Relative Efficiency Benefits of Wholesale and Retail Competition in Electricity: An Analysis and a Research Agenda

Douglas R. Bohi and Karen L. Palmer

Resources for the Future
Washington, D.C.

NREL Technical Monitor: Paul Galen

National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, Colorado 80401-3393
A national laboratory of the
U.S. Department of Energy
Managed by Midwest Research Institute
for the U.S. Department of Energy
under Contract No. DE-AC36-83CH10093

Prepared under Subcontract No. AAH-5-15202-01

March 1996
## CONTENTS

### INTRODUCTION

#### THE WHOLESALE COMPETITION SCENARIO

- Market Participants
  - Gencos
  - Independent System Operators
  - Discos
  - Consumers
- Price Cap Regulation of Discos
- Variations in the Regulation of Discos
  - Cost-of-Service Regulation of Discos
  - Mandated Time-of-Day Pricing
- Hedging Electricity Prices Through Financial Contracts
- Guarding Against Market Power in the Power Pool

#### THE RETAIL COMPETITION SCENARIO

- Mechanisms for Price Hedging
- Variations in the Regulation of Linecos

### FACTORS AFFECTING MARKET EFFICIENCY

- Spot vs. Contract Prices and Contracting Efficiency
- Efficiency Issues in Contracting
- Comparing Wholesale and Retail Competition
- The Array of Products and Services Offered to Consumers
  - Differentiated Generation
  - Differentiated Quality of Service
  - Differentiated Energy Services
- Comparing Wholesale and Retail Competition
- Technological Improvements for Consumers and Suppliers
  - Supply-Side Innovations
  - Demand-Side Innovations
- Comparing Wholesale and Retail Competition
- Transaction Costs and Scale Economies in Distribution
  - Wholesale Competition
    - Transactions Costs
    - Scale Economies
  - Retail Competition
    - Transaction Costs
    - Scale Economies
- Comparing Wholesale and Retail Competition
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation Reserve Margins</td>
<td>31</td>
</tr>
<tr>
<td>The Role of Regulation</td>
<td>31</td>
</tr>
<tr>
<td>Comparing Wholesale and Retail Competition</td>
<td>32</td>
</tr>
<tr>
<td>Investment in Transmission Capacity</td>
<td>32</td>
</tr>
<tr>
<td>The Incentive to Free Ride</td>
<td>33</td>
</tr>
<tr>
<td>Comparing Wholesale and Retail Competition</td>
<td>33</td>
</tr>
</tbody>
</table>

### RELATIVE IMPORTANCE OF THE FIVE FACTORS

### RECOMMENDED RESEARCH STRATEGY

**Task 1:** Assess the Relative Efficiency of the Contract Markets in the Wholesale and Retail Models

**Task 2:** Assess the Relative Benefits of Product Differentiation in the Wholesale and Retail Models

**Task 3:** Assess the Magnitude of Scale Economies in Distribution in the Wholesale and Retail Models

### REFERENCES
RELATIVE EFFICIENCY BENEFITS OF WHOLESALE AND RETAIL COMPETITION IN ELECTRICITY: AN ANALYSIS AND A RESEARCH AGENDA

By

Douglas R. Bohi and Karen L. Palmer
Resources for the Future
Washington, D.C.

INTRODUCTION

A central issue in the debate over restructuring the electric power industry is the extent to which the market should be opened to competition. One aspect of this debate is whether competition ought to be restricted to the wholesale power market or extend all the way to final retail consumers. The purpose of this paper is to begin to explore the potential differences in economic efficiency that would arise between wholesale and retail competition in the electric power industry. The paper first defines the two market-structure scenarios being compared, then describes the factors responsible for differences in efficiency in the two scenarios. This is followed by an assessment of the relative importance of the factors and, finally, recommendations for pursuing further research.

Since the outcome of the analysis depends on the assumed nature of the two competitive scenarios being compared, it is important to be clear and careful from the start about exactly what is being compared and why. To begin with, the focus of the comparison is the end states of two hypothetical market structures, ignoring any transition issues involved in getting from today’s market structure to the end states. An illustration of the two end states, labeled wholesale and retail competition, is given in Panels A and B of Figure 1. The two scenarios, which are described more fully below, both contain identical independent generation (Gencos) and Independent Service Operator (ISO) sectors, where the ISO incorporates the dispatch, control, and transmission functions. The essential difference between the two scenarios, therefore, is in the distribution stage of

---

1 The authors are grateful to David Kline, Peter Fox-Penner, Paul Galen, Bill Hogan, Val Jensen, Charles Stalon, and Scott Wright for their comments.

2 Transmission services and maintenance of the transmission system may be handled by one or more separate regulated entities, outside of the domain of the ISO. The main reason for separating these functions is that they correspond to the existing jurisdictional boundaries of federal and state regulation. Federal and state regulators might pursue different policies with respect to the entities under their jurisdiction, and these different policies would be difficult to implement in practice if the functions were aggregated into a single entity. However, this does not prevent us from considering a wide range of policies applied to the different functions that are aggregated within a single regulated ISO.
the industry. The wholesale competition scenario is characterized by a set of distribution companies (Discos) that provide a bundled service (power plus wires) to consumers, while the retail competition scenario is characterized by unbundling the wire service from the power service. A separate set of companies (Linecos) perform the wire service, while consumers may purchase their power either directly from the spot market run by the ISO or from one of several competing marketing companies (Marketcos).

The choice of market structures is based on five considerations. First, restructuring with an ISO is necessary in either scenario in order to obtain effective competition. Second, consolidating the transmission grid under a single operator is necessary to eliminate the “loop flow” problem. Third, the two scenarios embody the essential features of the California Public Utility Commission (PUC) proposal for restructuring, so that the results of our analysis are relevant to the California proceeding. Fourth, making the distinction between the two scenarios at the distribution stage allows for a natural progression from wholesale to retail competition: wholesale competition may evolve into retail competition simply by altering the structure of the distribution function. This has the advantage of allowing regulators the opportunity to observe how competition in the wholesale market is working before making the move to retail competition. Finally, the sharp distinction between the two scenarios makes it very clear where the difference in efficiency benefits may arise.

THE WHOLESALE COMPETITION SCENARIO

The structure of the electricity market envisioned under the wholesale competition scenario described below represents an important departure from the current U.S. industry structure. Under the wholesale competition scenario presented here and illustrated in Panel A of Figure 1, the three major functions of the previously vertically integrated

---

3 Restructuring with an ISO eliminates the connection between the dispatcher/transmitter and the ownership of generation, and therefore eliminates the incentive to discriminate against the generation owned by competitors. Open-access requirements (such as the Federal Energy Regulatory Commission’s recent mega-Notice of Proposed Rulemaking) attempt to achieve a similar goal without deintegrating utilities, but require extensive regulation to monitor and enforce the requirements because the incentive link is not broken. For this reason the “regulatory” solution is not as effective as the “structural” solution to the market power problem.

4 See the California Public Utility Commission proposal for comment in docket R.94-04-031/194-04-032 issued on May 24, 1995. The proposal actually combines features from both of the market structure scenarios discussed here. For example, the proposal calls for a wholesale power pool in which final customers would not have, at least initially, direct access to power, but where customers could obtain real-time prices for power and enter into contracts directly with generators. In time it is envisioned that consumers would have direct access to the pool or to generators.

5 Since the 1978 Public Utility Regulatory Policies Act and the onset of state commission requirements for least-cost planning in the late 1980s, integrated utilities have increasingly turned to non-utility generators (NUGs) as new sources of generated electricity. However, limitations on transmission access of NUGs to multiple purchasers have prevented the development of a more fully competitive wholesale market for generation.
electric utilities are performed by different firms. This is primarily accomplished by vertically disintegrating existing utilities such that the generation function is separated from the distribution function and by creating a new entity that is responsible for dispatch, control, and operation of the transmission function. Under this model, generators compete for the right to supply electricity to a market that serves all of the demand from all of the local distribution utilities in a particular region. The following sections describe the functions of each of the different entities participating in this restructured electricity market as well as the relationships among these entities.

Market Participants

Gencos. Generation is supplied by a number of Gencos that compete for the business of the Discos responsible for delivering electricity to customers. The Gencos compete with each other by offering to supply to the market a given volume of electricity at a given price. The market is segmented into half-hour blocks with an array of offer prices and volumes for each segment. Transactions in this market are coordinated by means of a centralized spot market, which is run by the ISO described in the following section. Open entry into generation is allowed and, assuming generation markets are truly competitive, generators are free from regulation. Thus, the spot-market price that Discos pay for electricity is a market-determined price.

Independent System Operators. The ISO is a regulated monopolist that is responsible for the operation of the centralized spot market for electricity and for the operation of the integrated transmission grid. As spot-market coordinator, the ISO solicits offers from generators regarding the prices and amounts of power they are willing to sell during each half-hourly segment of the day (usually for the following day). The ISO also solicits purchase bids from distribution companies that specify the projected amount of load at each half-hourly segment during the day, and the prices at which the Discos are willing to purchase the projected amounts of electricity. The ISO is then responsible for ordering the supply and purchase bids and finding the market-clearing price for each half-hourly

---

6 The market structure illustrated in Panel A of Figure 1 is a simplified version of that presented in Figure 6 of Hogan and Ruff (1994), except that in this version the distribution company is the only entity authorized to sell electricity at the retail level.

7 We do not preclude the possibility of Discos owning Gencos in this model. Because most generation is dispatched centrally by the Independent System Operator, a Disco cannot discriminate against nonaffiliated Gencos. We do assume that Discos are prevented from signing contracts with affiliated Gencos and that the regulated Disco and the unregulated Genco must be in separate subsidiaries to minimize the prospects for cross-subsidies.

8 The spot market must be run by the ISO, because prices are determined by the optimal economic dispatch model operated by the ISO. In addition, a regulated entity may lower risks to participants in the market, in the same way as those in the stock and futures markets. For example, operating rules are designed to protect the interests of investors and to prevent the market operator from developing a financial interest in any of the parties that trade in the market.
Figure 1
INDUSTRY STRUCTURES WITH WHOLESALE AND RETAIL COMPETITION

(A) Wholesale Competition
(B) Retail Competition
segment of the day. All spot-market transactions will take place at the spot-market price. Thus, all Discos will pay the spot-market price plus some markup to cover transmission costs and the ISO's operating costs, and all generators will receive the spot-market price in effect during the current period.10

The ISO is the central dispatcher and control operator. As such, the ISO is responsible for real-time balancing of the transmission grid, for ancillary services such as voltage support and line-loss compensation, operating reserves, and planning for emergencies due to equipment failures. Using the offers for generation made to the spot market, the ISO will dispatch generators in a least-cost manner when balancing demand and supply in real time. The ISO will also contract with Gencos to provide spinning reserves to be used for system balancing and for emergency situations. The ISO will recover these costs in the charges it imposes on spot-market customers. The ISO will also be responsible for transmission and grid-support services, including expansion of the system.

If jurisdictional sensitivity between the Federal Energy Regulatory Commission (FERC) and state regulators requires the separation of the functions of the ISO along traditional jurisdictional lines, then the responsibility for transmission siting and expansion would remain with the states. The ISO would handle dispatch and transmission pricing, while a separate Gridco (owner electric grid) would be responsible for the physical transmission system. It may be desirable to allocate some transmission-support functions to the Gridco as well. For our purposes, however, these regulatory distinctions can be ignored. They do not affect our ability to consider different forms of regulation applied to separate functions of the ISO as if different regulatory policies were in place. Combining the Gridco and ISO into one entity is simply a matter of expository convenience.

