomena not known to be caused by something else.

Ambrose Bierce
The Devil's Dictionary, 1881–1906

WATTS MEASURE POWER, BUT VOLTS AND AMPS ARE THE NUTS AND BOLTS OF ELECTRICITY. The economics of power flows can be understood without their help, yet they underlie every important physical phenomenon in the power marketplace. This chapter uses them to explain power flow, transmission losses, and the reason Westinghouse's AC networks triumphed over Edison's DC networks.

Chapter Summary 5-1: Voltage is pressure and electrical current is like a flow of water; with more pressure, more current flows. Power delivered is voltage times current (volts times amps) and is measured in watts. The power lost in a transmission line of a given voltage is proportional to the square of the power flow, but if the voltage is doubled the same power can be delivered with $\frac{1}{2}$ the current and $\frac{1}{4}$ the loss. Transformers make it easy to raise an AC voltage and this allows the transmission of power with very little loss.

Section 1: DC Power lines. Direct-current power lines, used in the past and again gaining importance, provide the simplest example of power transmission. Power transfer, W, equals $V \times I$, which is the voltage, V, of the transmission line times the current flow, I. Transmission losses are proportional to the square of the delivered power and inversely proportional to the square of the power-line voltage. Consequently, two equal loads cause four times the loss caused by one. This makes a meaningful assignment of losses to loads impossible.

Section 2: AC Power lines. In the United States, alternating current (AC) completes a *cycle*, two reversals of direction, 60 times per second. Transformers,

Chapter 5-2

Physical Transmission Limits

We're set up for direct current in America.
People like it, and it's all I'll fool with.

Thomas Edison
1884

WITHOUT TRANSMISSION LIMITS, POWER MARKETS WOULD HAVE AMPLE COMPETITION AND NO NEED FOR CONGESTION PRICING. Designing power markets would be far easier. A competitive market will only account for the physical limits on all power lines and transformers if property rights are designed to represent these limitations properly.

Transmission limits are of two types: (1) physical limits, and (2) contingency limits. Physical limits are the basis of the contingency limits, but contingency limits are stricter and are the relevant limits for trading. A contingency limit ensures that a line's physical limit will not be violated if some other line or generator goes out of service unexpectedly. All limits, whether physical or based on contingencies, can be expressed at any point in time as a simple megawatt limit on power flow that is allowed over the power line or transformer in question. But contingency limits present a complex problem for trade because a trade can affect the contingency limit on lines it uses. (This problem is not discussed in Part 5.)

Two facts about limits are economically important, the megawatt limit itself, and how predictable it is. Unpredictability makes forward trading difficult, and this chapter explains some of the reasons that limits vary.

Chapter Summary 5-2: Power lines have physical limits that restrict the amount of power they are allowed to carry. These limits prevent overheating of the wires, instability of the power flow, and low-voltage conditions at the load end of the lines. Any of these limits can be thought of as a simple limit on real power flow, but because the basis of these limits is complex, they may vary according to how the system is operating at a particular time.

Chapter *5-3*

Congestion Pricing Fundamentals

The popular belief is that radium constantly produces heat and light without any appreciable loss in its weight. . . . there exists a form of energy of which we have as yet no knowledge, but which may yet become available to us as a result of further discoveries.

George Westinghouse
1911

PHYSICAL IMPEDIMENTS TO TRADE CAUSE COMPETITIVE PRICES TO DIFFER; THE DIFFERENCE IS THE PRICE OF CONGESTION. Power lines, because of their limited capacity, often cause energy prices to differ between locations. Congestion prices were not invented for electricity grids, need not be centrally calculated, and occur on their own in competitive markets.

Chapter Summary 5-3: If there are binding physical transmission limits between different locations, a competitive bilateral market with physical transmission rights will trade power at different prices in different locations. These competitive locational prices of power (CLPs) are unique, are the same as nodal prices (LMPs or LBMPs), and are the only efficient prices.

Section 1: Congestion Pricing is Competitive Pricing. If the transmission line between two locations is inadequate to handle the desired trade between those two locations, the downstream location will be forced to buy power from more expensive local generators, This will raise the local price of power relative to the remote price, which is a standard competitive result and has nothing to do with centralized computation.

Section 2: Benefits of Competitive Locational Pricing. Like all competitive prices, CLPs minimize the cost of production and reveal to consumers the true cost of their consumption. CLPs at each location equal the system's marginal cost of providing power at that location.

Chapter *5-4*

Congestion Pricing Methods

Investigations . . . indicated such great simplification in wireless telephone apparatus that we may, within the quite near future, have placed at our disposal a simple portable apparatus which will permit wireless conversation to be carried on over a considerable area. This will prove of great value in sparsely settled districts.

George Westinghouse
1911

THE POINT OF CENTRAL CALCULATION IS TO FIND THE PERFECTLY COMPETITIVE, BILATERAL-MARKET PRICES. If competition is strong in a centralized market, bids will be honest and the data used in the central computation accurate. Though bilateralists often object to "nodal" or "marginal-cost" locational prices, bilateral theorists know that theoretical nodal prices are exactly the prices a bilateral market would produce if it worked perfectly. The real debate is not over the prices but over which system will do a better job of finding them.