The question of who owns the transmission grid is left intentionally ambiguous. What is important for our purposes is that control of the grid is strictly in the hands of the ISO, including decisions about system expansion. Both Gencos and Discos may own some transmission capacity, and earn a return on those assets, but neither is able to determine how the system is used. Access to transmission capacity is essentially open to all Gencos and Discos, subject to physical limits on transmission capacity and willingness to pay.

The key to efficient allocation of the existing transmission system, and to efficient expansion of the system, lies in the pricing of transmission service. Ideally, transmission

---

9 Discos also could submit bids for power interruptions, essentially offering to lower their power demand by a certain quantity if the price of power exceeds a certain level. Discos would be able to implement these interruptions by having arrangements with their customers to interrupt demand in peak periods such as air-conditioning shut-off programs utilities operate today or other types of DSM contracts.

10 Participation in the spot market run by the ISO need not be mandatory. If unaffiliated Gencos and Discos wish to enter into a bilateral contract for power delivery they should be allowed to do so as long as there is sufficient transmission capacity to accommodate the transaction, the ISO is informed of the quantity terms of the transaction, and the terms for settling imbalances between contract flows and actual flows are well established. However, it is unlikely that parties will find it in their mutual advantage to transact outside of the spot market. One party can benefit from an outside transaction only if the other party is disadvantaged relative to the spot market.
prices would be set in a way that encourages greater use of the system when it is uncongested, discourages use when it is congested, and encourages expansion of the transmission grid to alleviate congestion whenever the cost of expansion is less than the cost of congestion.\textsuperscript{11} As indicated below, the public-good properties of the network make it difficult to achieve all of these goals through the pricing mechanism alone.

One alternative to traditional embedded cost pricing of transmission service is a two-part approach. The first part is a congestion price that is achieved as part of the solution of the economic dispatch model. Bids and offers for power at nodal locations are used by the ISO as inputs into the central dispatch model, and the solution gives the value of power at different nodes. Differences in nodal prices reflect the value of transferring power between nodes (Hogan and Ruff 1994).\textsuperscript{12} The second part of the transmission price is based on the embedded cost of transmission assets and is intended to cover the fixed costs of the existing system. This cost would be collected from buyers and sellers of power who wish to purchase the right to insert or extract power from the grid at specific nodal points.

The Independent System Operator is clearly a natural monopoly and therefore must be regulated. Transmission pricing and access currently falls under the jurisdiction of the FERC, while maintenance and construction of the system falls under the jurisdiction of the states. These jurisdictional boundaries may be honored by dividing the functions of the ISO along the same jurisdictional lines. Alternatively, since the boundaries of optimal power pools are unlikely to correspond with state borders, the ISO could be regulated entirely by the FERC or by a yet-to-be-invented regional regulatory authority. Who regulates the ISO is less important for our purposes than how the ISO is regulated. Other than pricing of transmission service, we shall presume that the ISO is subject to price-cap regulation where the initial rate is set such that revenues cover costs, and rates in subsequent periods are capped according to a predetermined formula. (A more complete description of the properties and mechanics of price-cap regulation is presented in the following section.) Given the possibility of automating many of its tasks and the availability of inexpensive, high-powered computers, the costs of coordinating the spot market or power pool are expected to be modest. The possibility of automation and computer record-keeping should also facilitate oversight of the ISO's operations by the regulator.

\textbf{Discos.} Local distribution companies, or Discos, are regulated utilities responsible for distributing electricity purchased from the pool to retail electricity customers. Discos participate in the electricity spot market as wholesale purchasers of electricity and then sell this electricity bundled with distribution services directly to electricity customers. Each

\begin{footnotesize}
\textsuperscript{11} Transmission should also be priced in a way that internalizes any loop-flow externalities. In the system envisioned here, all of the loop-flow problems within the area controlled by the ISO would be internalized. However, there may still be loop-flow problems across interconnected grids that must be accounted for in pricing transmission between parties connected to different interconnected grids.

\textsuperscript{12} Oren et al. (1995) show that, in a network with parallel paths, efficient use of the system may require power to flow from higher valued locations to lower valued locations, thereby indicating a negative imputed value for the transmission service under the Hogan and Ruff model. This example illustrates why pricing of transmission service according to contract paths can lead to inefficiencies.
\end{footnotesize}
Disco has a monopoly franchise to supply electricity to all consumers in its service territory. In exchange for this franchise the rates charged by the Disco are regulated. Unless indicated otherwise, we will assume that the Discos are subject to price-cap regulation. We describe the properties of a price-cap regulatory regime below, immediately following the discussion of the role of consumers. In a later section we consider the implications of cost-of-service regulation of Discos for our analysis of wholesale competition.

Consumers. Under wholesale competition, electricity consumers who do not generate their own electricity must buy their electricity from the local regulated Disco. This requirement means that consumers do not have a choice regarding electricity suppliers. However, large electricity consumers such as large industrials may find it less expensive to generate their own electricity than to buy from the utility. These consumers would rely on the local utility for backup service only. A growing number of utilities are trying to combat the loss of customers to cogeneration by offering discounted rates to customers when they threaten to leave the system (Cross 1995). If wholesale competition results in lower electricity rates to all classes of customers, then losses of large customers to cogeneration should slow. However, if current utility practices serve to subsidize large customers at the expense of small captive customers, or if revenues from large customers are subsidizing the costs of serving small customers, competition will tend to eliminate these subsidies and the rates to the subsidized class of customers will rise.

Price-Cap Regulation of Discos

In accordance with recent electric utility restructuring proposals issued in California and Michigan, we assume that the segments of the electricity industry that continue to be regulated after restructuring will be subject to an incentive-based regulatory regime instead of traditional cost-of-service or rate-of-return regulation. Regulators in these and others states are rejecting continued reliance on cost-of-service regulation due to its poor incentives properties and the high costs of implementing such regulation. By linking rates to costs, cost-of-service regulation provides little incentive for regulated firms to reduce their costs and may bias the firm's choice of inputs. Moreover, the rate-setting process itself is costly and places a strain on limited regulatory resources.

Unfortunately for our purposes, none of the current restructuring proposals endorses a specific form of incentive-based regulation (IBR). While there has been some experience with incentive regulation that targets specific aspects of utility performance such as heat rates or baseload capacity utilization (Joskow and Schmalensee 1986; Berg and Jeong 1991), there has been very little experience with more general forms of regulation.
incentive regulation designed to replace (instead of merely augment) cost-of-service regulation.\textsuperscript{14} Given this lack of experience with more general IBR in electricity, we select a form of incentive-based regulation, price-cap regulation, that has been used with some success in the telecommunications sector (Mitchell and Vogelsang 1991). We incorporate some modifications, including some proposed by other authors (Brown, Einhorn, and Vogelsang, 1991; Vogelsang, 1994) that may make this form of regulation more suitable to the electric utility industry.

Thus, our analysis of wholesale competition assumes that regulated Discos are subject to price-cap regulation by the state utility regulator. Price-cap regulation places a ceiling on the rate of growth in the average price per kilowatt-hour (kWh) of electricity charged by the Disco. The initial price cap is usually based on the rates in place when price-cap regulation is adopted or on some cost-based rate calculation. Future values for the price cap are calculated using the following formula:

\[ p_{t+1}^* = p_t^* (1 + I - \text{prod} + d) \quad \text{and} \quad p_{t+1} \leq p_{t+1}^* \]

Specifically, the rate of price growth cannot exceed some predetermined measure of cost inflation (\(I\)) minus a predetermined expected rate of productivity growth (\(\text{prod}\)) plus an adjustment for other exogenous shocks (\(d\)) such as tax changes or unanticipated new environmental regulations. The regulated Disco is free to charge prices below those implied by the cap and to adjust freely prices for various individual services as long as the “average price” across all services (or across all services within a category) does not exceed the price ceiling. To prevent the Disco from strategically manipulating the price cap by influencing the quantity weights, average prices in each year are calculated using previous period quantities as weights, or

\[ p_{t+1} = \frac{\sum_i q_{i,t} \cdot p_{i,t+1}}{\sum_i q_{i,t}} \]

where \(p_{i,t+1}\) is the price of service \(i\) in period \(t+1\), and \(q_{i,t}\) is the quantity of service \(i\) sold in the previous period.\textsuperscript{15}

Because large differences in the expected and actual rates of inflation or productivity growth may lead to large profits or even losses, it may become necessary at some point to make additional modifications to the price cap and to these parameters in

\textsuperscript{14} In the mid 1980s, the Arizona Corporation Commission and Tucson Electric Power (TEP) entered into an agreement that initiated a five year regulatory experiment whereby TEP was subject to price cap regulation. This experiment ended in 1989 with a return to cost-of-service regulation and rate hearing request filed by TEP (Isaac, 1991).

\textsuperscript{15} Moreover, Bradley and Price (1988) and Vogelsang (1989) have shown that using past period quantities combined with current period prices of individual service to construct an average price which is subject to the price cap will lead the regulated firm to select prices that converge to Ramsey prices. This result will hold as long as demand for the individual services are independent (Brennan 1989) which is unlikely to be the case for electricity, particularly if the different services refer to different time periods during the day.
the price-cap formula. Therefore, after a prespecified period of time, the regulator will
take a closer look at the relationship between revenues and costs and make adjustments to
the current price cap that yield reasonable profits. Also at that time, the regulator will
evaluate recent productivity growth and input cost growth (perhaps at the industry level
instead of at the firm level to avoid encouraging strategic behavior in anticipation of future
reviews), and adjust the price-cap formula for use in future periods to reflect any changes
in expectations. The process and information used to make these adjustments are similar
to those used for a rate case under cost-of-service regulation. However, presumably these
reviews will be less frequent and less extensive than they were under cost-of-service
regulation.

One concern that consumers and others have raised about price-cap regulation
where the cap is applied to an overall average price is that it may not lead to "fair" prices
for individual services. For example, placing a single cap on the overall average price per
kWh could lead the regulated Disco to lower prices (relative to base-period rates) to
customer classes with high elasticity of demand (largely industrials) and raise prices to
customer classes with low elasticity of demand (largely residential). While such rate
changes may bring rates closer to Ramsey prices, they may not be politically acceptable,
particularly to those who believe that current rate structures are inherently fair. In order
to prohibit shifts of this type in the rate structure and to win political support for a price­
cap regulatory regime, the regulator may establish separate price caps for bundles of
services, where the bundles are defined by customer class. Indeed, the Federal
Communications Commission (FCC) did just that when it established price-cap regulation
for long-distance services offered by AT&T, establishing separate price caps for
residential, commercial, and 800 services (Mitchell and Vogelsang 1991). Following this
example, we assume that the Disco regulator will define separate price caps for residential,
industrial, and commercial customers.

We also assume that the Disco regulator follows the FCC in developing a method
for treatment of new services under price-cap regulation. Under FCC regulation of
AT&T, truly new services -- those which are more than modifications to existing services
-- are excluded from the price cap when first offered; thus, their rates are unregulated.
However, these services are folded into the average rate calculation at the time of the first
price-cap adjustment following 1-full year of sales of the new service. At that time, the
quantities for the new service used in the calculation of average prices are based on sales
during that first year. For new services that are essentially revised versions of existing
services, the previous year's demands for the existing services are used to weight the
prices of these new services in the price-cap formula.

The price-cap mechanism may present a credibility problem if regulated firms
believe that the regulator will not allow firms to go broke or to earn historically high rates
of profits (Brennan 1989) and therefore cannot commit to a particular price-cap rule. For
example, if the prices allowed under the price cap formula result in very high profits, then
there may be political pressure on the regulator to adjust the price cap downward before
the designated adjustment time. Electricity regulators in the UK have already succumbed
to such pressure and have announced their intention to lower the existing base-price cap
level well in advance of the next scheduled adjustment date (The Economist, March 11,
1995, p. 74). Conversely, the regulator may not be willing to allow a regulated firm to
lose money and thus may end up raising price caps in order to avoid bankrupting the firm. Both types of adjustments revert to setting regulated prices that are based on costs, which in turn may limit the firm’s incentives to reduce its costs. One method that has been proposed for mitigating this adverse incentive is for the regulator to precommit to some level of profit-sharing at the time of price-cap adjustment whereby the regulated firm knows that it will get to keep some prespecified portion of its profits at that time, which will buoy its incentive to continue to minimize costs (Vogelsang 1993). Such a provision could be part of a price-cap regime governing Discos.

Another concern about price-cap regulation is that it can provide an incentive for the regulated firm to cut costs by reducing service quality. Under cost-of-service regulation, the regulated firm will have no incentive to skimp on product quality since all cost increases are passed on to consumers. In order to guard against this type of behavior under a price cap, the regulator will need to establish additional mechanisms for monitoring and controlling service quality.

Variations in the Regulation of Discos

Cost-of-Service Regulation of Discos. In contrast to price-cap regulation, the application of traditional cost-of-service regulation to Discos may inhibit their incentives to actively pursue the lowest cost power supplies, since the Disco doesn’t earn a rate of return on power purchases and all cost savings are ultimately passed on to consumers in the form of lower retail electricity prices. This indifference to reducing costs is mitigated somewhat by regulatory lag which results in some profits to the Disco from cost-reducing activities during the period between the realization of cost savings and the associated adjustment in the regulated price. The possibility of being municipalized also could provide an incentive the keep costs down (Vince and Fogel 1995), as would the threat of large customers to bypass the Disco via self-generation of power. To the extent that regulation makes Discos indifferent to reducing the cost of their power purchases, Gencos may have a limited incentive to reduce their costs or to bid aggressively in the electricity spot market. This may limit the efficiency gains from the introduction of wholesale competition.

Mandated Time-of-Day Pricing. Another modification that could be made under either price-cap or cost-of-service regulation is the requirement that Discos time-differentiate retail electricity prices or at least offer time-differentiated retail prices as an option. The regulatory requirement would likely take the form of specifying a minimum amount of variation ranging from evening and daytime rates that reflect the average of spot-market prices to retail prices that vary half-hourly with the electricity spot price. The idea is to

---

16 If the price cap exceeds marginal cost, the regulated firm may chose to increase product quality in order to increase demand (Brennan 1989).

17 In the recent California PUC restructuring proposal (docket R.94-04-031/I94-04-032, May 24, 1995), the PUC proposes requiring utilities to offer consumers the option of paying spot-market prices for electricity instead of more traditional average prices. The proposal includes a 6-year proposed timeline beginning in 1996 for the installation of new customer metering equipment that would facilitate real-time pricing.
take advantage of some of the variation in spot-market pricing without actually allowing retail customers to purchase in the spot market. Furthermore, to the extent that differentiating prices by time of day reduces the Discos' purchased-power costs, the price-cap-regulated Disco may adopt such a pricing practice voluntarily in order to increase its profits. Voluntary adoption of time-of-day pricing is less likely to occur under cost-of-service regulation.

Mandated time-of-day pricing has the effect of increasing the range of differently priced services -- kWh consumed by time-of-day -- that must be aggregated in the construction of an average electricity price that is subject to the regulated price cap. Because such time-of-day service doesn't currently exist, the initial quantities that would be used to aggregate prices across different rate blocks would most likely be based on current electricity usage patterns. These weights would be updated in subsequent years to reflect changes in demand during different time periods throughout the day associated with the time-differentiated rates. An alternative method to assure that changes in spot-market prices of power purchased by the Disco can be passed through in retail time-of-day electricity prices would be to automatically allow the Disco to pass through these changes by essentially incorporating a purchased-power adjustment factor into the set of exogenous cost factors automatically included in the price-cap adjustment equation. The problem with this approach is that it would reduce the Disco's incentive to seek out the lowest possible sources of power.

The extent to which a Disco is required to adopt time-of-day prices will also affect its incentive to sign a contract-for-differences (CFD) with its suppliers and the types of CFDs that arise. For example, if a Disco is required to adopt a retail pricing structure with night and day prices that are seasonally differentiated, it may seek a CFD with a similar pricing structure. If the Disco is required to adopt more disaggregated time-of-day pricing, it may decide not to enter into a CFD at all since much of the purchased-power price risk is being passed through to consumers. Of course, if time-of-day pricing is optional for customers, the Disco will probably want CFDs covering some portion of its power purchases, as long as it believes that it can recover the higher average cost of electricity without violating the price cap.

### Hedging Electricity Prices Through Financial Contracts

In the wholesale competition model pictured in Panel A of Figure 1, all short-run wholesale electricity transactions take place through the centralized spot market run by the ISO. The market-clearing price in this spot market will vary over the course of the day as demand for electricity rises and falls, because higher cost generating units are brought on line and taken off line in response. Electricity spot prices may also vary with fluctuations in input prices such as increases in the market price of oil or natural gas due to changing fuel market conditions that are reflected in the bids of electricity generators. Changes in environmental regulatory laws and the possibility of new electricity-using or producing technologies coming on line also contribute to the price risk in electricity spot markets.

Reliance on spot-market prices alone imposes risks on both Gencos and Discos. The Gencos would face uncertainty about their profit margins, while, depending on the treatment of spot-price fluctuations in the price-cap model, the Discos may face
uncertainty about their ability to pass through price changes and the possibility of losing customers to self-generation.\textsuperscript{18} In order to mitigate some of the price risk, Gencos and Discos will have available to them different financial mechanisms that will help to hedge against price fluctuations. The two primary opportunities for hedging against price risk include contracts-for-differences (represented by the dashed line in Panel A of Figure 1) and trading in electricity futures markets (represented by the diamond box and dotted lines in Panel A of Figure 1).

In order to smooth price fluctuations associated with fluctuations in demand, a Disco may arrange a fixed-price (or quasi-fixed-price) electricity contract with a Genco. Recognizing that all electricity transactions take place at the market-clearing spot price, this contract, known as a CFD, requires the Genco to pay the Disco the difference between the spot price and the contract price whenever that difference is positive and requires the Disco to pay the Genco the difference whenever that difference is negative.\textsuperscript{19} The terms of the CFDs will vary, potentially allowing for more than one contract price that depends on the time of day or the season of the year, but the extent of potential price fluctuation under the contract will be significantly less than that found in the spot market. These contracts might also provide a way for the Genco to recover some of its fixed costs that will not be completely recoverable in the spot market when competition drives the spot price down to the short-run marginal cost of electricity supply. Depending on how they are regulated, Discos may be willing to pay a premium that covers these costs in order to avoid the potentially large price fluctuations inherent in the spot market.

Discos, Gencos, and consumers may also rely on transactions in the futures market to hedge against fluctuations in electricity prices. As in futures markets for natural gas and other commodities, electricity futures contracts will be standardized by specifying a given quantity of electricity (megawatt-hours) to be delivered to a specific location at a specific time, usually between 1 and 18 months in the future. In the wholesale competition model with a centralized power pool, the delivery point will be at a specific point on the power grid that lies within the jurisdiction of the pool. Gencos, Discos, or other participants in the futures markets may buy and sell multiple units of electricity futures CFD dates in the future. The prices of the futures contracts will fluctuate over time depending on the spot-market price and other factors.

Transactions in electricity futures can serve as a hedge on price risks in a number of situations. For example, if a Disco anticipates a large increase in electricity prices as a result of a fuel-market shock, it might purchase some quantity of electricity futures to hedge the price. If the futures price does rise, the Disco can either take delivery at the

\textsuperscript{18} Since the price cap is a cap on average price, prices at specific times of the day need not fall below the cap. However, if the price cap formula does not automatically pass through changes in the average spot price of electricity, the Disco faces the risk of not being able to cover its costs. However, this risk is not necessarily reduced by signing a CFD, because such a fixed-price contract will tend to have a higher average price per kWh.

\textsuperscript{19} The California PUC restructuring proposal (docket R.94-04-031/I94-04-032, May 24, 1995) does not maintain our sharp distinction between wholesale and retail competition because retail customers are allowed to sign contracts-for-differences with power generators and, in a phase-in process that depends on customer size, are allowed to purchase power at spot prices determined in the centralized power pool.
futures price or, more likely, sell the futures, thereby zeroing out its position in the futures market, and use its revenues to purchase electricity from the pool at the new higher price. A Genco could also use the futures market to hedge against a lower electricity price or against the possibility that its units might not be dispatched. By selling futures, the Genco will obtain some revenues that it can use to cushion the financial blow should the price of electricity fall. In the event that the price does fall, the Genco can do a reverse trade by buying a futures contract and offsetting its position in the market. The Genco will earn some revenues even if its units are not dispatched.

The futures market also allows an opportunity for outside parties to speculate on the price of electricity. In contrast to the hedging transactions described above, where the Gencos and Discos take opposite positions in the futures market and in the commodity market, a speculator takes a single position in the futures market. A buyer of a futures contract is betting that the spot price of electricity will rise in the future, while the seller of a futures contract is betting that the spot price will fall.

**Guarding Against Market Power in the Power Pool**

Recent debates about the role of centralized power pools in a restructured electricity industry have focused much attention on the potential for power pools to facilitate the exercise of market power within wholesale electricity markets (Hogan et. al. 1995). Critics of the development of a market structure involving a centralized pool suggest that either the pool operator, large Gencos, or large Discos will be in a position to manipulate the market-clearing prices in the spot market. The pool operator would be in a good position to manipulate the market, but to act in this way presumes that the operator will stand to gain from the outcome. Large buyers and sellers could manipulate the spot price simply because of the weight of their actions in the market. These concerns suggest that regulators need to be careful to protect against the possibility of market power when designing and setting up the power pool.

In addition to oversight by a regulator, it is important to maintain true independence on the part of the ISO to ensure that the system is used efficiently. The ISO should own no interest in a Genco, a Disco, or in the transmission system. Only by separating these functions sharply and distinctly will the ISO have no incentives to manipulate the spot market or to discriminate in providing access to transmission facilities. The ISO is purely a market facilitator that neither buys nor sells electricity.21

---

20 Incentives to manipulate the spot market could also be eliminated by establishing a competitive joint ownership of the ISO by generators and distribution companies. See Braman (1994) for a discussion of such a possibility.