Chapter Summary 5-4: Central computation finds the optimal dispatch and then computes prices from the marginal benefits of a free megawatt at each location. Transmission constraints make power more valuable in some locations than others. Bilateral traders never consider the optimal dispatch but look only for profitable trades. Arbitrage produces a single price at each location, but transmission constraints can prevent it from leveling prices between locations. These two different processes lead to the same quantities being traded and to the same prices because perfectly competitive bilateral trade is efficient.

Section 1: Centralized Computation of CLPs. Central computation of CLPs makes the ISO the trading partner of every buyer and seller and does not price congestion separately. Power flows in a looped network are governed by the impedances (resistances) of the lines. This is illustrated with a three-line looped network with one constrained line. If the constraint is binding, all three buses will have different prices, and these will cause generators to minimize the total cost of production.

Section 2: Comparing Bilateral and Central Congestion Pricing. Competitive bilateral trading produces the same locational prices in the example network as central computation. The bilateral system collects congestion rent by selling

Chapter 5-5

Congestion Pricing Fallacies

Furthermore, the use of electricity will conserve the coal deposits of the world.

George Westinghouse
1911

NOT REALIZING THAT CONGESTION PRICING *IS* COMPETITIVE PRICING, ADVOCATES OF COMPETITION CRITICIZE IT AS UNFAIR. Competitive locational prices, like all competitive prices, contain scarcity rents that cover fixed costs of generators as well as congestion rents that cover the fixed costs of the grid. None of this revenue is wasted, and occasional high prices caused by congestion send the right signals to investors to build new generators, to customers to use less power, and to the ISO to build needed lines.

Section 1: Are Competitive Locational Prices Too High? Congestion is managed by redispatching expensive local generation in place of cheaper remote generation. Usually the additional production cost is much less than the cost to consumers, and this is often cited as a failure of competitive locational pricing. The high prices are necessary, however, to cover the fixed costs of an efficient set of generators and to cover fixed transmission costs.

Section 2: Congestion Taxing. Critics of congestion pricing usually suggest congestion taxing, instead of congestion pricing, though not by that name. This three-step plan first pretends congestion does not exist and finds a single price to clear the entire market. Then the ISO redispatches around the congestion, at a cost, and finally taxes consumers to pay the redispatch cost. If generators would bid honestly, the tax would be small, but the system induces dishonest bidding that results in a higher average price for power than does standard congestion pricing.

5-5.1 ARE COMPETITIVE LOCATIONAL PRICES TOO HIGH?

Fallacies of congestion pricing tend to focus on the occasional high prices caused by local scarcity when import lines are congested. Perhaps the most articulate exposition of these misconceptions is found in Rosenberg (2000).

Chapter 5-6
Refunds and Taxes

When asked by Gladstone if electricity would have any practical value, Faraday replied, "Why sir, there is every possibility that you will soon be able to tax it."

Michael Faraday

HOW SHOULD THE ISO REFUND CONGESTION RENT AND COLLECT REVENUE SHORTFALLS? The answer is surprisingly simple. Over- and under-collections of revenue should be added together and collected as a flat energy charge or tax. Whether to collect it from generators or loads should be decided as matter of convenience as the tax incidence of both choices is the same.

This answer depends on defining "prices" as charges intended to allocate resources efficiently and "taxes" as charges intended to raise revenue. The answer is most convincing if prices have been designed as efficiently as possible. Some will argue that because prices are not yet right, taxes should be designed to compensate for their shortcomings. But if taxes begin to take on the role of prices, improved prices will then be disputed on the basis that these would interfere with the incentives of the tax structure. The best designs will result from separating the two tasks: taxes should only raise revenue to pay for fixed-costs of the system, while prices should be designed with only efficiency in mind.

Chapter Summary 5-6: Pricing should be used to the greatest practical extent for two reasons: (1) to maximize efficiency, and (2) to minimize taxes. Taxes should be designed to minimize deadweight loss. This is accomplished easily with a flat charge on energy.

Section 1: Pricing versus Taxing. System prices are defined as charges intended to increase efficiency, while system taxes are defined as charges intended to raise revenue. As many services and externalities as possible should be priced efficiently, both to raise revenue and to increase efficiency. The inevitable revenue shortfall must be recovered through taxes which should be designed to reduce efficiency as little as possible. A flat energy tax does this so well there is little chance for improvement.

Chapter *5-7*

Pricing Losses on Lines

Reason and Justice tell me that there is more love of man in electricity and steam, than in chastity and refusal to eat meat.

Anton Chekhov
letter, concerning Tolstoy
1894

LOSSES ARE MORE THAN THE SUM OF THEIR PARTS. If one power flow loses 10 MW to heating power lines and another loses 20 MW, the two flows together will lose 58.3 MW. Billing for the extra 28.3 MW of losses has caused much controversy. To make matters worse, economists suggest charging the first power flow the cost of replacing 38.9 MW and the second the cost of replacing 77.7 MW. Together this is double the actual cost of losses.