21 Opponents of a voluntary centralized spot market run by a designated ISO have suggested that the ISO will have an incentive to steer market participants away from physical bilateral contracts toward reliance on the spot market. If the costs of facilitating bilateral contracts are higher than the costs of facilitating spot-market transactions, there may be efficiency arguments for steering Discos away from one form of transaction and toward another. These additional costs need to be weighed against any additional value that might be created by product differentiation or other service enhancements achieved through bilateral contracts.
To mitigate market power on the part of Gencos and Discos, it may be necessary to break up the generation and distribution components of existing utilities to form smaller units. In addition, to facilitate competition in some regions, it may be necessary to alleviate transmission constraints and to introduce additional competition by expanding the market's boundaries.

THE RETAIL COMPETITION SCENARIO

Under retail competition, electricity consumers may choose the supplier of their electricity. Consumers are no longer required to buy bundled distribution, transmission, and generation services from their local Disco.

The industry structure under retail competition is illustrated in Panel B of Figure 1. The structure of the industry is identical to that described under wholesale competition up to the distribution stage of the industry. Both structures contain the same Gencos and ISO, and control over the transmission system is the same. The key difference between the two structures is that under retail competition the spot market is open to retail electricity customers. The bundled power and distribution service provided by Discos is replaced with unbundled wire service provided by regulated Line Companies (Linecos), and unbundled power service that may be purchased directly from the spot market or from an independent marketing company (Marketco). The Linecos have a monopoly franchise within their service territory and are subject to rate regulation by the state utility regulator. Unless indicated otherwise, we will assume that the form of regulation applied to the Lineco is price-cap regulation similar to that applied to Discos under wholesale competition.

In contrast to the Linecos, the Marketcos are unregulated entities. There are few barriers to entry into the Marketco business, so this stage of the industry is expected to be highly competitive. To further this objective, Linecos are prohibited from owning a Marketco subsidiary in order to avoid the incentive for providing preferential service to affiliated Marketcos. The Marketcos purchase electricity directly from the pool and then

---

22 We are abstracting from the political and economic issues that arise in obtaining the market structures under review.

23 If under wholesale competition the regulated Disco is required to offer retail electricity prices that vary with the spot prices, an alternative discussed above, then the differences between the retail and wholesale competition scenarios are more subtle.

24 There may be a transition period during which Linecos may be allowed to offer bundled retail electricity sales and distribution services to some classes of customers (such as residential customers) who may face high transaction costs associated with dealing directly with the pool, but who have yet to attract the attention of aspiring Marketcos. Until the supply side of the Marketco business evolves, the Lineco will be allowed to serve these customers at a regulated rate. Once competitors are attracted to the market, Linecos will revert to being providers of distribution services only. As before, we ignore the problems that may arise in achieving these structural changes, because it is the end result we are comparing and not the process of getting there.
sell it to customers, either bundled or unbundled with distribution services, which the Marketcos purchase from the Linecos.  

**Mechanisms for Price Hedging**

The mechanisms for price hedging in the retail competition model are the same as those envisioned in the wholesale competition model: CFDs between buyers and sellers in the market for power and an active electricity futures market. The major difference between the two models is that, under retail competition, electricity consumers and Marketcos are the ones who sign CFDs with Gencos, and these agents are also expected to participate in the electricity futures market. Because Marketcos are unregulated profit-maximizers, their incentives with regard to these hedging activities are likely to be different from those of regulated Discos. Consumers may also have different attitudes toward price risk than Discos. Thus, the types of contracts and futures-market activity that evolve under retail competition are likely to be different from those that evolve under wholesale competition.

**Variations in the Regulation of Linecos**

In general, Linecos are expressly prohibited from owning Marketcos because of potential problems of discrimination in service quality or access between affiliated and unaffiliated Marketcos. However, prohibiting Linecos from entering into the electricity-marketing business is not the only way of combating these potential problems. For example, it may be possible to impose accounting practices and barriers to information flows between a Lineco and its Marketco subsidiary that effectively prevent this type of behavior. Indeed, natural gas pipelines are allowed to diversify into the gas-marketing business and in doing so have satisfied the FERC that such diversification will not compromise their ability to deal fairly with competing gas marketers. Therefore, as an alternative case, the analysis should consider the implications of allowing Linecos to diversify into retail power marketing for the relative efficiency of retail competition and wholesale competition.

---

25 In the unlikely event that there is only one Marketco actively doing business in a particular Lineco’s service territory, the Marketco would be required to offer both a bundled and unbundled service to all retail customers to prevent it from enjoying monopoly rents as the second-hand monopoly provider of distribution services. With access to the unbundled service, customers could purchase power from the Marketco and distribution services from the Lineco, with the latter provided at regulated rates. In this case the Marketco is subject to antitrust regulation rather than traditional utility regulation.

26 If the Lineco were subject to cost-of-service regulation, allowing entry into unregulated markets could create the additional problem of using regulated revenues to subsidize participation in unregulated markets. Under price-cap regulation, however, regulated rates are divorced from costs and, therefore, the opportunities to cross-subsidize by shifting costs are eliminated.
FACTORS AFFECTING MARKET EFFICIENCY

Spot vs. Contract Prices and Contracting Efficiency

Economic efficiency requires that consumers pay prices that are equal to marginal costs. In this case, consumers will make appropriate decisions about the volume of consumption and pay a price that just covers the cost of production.

The spot and contract markets for electricity operate the same way in both market scenarios. The only difference is the party at the distribution end of the transaction: Discos versus Marketcos and consumers. The question to be addressed in this section is whether this difference in the two market structures has any implications for efficiency.

Efficiency Issues in Contracting

Spot-market prices may be expected to vary significantly over the course of the day and year in response to changes in the level of consumption, as generating units with different marginal costs are brought on or taken off line, and in response to changes in the cost of inputs such as fuel prices. Both generators and consumers, or the Discos and Marketcos operating on behalf of consumers, may be uncomfortable with the variation in spot prices and wish to enter into a CFD that stabilizes the transaction price. For example, the generator may wish to settle on a fixed price that will cover total costs, and the consumer may wish to know with certainty the cost of electricity as an input in production or for final consumption. Thus, the transaction price of electricity may differ from marginal cost of production.

The deviation in the contract price away from the spot price need not represent a loss of market efficiency, however. The reason is that price risk is another form of cost that must be taken into account in determining efficiency.

If both parties were risk neutral, they would be indifferent between purchasing at a variable spot price and purchasing at a contract price that equaled the expected value (i.e., the mean of the probability distribution) of the spot price. This means that the parties are comfortable with variation in the spot price and there is no basis for arranging a contract.

If one party is risk averse, on the other hand, that party would be willing to pay for a contract that smoothed the acquisition cost of electricity. If the consumer is risk averse, the contract price would be higher than the expected value of the spot price by an amount equal to the cost of shifting price risk to the generator. If the generator is risk averse, the contract price would be less than the spot price by an amount necessary to compensate for the risk. In general, economic theory predicts that risks will be traded when the parties to a contract have differences in the willingness to accept risk. Contracts contribute to economic efficiency by shifting risk to those parties best able to absorb and manage it.

In principle, the amount necessary to compensate for the risk would be the cost of insurance required to protect the generator. The cost of insurance would be subtracted from the spot market price paid by buyers.

Two other motives for contracting are not expected to be important in this market: transaction costs and strategic behavior. Transaction costs in the spot market should be low, and buying from and
Comparing Wholesale and Retail Competition

An efficient contract is one that encourages Gencos to minimize their costs, encourages consumers to conserve on the use of power, and achieves an optimum trade-off between risk-bearing and the price required to shift risk to another party. A number of factors will determine whether efficiency may differ between the two market structures.

First, Discos and Marketcos are agents operating on behalf of consumers and may differ in the effectiveness with which they represent the interests of consumers in contracting with Gencos. The trade-off between risk and price will differ across consumers, and competition among Marketcos will encourage them to offer a variety of trade-off terms that appeal to customers. Those Marketcos that are not responsive to the interests of consumers will lose customers to Marketcos that are more enterprising. In addition, in negotiating terms with the Gencos, Marketcos that fail to obtain the best terms will not be competitive with those who are more successful and again will lose customers to the more enterprising Marketcos.

The lack of competition among Discos for customers will blunt the incentive to offer customers a wide variety of risk terms, and will reduce the pressure to negotiate with Gencos the best possible terms on behalf of their customers. The pressure of regulation replaces the pressure of the market as the mechanism for inducing an efficient outcome. Thus, the nature of the regulation is important in determining the extent to which Discos will maximize the gains from contracting on behalf of their customers.

Price-cap regulation will shift the risk of price fluctuations in the spot price onto the Disco and encourage the Discos to use contracts for the purpose of price hedging. In addition, price caps give Discos the incentive to bargain for lower contract prices, since any difference between the cap and the acquisition price will add to profits. The amount of price variability allowed in the contract will depend on the cost of insurance and the degree of risk aversion on the part of the Disco. The incentive provided by price-cap regulation will be to truncate the upper part of the distribution of spot prices, in order to minimize price variation in the upward direction.

It seems unlikely that price-cap regulation will allow the optimal amount of price variation, however. The reason is that the amount of price variation that is consistent with the interests of a wide range of consumers is unlikely to be captured in a limited number of price caps. This is a topic for further research, however, keeping in mind that transaction costs will place a practical limit on the variety of price-risk trade-off terms provided.

Second, efficiency in contracting is enhanced if there are many potential parties available to negotiate contracts. This condition underlies the competitive nature of the contract market because a larger number of parties implies a greater range of options in contract terms and less opportunity for the exercise of market power. The number of Gencos is the same in both market structures, but the number of power buyers is larger in the retail model. Each Disco in the wholesale model could contain many consumers and Marketcos that would contract individually with the Gencos. Moreover, because selling into the spot market removes the problem of strategic behavior, as arises in the case where a generating plant might be held captive to a specific buyer.

29 Consumers' interests are represented directly when they enter into contracts with Gencos.
Marketcos are more likely to represent the varied interests of consumers in their contracting negotiations, the variety of contracts that develop in the retail model may be wider than that in the wholesale model. Third, efficiency is enhanced when the transaction costs involved in creating and enforcing contracts are minimized. Contracts that are otherwise desirable to both parties may not be possible to consummate if transaction costs are too high. Transaction costs will increase directly with the number and variety of transactions that take place. Thus, this condition will have the opposite effect on efficiency of the first two, because the number and variety of contracts in the retail model would exceed that of the wholesale model.

The cost of acquiring information about contract options and their implications is also a form of transaction cost that can be accomplished more efficiently by some parties. For example, Gencos, Discos, and Marketcos are members of the electricity industry and will have better information than consumers. In addition, because of the fixed costs involved in acquiring information, larger entities will be able to spread those costs over more units of electricity than smaller entities. These considerations will limit the flexibility of consumers and Marketcos to enter into unique contracts, and will encourage the standardization of contracts that may apply to broader aggregates of consumers. The larger the magnitude of transaction costs and information costs, the narrower the difference in efficiency of the wholesale and retail models.

Fourth, efficiency is enhanced if risks are borne by the party that is most willing and able to accept the risk. Two considerations come into play here. First, risk should be borne by the least risk-averse parties in order to minimize the cost of insurance involved in power transactions. This condition is more likely to be achieved in the retail model because of competition among Marketcos. As a consequence, they will be more inclined than Discos to offer terms that coincide with variations in risk aversion of customers. Those customers who are least risk averse will be able to absorb price risk in exchange for lower average prices; other customers will pay the cost of less price risk. In both models, the allocation of price risk to the Gencos would require higher rates of return to investors to compensate them for risk bearing, and the allocation of risk to distributors would require higher prices for power as compensation.

The second consideration is that risk should be allocated to those parties with the most flexibility to adjust to higher prices. The primary source of price risk comes from variations in consumption and in the marginal cost of generation. Gencos have relatively little ability to control generation costs once the plant is built and the technology and fuel requirements are set. Consumers have some flexibility to alter consumption patterns to avoid high-price periods, and to allow for some degree of interruptibility of consumption when the price of power peaks. Again, the retail model is more likely to encourage these responses because of the incentive to offer terms that match consumer interests, so that

---

30 Buying contracts in the futures market may be a more feasible way to hedge against price fluctuations when the transactions costs of bilateral contracts are too high. The hedge against future price changes is accomplished by taking opposite positions in the futures market and in the commodity market. Both parties will pay a transaction cost and the price hedge will not be perfect (i.e., some price risk will remain). Moreover, futures markets typically become illiquid when the term of a contract exceeds a year, so that hedging is restricted to a shorter term than may be desired under a bilateral contract.
consumers with the greatest flexibility to alter consumption will absorb the most price risk; however, a definitive answer depends on the nature of regulation of the Discos.

The Array of Products and Services Offered to Consumers

Economic efficiency is enhanced when consumers are offered more diversity in the array of products and services that are available, and when that diversity is responsive to consumers' willingness to pay for differentiation. We will consider the following three classes of differentiated electricity service in the context of the two competition scenarios.

Differentiated Generation

Generation facilities are differentiated by technology, fuel type, and environmental impacts, among other factors. Under the wholesale competition model, consumers are not allowed to enter into contracts directly with Gencos. Therefore, short of generating their own power, it would be very difficult for individual consumers to purchase differentiated generation unless the Disco has an incentive to contract with different sources of generation. The Disco could offer a menu of generation sources, such as generation supplied by various conventional sources or by "green" (i.e., renewable) technologies, where supply would be arranged through contracts with each generator in relation to consumer demand.

Under price-cap regulation, the Discos may be discouraged from offering a full range of generation options. Those options with marginal costs in excess of the price cap would be excluded, for example, even if consumers were willing to pay a "premium" price. In addition, if the volume of consumer demand for specialty sources of generation were small, the transaction costs involved in arranging a multitude of contracts with many generators could be prohibitive.

Under retail competition, consumers can enter into CFDs with Gencos. These contracts provide a mechanism for consumers to express directly their willingness to pay for power supplied by a particular type of Genco, and provide Gencos with a better opportunity to differentiate themselves in the marketplace. One of the provisions of the contract might be that the Genco bid low in the spot-market auction in order to increase the chance of being dispatched. Shortfalls in revenue resulting from such a bidding strategy would be made up in the contract price. Marketcos can use similar contract forms to affiliate themselves with particular types of Gencos. Because consumers can contract directly with the Genco or Marketco of their choice, it is no longer necessary to obtain a consensus among all the consumers in a particular geographical jurisdiction to participate in the same contract, as in the wholesale competition model.

---

31 Consumers preferring a specific source of power (e.g., green power) are unlikely to increase their consumption of preferred power relative to generic power. Thus, if green power has a marginal cost of production higher than that of generic power, it is unlikely that the Disco could earn marginal revenues in excess of marginal costs by offering green power.
Differentiated Quality of Service

All of the power supplied to the spot market is of homogeneous quality and will differ only in terms of the price. However, the quality of service provided by the Discos and Marketcos can vary, depending on the flexibility, firmness, and duration of service provided to the customer under a specified set of contract prices. Flexibility refers to the degree of variability in contract demand, firmness refers to the prospects for interruption, and duration refers to the length of the contract. Higher quality of service generally implies a higher average cost of service.

Under either competitive model, residential consumers may be expected to have the highest quality of electricity service because they are free to vary their consumption of electricity as much as they desire and are free to terminate their relationship with the utility at any time. Likewise, they may be expected to pay the highest average price for electricity. The potential variations in service quality offered to industrial consumers are more numerous. Industrial customers would choose the preferred alternative according to the trade-off between the quality of service and the price that best suits their needs.

Under the wholesale model, the extent to which Discos will offer different levels of service quality to their customers is largely a function of how Discos are regulated and of the Discos’ use of CFDs to smooth out fluctuations in the spot price of electricity. Under price-cap regulation, the Disco will have an incentive to reduce its purchased power costs and therefore may be more inclined to offer interruptible service contracts than if it faced cost-of-service regulation. The incentives would be consistent with economic efficiency if the Disco can increase its profits by reducing power purchases at times when electricity is most expensive. If consumers respond to service interruptions by shifting demand for utility-supplied electricity from peak periods to off-peak periods, these contracts will have the additional effect of increasing demand for underutilized generation and distribution facilities during off-peak periods. On the other hand, a Disco that bears more of the risk of cost fluctuations under price-cap regulation may have a stronger incentive to sign a CFD with a Genco than it would under cost-of-service that, in turn, will reduce its incentive to offer interruptible service to its retail customers.

Under traditional cost-of-service regulation, the incentives to offer different levels of service quality come largely from the nature of the rate structure and regulatory lag. For example, if the costs of obtaining power during peak periods are high compared to the price at which power is sold during those periods, Discos will have an incentive to offer interruptible service contracts. However, the Disco’s incentive to offer interruptible service contracts may be limited somewhat if it is a party to a CFD that limits the purchased power price extremes that it faces during peak periods.

The ability of consumers to communicate their willingness to pay for electricity service of varying quality is linked to the incentives that Discos have for offering such arrangements. In general, the more flexibility regulators allow Discos in the definition of services and the design of rates within the parameters of regulatory restrictions on aggregate prices or profits, the more opportunities consumers will have to communicate their demands to Discos.

Under retail competition, consumers who purchase power directly from the electricity spot market run by the ISO may prefer the flexibility of deciding whether or not
to cut demand in response to increases in the spot price of electricity at the time they occur, rather than agreeing in advance to be interrupted when the spot price rises to a particular level. If consumers are paying spot prices for electricity, then they will have an incentive to reduce demand on their own during peak periods when the spot price of electricity is high.

However, some consumers may opt for a combination of a lower average electricity price and a higher probability of interruption, where the combination is negotiated in a contract with either a Genco or a Marketco. In both cases, the contract would include an agreement to be interrupted whenever the spot price reaches a certain level. To compensate consumers for making themselves available for interruption, they would be offered a lower fixed price.

**Differentiated Energy Services**

The differentiated energy services referred to here include cooling, heating, and lighting, as well as sophisticated load monitoring, remote control appliances, and other demand-side management (DSM) devices.

With regard to DSM, both Discos and Marketcos will have an incentive to find methods to reduce customer consumption when the spot price of electricity is high, especially for those customers who do not pay real-time prices for electricity. The relative strength of the incentive for Discos to offer DSM will depend on the reward system provided by the regulator, so that the savings created by DSM translate back into profits for the Disco. For Marketcos, the reward system is clear.

If the incentive structure imposed on Discos by regulators emphasizes the volume of power sales as the means to earn a profit, as in traditional regulation, the Discos will perform in a manner similar to that of vertically integrated utilities who make their profits by increasing generation. In this circumstance, Discos will be less inclined to offer consumers products that economize on power sales (except when the spot price of power is very high). Experience suggests that the regulator will not be able to offset the incentive to sell power without providing Discos incentives to encourage them to market DSM options.32

The extent of consumer demand for such services will depend on the variability of electricity prices they face. For example, if retail electricity prices fluctuate with the movement in spot prices, consumers will be eager to purchase energy service contracts that reduce their use of electricity during peak periods and limit their exposure to peak electricity prices. The prospects for price variability are narrowed under the wholesale model when compared to the retail model, and consumers that buy at average rather than real-time prices will not be inclined to conservation measures.

Under the retail competition model, where power is purchased from Marketcos or directly from the spot market, consumers will be inclined to seek combinations of power

---

32 One alternative to price-cap regulation that has been proposed to encourage utility investment in DSM is the Bill Cap mechanism (Regulatory Assistance Project 1994). This mechanism imposes a cap on average revenue per customer instead of average revenue per kWh, thereby providing an incentive to limit the sale of electricity to meet the cap.
and other products that will minimize the total cost of energy services. Consumers who participate directly in the spot market may be particularly interested in power-storage technologies that could help reduce electricity demand during high-priced peak periods. Consumers who sign CFDs with generators that assume a firm electricity supply may have less of an incentive to clip or shift peak demands, but will face a higher average price for electricity than those who deal exclusively with the spot market or those who negotiate some degree of interruptibility.

Incentives to supply differentiated energy services are also different under retail competition. In contrast to the regulated Discos whose profits arise from electricity sales only, the unregulated Marketcos stand to earn a profit whether they sell power or other products to consumers. Moreover, because the Marketcos are in competition with each other, they will attempt to differentiate their energy service offerings as much as possible to gain a competitive advantage. Finally, because Marketcos are allowed to use CFDs to hedge the price they pay for electricity, they will have a greater willingness to offer integrated energy services than will companies that must buy electricity at regulated rates from local Discos.

**Comparing Wholesale and Retail Competition**

There are several distinctions between the wholesale and retail competition models with regard to the incentives for product differentiation in general and for the specific classes of products identified here. On the demand side, it is very difficult for individual consumers to express their willingness to pay for power supplied by particular types of generators in the wholesale competition model; thus, some potentially valuable service distinctions may be excluded from the market under this scenario that would be available under retail competition. Also, the actions of consumers may differ depending on whether they are purchasing their power supplies and related equipment from a regulated Disco or from unregulated firms.

Distinctions regarding supply-side incentives also arise in the two models because Discos are regulated while Marketcos are not. First, the Marketcos have a profit incentive to provide consumers with either power or substitute (and complementary) products, while the Discos when regulated under price-cap or cost-of-service regulation only have an incentive to provide power (and complementary products). Second, there is easy entry into the market for new Marketcos, but not for competitors of the Discos. Competing independent energy service companies have limited incentives to enter because they are unable to use CFDs with generators to hedge the price of electricity. However, under the retail competition model, the increased competition among Marketcos will provide the stimulus on the supply side to create new products in order to attract consumers.

**Technological Improvements for Consumers and Suppliers**

The relative economic efficiency of the two scenarios depends in part on the implications of each scenario for dynamic efficiency in electricity and related markets. Dynamic efficiency is enhanced if market participants face the appropriate incentives to
develop new technologies that will lower the cost of supplying electricity, lower the cost of using electricity, or expand the range of productive uses of electricity. Incentives will be appropriate if innovative activities are pursued up to the point where their marginal benefit to society is equal to their marginal cost.

It is a well-accepted proposition that there are market failures related to research and development (R&D) and other innovative activities that lead to socially inefficient levels of innovation arising in the marketplace. In particular, imperfect patent protection and resulting spillovers to firms other than the innovator will lead firms to innovate less than the socially efficient amount because they are unable to internalize all the benefits of their efforts to society. However, the workings of the patent system are unaffected by the two scenarios considered here. Therefore, we will ignore this spillover problem. Instead, we will focus on other barriers to dynamic efficiency that are affected by regulatory and market structure differences between the two models.