Though charging transmission users for marginal losses often seems unfair to those being charged, it is what a competitive market in transmission would do. As always, the competitive approach increases market efficiency relative to the regulatory average-cost approach. The fact that it causes an "overcollection" is also an advantage as the revenue can be used to reduce the inefficient taxes that are needed to cover various fixed costs (see Chapter 5-6).

Chapter Summary 5-7: If transmission were provided by many small competing line owners, each with an unconstrained line, the price of transmission would equal the marginal cost of losses. This price would minimize the total cost of power.

Section 1: The Competitive Price Is Twice the Average Cost. Marginal-cost pricing means charging a power flow for the losses caused by a 1-MW increase of the flow times the amount of the flow. It collects about twice the cost of replacing total losses.

Section 2: Competitive Loss Pricing. A decentralized (bilateral) competitive transmission market would price losses at marginal cost thereby collecting about double the cost of the losses.

Section 3: Inefficiency of Average-Loss Pricing. Average-cost pricing of losses causes an inefficient dispatch and thereby increases the total cost of power. When generation is not otherwise constrained, the cost increase is approximately equal to the total cost of losses under an efficient dispatch.

Chapter 5-8

Pricing Losses at Nodes

Her own mother lived the latter years of her life in the horrible suspicion that electricity was dripping invisibly all over the house.

James Thurber
My Life and Hard Times
1933

LOSS PRICING, LIKE CONGESTION PRICING, CAN BE SIMPLIFIED BY USING NODAL PRICES. By charging each generator and load the loss price at its bus, all trades will be charged properly for their losses. Also, if politics prevents locational pricing for loads, generators will still receive correct price signals.

It is easiest to charge for losses by including loss prices in the competitive locational prices (CLPs), but if they are charged separately and billed after the fact, generators will learn to adjust their bids and the market will still handle losses quite efficiently. Even the restrictions that total loss charges not exceed the value of lost power and that losses be paid for in kind can be handled quite efficiently by shifting nodal loss prices.

Chapter Summary 5-8: By choosing a reference bus, marginal loss prices can be computed at every bus. These are relative prices, but that is all that is needed. If losses are not collected from loads, then the total collection of losses can be controlled by choosing the reference bus. One nearer to load increases collections, and one nearer to generation decreases it. Shifting the reference bus shifts all loss prices uniformly, which has no economic impact for reasons discussed in Chapter 5-6. In particular, loss charges can be adjusted so that generators pay marginal loss prices, but total collections equal the total cost of loss replacement.

Section 1: Nodal Loss Prices. Unlike bilateral loss prices, which have an absolute value, nodal prices are relative. System physics cannot attribute losses to a power injection but only to a point-to-point power transfer. Similarly, only the differences between nodal loss prices are meaningful.

Choosing a reference or "swing" bus allows both losses and loss prices to be computed on a nodal basis, but both give absolute answers only when used in pairs corresponding to an injection and a withdrawal.

Chapter 5-9
Transmission Rights

Today's scientific question is:
What in the world is electricity and where does it go after it leaves the toaster?

Dave Barry

FINANCIAL RIGHTS REFLECT ELECTRICAL REALITY; PHYSICAL RIGHTS REFLECT AN ILLUSION—THE NOTION THAT SUPPLIERS ACTUALLY DELIVER *THEIR* PRODUCT TO *THEIR* CUSTOMERS.[1] If supplier A sends power to load B and supplier B sends power to load A, their shipments may physically cancel each other on the connecting power line with the result that no power flows from A to B or from B to A. Instead supplier A's power goes to supplier B's customer and vice versa.

Suppose that instead of selling in their own regions, the Northern California generators decide to sell to Southern California and vice versa. Nothing physical changes. The same generators produce, the same loads consume, and the same amounts of power flow over the same paths. But with the new contracts, traders wish to own 10 GW of north-south rights and 10 GW of south-north rights. If the rights are financial, they just cancel out for the issuer. Whatever they pay to one set of rights they collect from the other. This calculation so perfectly mirrors physics that when financial rights are summed to find out if the total set is feasible, they are first converted to power flows and then summed by the engineers.

Issuing 10 GW of physical rights in each direction is next to impossible. The physical path may be limited to 2 GW. What if 4 GW of south-north rights were not exercised? The path would be burned out if the 10 GW of north-south rights were exercised, so 2 GW of north-south flow would be cancelled. Such rights are not very firm. To ensure that physical rights are firm, the issuance of such rights is limited to 2 GW in each direction on a 2 GW path. This forces trade to fit the limited concept of goods moving from supplier to customer without the possibility of automatic rerouting according to the far more efficient laws of power flow. Financial rights automatically cancel and reroute just as do power flows.

1. This is not the contract-path fiction. It accounts for paths but not for counterflow cancellations.