The incentives for and barriers to technological improvements will differ between innovations developed for the consumer market and innovations developed for the power supply market. The former category includes new energy-efficient appliances and equipment, new ways to use electricity such as electric vehicles, and new ways to store electricity at the customer's location in order to shift load from high-cost periods to low-cost periods. The latter category includes more efficient generation technologies, new methods of monitoring transmission system use and electricity consumption, new methods of metering electricity use, and new electricity storage technologies that might reduce supply costs in peak periods. The boundaries between these two categories are overlapping. For example, consumers may place some value on the introduction of a new electricity supply technology that goes beyond its impact on the cost of electricity. More specifically, consumers might place a high value on the development and adoption of a new clean coal technology that allows local Gencos to burn local coal with minimal impact on the environment, even if the impact of that technology on the price of electricity were negligible or positive. Thus, some innovations may fall into both categories.

In the next two sections, we review the barriers to innovation under each market structure with regard to supply-side and demand-side innovations. The third section summarizes the differences between the two models. Most of the barriers to innovation identified here are dynamic analogs to the disincentives for static product differentiation discussed in the previous section. Where this is the case, the discussion will be more abbreviated.

---

33 In recognition of the nonappropriability problems associated with R&D and the existence of economies of scale associated with some major types of research efforts, the electric utility industry established the Electric Power Research Institute (EPRI) to conduct research that would be of value to all participants in the industry. EPRI is funded by contributions from member utilities and the results of its research efforts are available to all members. As the electric utility industry becomes more competitive, the expectation exists that support of EPRI's research efforts will decline. The implications of the two scenarios considered here for support of EPRI's research activities are difficult to predict. For purposes of this discussion, we ignore the impact of a shift from wholesale competition to retail competition on the level of EPRI-supported research.
Supply-Side Innovations

Under wholesale competition, the major barrier to the development of new supply-side technologies is the regulatory regime that governs the Discos. The creation of a competitive generation market under both wholesale and retail competition creates an incentive for Gencos to seek and adopt new technologies that reduce the costs of generation. Discos’ incentives to seek and adopt new technologies will depend on how they are regulated. Under price-cap regulation, the Disco will have an incentive to adopt new technologies for storage, monitoring, metering, or distributing power that reduce its costs because it will be able to retain some of the profits from this decision. The Disco’s desire to minimize its costs under price-cap regulation will also provide incentives for its suppliers to develop new technologies. Also, the Disco will not be biased away from any particular type of innovation because price-cap regulation does not discriminate across inputs.34

On the other hand, cost-of-service regulation will tend to dampen the regulated firm’s incentive to develop or adopt a new cost-reducing technology because all cost reductions are ultimately passed on to consumers in the form of lower prices. This dampening effect may be enhanced by inflexible regulatory depreciation rules that extend the life of a regulatory asset beyond its economically useful life. On the other hand, regulatory lag will provide a positive incentive for adopting cost-reducing innovations. Indeed, if the regulator allows the regulated firm to incorporate any R&D expenditures into its revenue requirement, then the firm will have an incentive to search for new technologies, although it may not have the appropriate incentive to discriminate across various opportunities. Cost-of-service regulation could also bias the Disco away from adopting innovations that are less capital intensive than current technologies if the allowed rate of return exceeds the cost of capital to the Disco.

The introduction of retail competition will probably have a very small effect on the incentives related to pure supply-side innovations. The status of Gencos is largely unchanged in this regime. However, if the generation market becomes more competitive when purchases by regulated Discos are replaced by purchases by unregulated Marketcos and Consumers, there may be a greater incentive for Gencos and their suppliers to innovate.35 Marketcos that become major players under this regime will tend to focus their innovative efforts on consumer-oriented or demand-side innovations. Of course, to the extent that Marketcos rely on sophisticated monitoring or metering technologies for the delivery of energy service products, they will be interested in innovations in these technologies that reduce their costs of providing existing services or expand the range of

34 In the past, many state PUCs have experimented with forms of incentive regulation that focus on particular aspects of the integrated utility’s behavior, such as its heat rate (Joskow and Schmalensee 1986). Throughout this paper we assume that incentive regulation focuses primarily on price growth (either through a price cap or a “yardstick” method) and therefore, it provides no incentive for the Disco to be biased either for or against a particular input.

35 The introduction of retail competition may also create an impetus for innovation in generating technology if consumers are looking for a particular type of generation. This case is discussed under demand-side innovations below.
services that they are able to offer. Competition among the Marketcos should inspire research of this type.

**Demand-Side Innovations**

Under price-cap regulation there are few regulatory barriers to demand-side innovations. Moreover, because new services are exempt from the price cap and therefore from price regulation for at least 1-year after they are adopted, the particular brand of price-cap regulation proposed here may be particularly encouraging of demand-side innovations that lead to new services. However, there are nonregulatory barriers to demand-side innovations under wholesale competition that result primarily from the limited ability of energy service companies (ESCOs) to hedge the price of electricity for long periods of time and the limited ability of consumers to express their demand for particular products.

While ESCOs will presumably have access to the electricity futures market, they will not be able to enter into longer-term CFDs with Gencos. The absence of this longer-term contract option raises the risk associated with the development of new energy-efficient appliances or new electricity-using technologies, thereby lowering the incentives for risk-averse ESCOs and other equipment suppliers to innovate.

Analogous to the product differentiation story above, the inability of consumers to express their demand for differentiated generation may limit the incentives of Gencos and their suppliers to seek out new generation technologies for which consumers may have a preference, such as certain types of renewables. The retail pricing structure of electricity will also affect the types of products and services demanded by electricity consumers. For example, if regulated Discos don’t differentiate retail electricity prices by time of day, then consumers will be unlikely to demand new small-scale energy storage technologies that would assist in shifting demand from peak periods to off-peak periods.

Under retail competition, both of these nonregulatory barriers are substantially reduced if not eliminated. Allowing Marketcos to sign CFDs with Gencos will reduce the level of price risk and create incentives for developing new products. Allowing retail consumers and Marketcos to sign CFDs directly with Gencos will provide a means for individual consumers or groups of consumers to express their willingness to pay for generation technologies with particular characteristics which, in turn, should increase incentives for innovation. Competition among Marketcos should also provide an impetus for developing new energy-using or energy-saving technologies that consumers will find valuable.

**Comparing Wholesale and Retail Competition**

Overall, the incentives for innovation appear likely to be stronger under retail competition than under wholesale competition, particularly for innovations related to the consumer market. The stronger incentives for innovation under retail competition come primarily from three sources. First, any increase in the competitiveness of the generation market associated with the switch from wholesale competition to retail competition will provide increased incentives for the development and adoption of new generation
technologies. The magnitude of this impact is likely to be greater if Discos are regulated under cost-of-service regulation than under incentive-based regulation.

Second, allowing Marketcos and ESCOs to hedge electricity prices using CFDs will increase their willingness to supply energy services that require electricity as an input. Growth in the market for these services presumably will bring with it growth in innovation to develop new equipment that will lower the cost, improve the quality, or increase the range of energy services offered in the market.

Third, allowing consumers direct access to generation markets will provide them with the means to express their willingness to pay for differentiated products and services including different types of generation, different qualities of electric service, and different bundled energy services. As these demands are revealed to existing and potential market participants, these agents will be in a better position to develop and offer new services and products that correspond to consumer demands.

Transaction Costs and Scale Economies in Distribution

Differences in transaction costs that arise in arranging for the distribution of power to final consumers is another possible source of difference in the relative economic efficiency of the two competitive scenarios. Transaction costs refer to the costs of arranging and consummating transactions, including the cost of resolving disputes, in contrast to the cost of producing, transmitting, and distributing electricity. Higher transaction costs imply a cost advantage for market structures that allow for a smaller number of transactions, such as those achieved through vertical and horizontal integration. In addition, there may be economies or diseconomies of scale in distribution that will affect the cost advantage of larger or smaller firms.

Wholesale Competition

The Discos that appear in the wholesale model are meant to correspond initially to the distribution functions of existing vertically integrated utilities. Once divorced from their parent companies, however, there is no reason to believe that this original size will continue to be appropriate. Independent Discos will rely on the market instead of internal transactions for obtaining wholesale power supplies and, as a result, will have different cost structures than their integrated predecessors. If the regulator is successful in providing the Disco with the appropriate incentives to achieve a size that minimizes the combination of transaction costs and other costs of delivering electricity, then economic efficiency considerations will become paramount in determining whether separate distribution companies will combine together or disintegrate into smaller units to form the Discos shown in Figure 1. Two important considerations include the size of transaction costs and the extent of economies or diseconomies in the distribution function.

Transaction Costs. Discos participate in transactions with several different parties and these transactions involve different types and levels of potential transaction costs. These transactions include power purchases from the spot market operated by the ISO, CFDs...
with the Genco, bundled power sales to electricity consumers, and regulatory negotiations
and other interactions with the regulator. In the following paragraphs, we consider each
of these transactions and the factors affecting the size of any associated transaction costs.

Spot-Market Transactions. The transaction costs of dealing with the centralized
spot market should be quite low, particularly since Discos presumably have experience
with forecasting load and anticipating consumer responses to electricity price changes.
Transactions in the spot market will be defined in standardized terms including quantity,
spot price, and duration and, with the exceptions of differences resulting from transmission
constraints, the price of all kWh’s purchased from the spot market at any particular time
will be the same. Of course, the vertical disintegration of Gencos and Discos expected
under the wholesale-competition model eliminates any vertical economies from the joint
supply of generation and distribution services that might have existed with traditional
vertically integrated utilities. Recent research by Kaserman and Mayo (1991) and Kwoka
(1995) suggests that there may be vertical economies from integrating generation and
distribution. However, these authors all indicate that their research remains silent on
whether the benefits of competition in generation are likely to outweigh any increased
costs from vertical disintegration.

CFDs with Gencos. CFDs between Discos and Gencos involve a wide range of
potential transaction costs. These costs include the cost of identifying potential
contracting partners, the cost of negotiating the contract, and the cost of enforcing the
contract. Many of these costs, such as the cost of the necessary legal talent to write an
enforceable contract, may be independent of the size of the Disco. Thus, larger Discos
will be better able to bear these costs by spreading them across a larger group of
customers than smaller Discos can.

Power Sales to Consumers. The transaction costs associated with selling power to
consumers include the costs of establishing service and billing accounts for individual
consumers, the costs of metering power use by consumers, and the costs of handling
consumer complaints about service quality. The implications of these costs for efficient
firm size, all other things equal, are difficult to predict. Some transaction costs, such as
the cost of metering power use, will vary with the number of customers. Other costs will
be relatively fixed and the cost per consumer will decline as the size of the Disco increases.

Regulatory Negotiations and Reporting. The costs of interacting with regulators
will depend on the nature and scope of the regulation faced by the Disco. In general, the
costs will be higher the more comprehensive the regulation and the more cost information
required to set regulated rates. Thus, a Disco probably faces higher regulatory transaction
costs than a Lineco (see below for more discussion), and a Disco subject to cost-of-
service regulation faces higher regulatory transactions costs than one subject to incentive
regulation. At least some of these costs are relatively fixed in relationship to the size of
the Disco, suggesting that larger Discos will bear them more easily than smaller Discos.

Most of the transaction-cost considerations identified above suggest that larger
Discos may have lower aggregate transaction costs per kWh than smaller Discos.
However, regulators may be reluctant to allow Discos to consolidate for fear of creating
market power on the buyers’ side of the electricity spot market. Therefore, any
transaction-cost efficiencies that might result from consolidation of Discos must be
weighed against the inefficiencies resulting from creating a monopsony (or oligopsony) situation in the electricity spot market.

Scale Economies. In addition to transaction-cost considerations, the efficient size of the Disco also depends on the extent of economies or diseconomies of scale in the supply of distribution services. A few econometric studies have addressed this question (Nelson and Primeaux 1988; Roberts 1986; Kaserman and Mayo 1991). Although all of these studies find evidence of economies of scale associated with increasing the number of kilowatt-hours distributed to customers,36 Roberts finds that there are no economies or diseconomies of scale associated with increasing the number of customers in a distribution company’s service territory holding output per customer and density constant. Nelson and Primeaux find evidence of diseconomies of scale associated with increasing the number of customers served by the utility. The latter set of findings, based on data for a sample of municipal utilities, suggest that there may be gains associated with reducing the size of utility service territories.

While suggestive, these results are not directly transferable to the question of optimal size for stand-alone Discos for several reasons. First, these studies use data from the late 1960s through the very early 1980s that predate many important developments in generation technology and other regulatory developments that affect utility costs. Second, the data used in all of these studies come largely from vertically integrated utilities that engage in generation, transmission, and distribution. Third, any empirical study based on existing regulatory regimes and market structures may have limited applicability to the wholesale competition model contemplated here. Despite the third limitation, it still could be valuable to investigate this question of economies of scale in distribution using a methodology similar to that developed in Roberts, but with more recent data for electric utilities, probably focused on municipals that engage exclusively in distribution.37

Retail Competition

The separation of the distribution and marketing functions under retail competition increases the number of transactions and, therefore, the number of possible transaction costs at the distribution end of the market. The expanded range of transactions and the factors affecting associated transaction costs are described in the paragraphs below. The implications of this vertical disintegration for scale economies in distribution are less clear, as can be seen from the discussion below.

Transaction Costs. The expanded market access under retail competition means more transactions and more potential transaction costs. In particular, the range of transactions

36 The finding that there are scale economies in the volume of power delivered to a fixed number of customers suggests that regulators should not allow competition between distribution companies for the business of existing customers.

37 The data set for municipal utilities constructed by John Kwoka contains 1989 data for 230 municipal utilities that are exclusively distribution utilities. These data might be appropriate for such a study.
is expanded for three sets of market participants: Linecos, Marketcos, and consumers. Regulated Linecos sell services to Marketcos and consumers and negotiate with and report to regulators. Unregulated Marketcos purchase power from the spot market, sign CFDs with Gencos, and sell their services to consumers. Consumers who don’t deal with Marketcos can purchase power directly from the electricity spot market and sign CFDs directly with Gencos. In the following paragraphs, we consider transactions involving Linecos, Marketcos, and consumers, but not Marketcos and Linecos. We once again focus on the factors affecting the size of transaction costs in each of these situations.

**Lineco Sales to Marketcos.** Linecos sell wire services to Marketcos for resale to consumers as a part of a bundled electricity service package. Because the rate at which these sales take place is presumably a tariffed rate approved by a regulator, the size of the transaction costs associated with these sales will depend in part on the amount of flexibility allowed by the regulator. The transaction costs of these sales are likely to be greater the more Marketcos the Lineco has to deal with; therefore, the size of these transaction costs has implications both for the optimal size and number of Marketcos and Linecos.

**Lineco Sales to Consumers.** The transaction costs associated with sales to consumers are similar to those identified above for Disco sales to consumers and are equally relevant for Marketco sales to consumers. The implications of these costs for the efficient size of Linecos are equally difficult to predict.

**Regulatory Negotiations and Reporting by Linecos.** The costs for Linecos of dealing with regulators are likely to be less than those for Discos, regardless of the overall regulatory structure. This follows because the scope of activities subject to regulatory oversight is much narrower for a Lineco than for a Disco. The extent of reduction in transaction costs remains an empirical question to be investigated, perhaps by appealing to other countries such as Norway that have implemented retail competition in electricity.

**Marketco Purchases from the Spot Market.** The size of the transaction costs here are likely to be small for the reasons stated above in the section on Disco transactions with the spot market.

**CFDs Between Marketcos and Gencos.** The discussion regarding CFDs between Discos and Gencos is relevant here as well. Large Marketcos are likely to have lower transaction costs per customer than small Marketcos. Free entry into the Marketco business suggests that there are likely to be several Marketcos serving customers in a particular geographic region that might have been served by only one Disco. However, because Marketcos need not specialize in a particular geographic area, and indeed may choose not to (particularly if they see their role as lowering costs by aggregating demands across heterogeneous groups of customers), they may be able to spread these transaction costs across a wide customer base.

**Consumer Purchases from the Spot Market.** In order to participate directly in the electricity spot market, customers must be able to predict their demand for electricity for each half-hour period throughout each day or, as in the United Kingdom, turn over this responsibility to the ISO to forecast on their behalf. Forecasting costs (or errors) are likely to be higher for small residential and commercial customers who have variable demand levels throughout the course of the day. The existence of these costs means that most of these small customers will arrange to purchase their power from a Marketco that
will perform these functions for the entire group of customers that they serve, presumably at a lower average cost than that required to serve each customer individually. Larger customers with more predictable loads will be better able to predict their demands at different times throughout the day and, therefore, will be more inclined to participate in the spot market directly. Thus, as long as there are Marketcos that will serve consumers, the fact that a consumer deals directly with the spot market is a signal that the related transaction costs are low.

**CFDs Between Consumers and Gencos.** If the transaction costs associated with entering CFDs directly with Gencos are high enough, then consumers are likely to seek alternatives such as purchasing electricity from Marketcos that can consolidate the business of many different consumers. Therefore, the only consumers likely to sign with Gencos are those who are unable to find the degree of price stability they want from a Marketco or those for whom the transaction costs of dealing with Gencos directly is quite low.

**Scale Economies.** The economies of scale associated with the Lineco function are likely to be similar, but not identical, to those for the Disco function. The major source of scale economies in both cases is the fixed cost of the distribution wires. However, there may be some fixed costs associated with being a power merchant that feed into the scale economies calculation for Discos that do not affect the costs born by Linecos. Alternatively, it could be the case that distributing power purchased by a wide number of independent agents including Marketcos and consumers could be more costly than distributing power by a distribution company (or Disco). The latter outcome would suggest that there might be vertical economies associated with being a power merchant that are foregone in the retail competition case. The relative importance of these different influences on the costs of Linecos and Discos is difficult to predict.

Any scale economies associated with the Marketco function are likely to be a result of high fixed-transaction costs of the types discussed in the previous section. Because these transaction costs are generally much smaller than the fixed costs associated with the physical distribution system, the minimum efficient size of a Marketco is likely to be much smaller than that of either a Disco or a Lineco. Therefore, there are likely to be several Marketcos functioning in any Lineco service territory.

**Comparing Wholesale and Retail Competition**

The wholesale competition model involves fewer transactions than the retail competition model because Discos provide both distribution service and merchant service. A reduced number of transactions suggests that transaction costs are likely to be lower under the wholesale competition model than under retail competition. On the other hand, the transaction costs associated with the regulatory function are likely to be greater under wholesale competition and the two costs will tend to be at least partially offsetting.

Many of the potential transaction costs in the retail competition model are mitigated by Marketcos, which substitute a small number of larger transactions between themselves and the spot market or Gencos for a large number of transactions between consumers and either of these institutions. Allowing free entry in the Marketco business
means that if consumers choose to deal directly with the spot market and with a Genco, the transaction costs associated with this deal must be more than offset by the value of operating alone as opposed to going through an intermediary.

The implications of these two scenarios for scale economies in power delivery are likely to be similar, but not identical. The efficient size for a Disco may be larger or smaller than the efficient size for a Lineco, depending on the existence or absence of vertical economies associated with the power merchant function. If there are large cost savings from being integrated into sales, then a Lineco may need to be slightly larger than a Disco in order to recover efficiently the costs associated with coordinating the distribution of power purchased by a number of independent purchasing agents.

The extent of scale economies in the merchant function itself is likely to be small, suggesting that in a given geographical region there are likely to be more Marketcos than Discos.

**Generation Reserve Margins**

Restructuring along the lines contemplated in Figure 1 will undoubtedly result in lower generation reserve margins compared to existing levels, simply because the implied horizontal aggregation across the operations of several utilities will enable the ISO planners to exploit the scale economies that come with shared reserves. The question at hand is not how reserve margins might compare with the existing situation, but how they might compare between the two hypothetical market structures. The answer appears to depend upon assumptions about regulation.

**The Role of Regulation**

At least two kinds of reserves are involved here: operating reserves, which refers *inter alia* to load-following capability and to the ability to compensate for an interruption in operating capacity, and planned reserves, which refers to the margin added to capacity above that necessary to meet forecasted peak demand. Operating reserves are a form of ancillary transmission service, which is the responsibility of the ISO. The ISO may be expected to contract for these services from the Gencos, at a level consistent with reliability standards established by the regulator, and pass on the cost to consumers through the spot price. The process for establishing reserves, and the reserve margin itself, should be the same in both the wholesale or retail models.

In contrast to operating reserves, the market is expected to provide investors the necessary incentives for additions to planned reserves. If past experience is any guide, the regulator will be skeptical about the ability of the market to provide for adequate generation reserves and may impose on the ISO a requirement to build a safety margin. Excessive margins (from investor’s point of view) will reduce if not destroy the market incentive for private investment in additional capacity. Consequently, capacity requirements imposed by the regulator may be offset by reductions in voluntary investment, and the attempt to add a safety margin will be frustrated by the market’s response.
This represents one area where an incompatibility may arise between regulatory goals and market incentives. There is no loss in efficiency if the regulatory requirement for reserves exactly matches the offsetting private investment in both magnitude and kind. However, if the regulator is right about the failure of the market to provide sufficient incentives to private investment in generation (which is a proposition that has not been supported as yet), there is a legitimate need to make up the deficiency and any attempt to do so will be at least partially self-defeating.

A related issue arises in relying on the market to achieve the optimal level of generation capacity because transmission investments are substitutes for generation investments, and the regulator necessarily will be involved in determining the amount of transmission investment. Increased transmission capacity will in general reduce the need for additional generation capacity and vice versa. For example, added transmission capacity will reduce congestion on the grid and, in turn, reduce locational differences in power prices. Consequently, the incentive to build generation to take advantage of locational price differences will be reduced.

Because generation and transmission are to some extent substitutes, investment in excess transmission capacity offers an alternative to the urge to build generation safety margins. In effect, excess transmission capacity enables the market to get the most out of available generation facilities.

**Comparing Wholesale and Retail Competition**

The two market structures are unlikely to lead to a difference in the volume of generation reserves. The need for operating reserves for reliability purposes is essentially the same in both models, while the need for planned reserves stems from concerns about the market's ability to provide adequate incentives.

The major difference between the two market structures lies in the incentives for investment in transmission capacity. As described in the next section, the retail competition model is likely to lead to less private investment in transmission capacity than the wholesale model. As a consequence, the regulator (operating through the ISO) has a greater responsibility in the retail model to make up any deficiency in investment. Any difference in efficiency depends on the ability of the regulator (and ISO) to fulfill this responsibility.

**Investment in Transmission Capacity**

Both models envision that the impetus for investment in transmission upgrades will come from the parties who will benefit from the investment, and that these parties will be expected to pay for the investment. The question is whether the optimal amount of investment is likely to take place and whether the answer will differ between the two market structures.

---

38 We are grateful to Bill Hogan for suggesting this issue.
The Incentive to Free Ride

When there is congestion on the transmission grid, spot prices for power will vary according to the location of the nodes that connect generators and customers to the grid. Some generators will receive prices for their power that are less than other generators and some customers will be required to pay higher prices for the power they receive compared to other customers. Power cannot flow from the low-cost generators to the high-valued customers to remove these price differences because of congestion on the grid. These differences in nodal prices represent congestion (i.e., scarcity) rents. Congestion rents provide the economic incentive for parties to get together and invest in additional transmission facilities that will relieve the transmission constraint. Ideally, the investment in new capacity will take place when the flow of congestion rents exceeds the cost of amortizing an investment that will remove the rents.

Working against this ideal outcome is the fact that investment in transmission that relieves a constraint somewhere on the grid will generally benefit all parties connected to the grid. When there are multiple parties that may be interested in investing in new facilities, some parties may hold back and wait for others to take action first. Those who hold back will try to “free ride” on the improvements paid for by others, causing the total amount of investment to be less than it should be to satisfy the benefit-cost criterion. Another way of looking at the problem is that the benefits of transmission expansion to all parties connected to the grid will exceed the benefits that accrue to any subset of parties that are expected to the investment. Applying the benefit-cost rule, it follows that individual investors will not invest as much as called for to achieve the full benefits to all parties. Thus, market incentives will lead to underinvestment in transmission capacity.

Comparing Wholesale and Retail Competition

The possibility of underinvestment is more serious in the retail model than in the wholesale model. The smaller number of entities extracting power from the grid at each node in the wholesale model reduces the possibility of free-riding. In the retail model, by contrast, many retail customers and Marketcos could be accessing the grid at each node, and each customer’s contribution to congestion will be small. The smaller shares will disperse the responsibility for relieving congestion and make it more difficult to reach agreement among the parties to invest in expansion. At the same time, the parties may be affected in different degrees by the presence of congestion, and what may appear to be strategic free-riding on the part of a customer could be an honest lack of concern.

More important, the magnitude of the power-purchase transaction generally will be greater in the wholesale model and, because of economies of scale in transmission investments, the tendency toward free-riding will be reduced. As the value of the power transaction grows, the cost of transmission upgrades becomes a smaller part of the overall cost of the transaction, and the parties involved would be less inclined to risk losing the transaction by free-riding. In the retail model, in contrast, Marketcos are expected to be smaller than Discos over the same customer territory, and some individual consumers will make their own deals. Consequently, these parties will not be able to exploit the scale...
economies in transmission to the same degree as in the wholesale model, and the incentive to free ride is stronger.

Even the complete elimination of the tendency to free ride will not eliminate the tendency to underinvest, however, because of the inability of investors to appropriate all of the benefits that accrue to others connected to the grid. To solve the underinvestment problem will require intervention on the part of the ISO to make up any deficiencies. The ISO would have to participate in transmission-planning decisions with interested parties and determine when and by how much the social benefits of investment exceed the private benefits of the interested parties. Because of scale economies and the lumpiness of transmission upgrades, moreover, the ISO may find that small additional investments will achieve large additions to capacity. The additional increments to investment provided by the ISO would be reimbursed by taxing all entities connected to the grid with a higher access charge.

RELATIVE IMPORTANCE OF THE FIVE FACTORS

A research strategy for studying the relative efficiency of the two market structures should attempt to assign priorities to the five factors just discussed. In the absence of empirical information on each of the factors, we rely on the qualitative implications suggested by the foregoing discussion.

(1) Efficiency of the Contract Market. Potential differences in pricing behavior between the two market structures could be the single most important factor affecting relative economic performance of the two market structures. In addition to the potential direct effects of efficient pricing on productive and allocative efficiency, differences in contract market performance could affect the incentives for product diversity, technical change, and capacity additions. Thus, any study of relative market performance should include a component that examines the performance of contracting between buyers and sellers of electricity. Central to the comparison of the two market structures will be the compatibility of regulation of the Discos with the incentive structure necessary to properly represent the interests of consumers.

(2) Product Differentiation. There are clear differences in the incentives for product differentiation in the two market structures. Direct access to consumers, and the presence of entities that unambiguously gain from product differentiation, suggest that the retail competition model will dominate in this regard. This result is also the common expectation based on the experience of other industries (e.g., telecommunications) that have moved in a similar direction. The issue that remains for further research is the potential importance of this factor for market efficiency and economic welfare. There is no way at present of projecting the quantitative magnitude of the relative benefits of product differentiation. However, the clear distinction between the two markets and the common expectation about relative performance make it important to study this issue further.
(3) **Technological change.** Differences in the incentives for technological change in the two market structures arise from direct consumer access and unregulated profits in the same way as the incentives for product differentiation. Indeed, technological change is the long-term, dynamic analog of product differentiation. However, it is even harder to assess the impact of technological change on efficiency and welfare than on product differentiation. Thus, for practical reasons we do not suggest including this factor in the recommended research strategy described below.

(4) **Transaction Costs and Scale Economies in Distribution.** The anticipated differences in transaction costs in the distribution stage of the two market structures is expected to favor the wholesale competition version. However, perhaps the more important issue is the extent of scale economies in distribution. The reason is that scale economies will determine the incentives to merge or disaggregate distribution companies, and thus determine the tendency toward competition on the buying side of the market. The issue is paramount in the retail competition model where competition among the Marketcos is a prerequisite to deregulation of this sector. Thus, we recommend directing further research toward the scale economy issue.

(5) **Generation Reserve Margins.** Any difference in the relative size of generation capacity in the two market structures is largely a matter of regulatory decisions rather than the result of differences in economic incentives created by the two markets. Moreover, the reasons for and implications of a public policy decision on the adequacy of generation capacity would be essentially the same in either market structure. Further research on the possibility of a market failure is of value to determine proper regulatory policy, but not to distinguish the relative efficiency consequences of wholesale and retail competition. Thus, we do not recommend further research on this issue in the present context.

(6) **Transmission Investment.** The problem of underinvestment in transmission capacity is more serious under retail competition than under wholesale competition. However, the solution to the problem is the same in either case: the ISO must step in to ensure an adequate level of investment. Consequently, the research issue of interest is not so much the relative efficiency implications as it is the relative importance of the responsibilities assigned to the ISO. This paper has not addressed questions about how the ISO will be created, operated, managed, and regulated, though these issues will be important to the success of restructuring.

**RECOMMENDED RESEARCH STRATEGY**

Task 1: Assess the Relative Efficiency of the Contract Markets in the Wholesale and Retail Models
Our discussion of the contracting issue emphasizes three broad topics that are related to the performance of contracting in the two market structures. The first topic concerns the specification of the type of regulation of the Discos in the wholesale competition model before drawing any conclusions about the relative performance of contracting in the two models. The second topic concerns the conditions necessary for competition in the market for contracts, such as the number of power sellers and buyers, the size of transaction costs in completing contracts, and the cost of acquiring information about contract options and prices. The third topic concerns the ability of different parties in the market to accept risk, as determined by risk aversion and flexibility to respond to price variation. Each of these topics suggests a separate research task.

The discussion of the subtasks avoids specifying particular approaches or methodologies for undertaking the work. In view of the absence of relevant data, it is unlikely that statistical methods will prove to be very useful. It is expected that microeconomic tools developed in the contract literature will be applied to reach qualitative conclusions. Where the theory is inconclusive and the data are nonexistent, researchers may turn to experimental economic methods.

**Subtask 1.1: Regulation of the Discos.** Assume for the sake of this subtask that the retail competition model is perfectly competitive with respect to the spot and contract markets. The question to be addressed in this subtask is the identification of the type of regulation of the Discos that is necessary and sufficient to encourage the Discos to enter into contracts with the Gencos that represent the interests of consumers in the same way as contracts arranged by Marketcos and consumers in the retail model. What will regulators have to do to ensure that contracts in the wholesale market mimic those in the retail market? Put another way, are there characteristics of the contracts in the retail model that cannot be replicated in the wholesale market by the actions of the regulator?

This project may proceed by first developing the incentives and actions of an unregulated Disco to determine the essential differences in contracting behavior with outcomes of the competitive model. The next step is to impose regulatory rules on the monopolist Discos to alter their behavior in the desired direction. The question is whether it is possible to alter behavior in the desired way, or if the rules will cause other distortions in efficiency.

Price-cap regulation, as noted earlier, is unlikely to achieve the competitive result because consumer interests are too varied to capture with a limited number of price caps. However, the differences that exist among consumers may not be important. For example, in practice transaction costs limit the number of different contracts that are offered in many fields. Where contracts are standardized, it may be inferred that the benefits of flexibility are more than offset by the reduced transaction costs that come with standardization. Thus, efficiency need not increase in direct proportion to the number of different contracts. Consumers may be aggregated into groups with similar interests without loss of efficiency, and these groups may correspond to the structure of price caps imposed by the regulator. This possibility should be taken into consideration in the foregoing analysis.
Subtask 1.2: Competitive Conditions in the Contract Market. This subtask deals with the assumption stated in the foregoing subtask: is it reasonable to assume that the conditions necessary for workable competition in the contract market in the retail model may actually exist? A number of factors will determine the degree of competition in the contract market. These include the number of independent parties seeking to negotiate contracts, access to and cost of obtaining information about market conditions, the magnitude of transaction costs, and others. This subtask should use the literature on contract theory to analyze the separate elements of a competitive market as they relate to the retail competition scenario. If possible, empirical applications of this literature should be explored as well to draw inferences that are relevant to electricity markets. (Assume for the purposes of this subtask that marginal costs of marketing power are increasing so that the tendency to consolidate Marketcos is avoided. This issue is explored in Task 3.)

Subtask 1.3: Ability to Absorb Risk. Parties that are best able to absorb risk are those that are least risk averse and have the most flexibility to adjust to price variation. This subtask will identify factors that determine risk aversion and flexibility and will associate those factors with the parties in the two market structures to determine how price risk in contracting will be best allocated. Using the work completed under subtask 1.1, this task will assume a form of regulation that comes closest to encouraging Discos to maximize the interests of customers.

A related issue is whether the contract market should be regulated. The conventional insurance industry is regulated in a number of ways, including, in particular, solvency requirements. Is there a need for similar regulation in the electricity market? For example, are the contract values large enough that default will destabilize the market and harm consumers? Is it reasonable to expect the market to develop its own safeguards, or will they have to be imposed by government mandate?

Task 2: Assess the Relative Benefits of Product Differentiation in the Wholesale and Retail Models

The discussion above suggests that retail competition will lead to greater product differentiation than would occur under wholesale competition. This conclusion results from three factors. First, under retail competition, consumers are better able to express their willingness to pay for differentiated electricity services because they participate directly in both spot and contract markets for electricity. Second, the Marketcos have a profit incentive to provide consumers with either power or substitute (and complementary) products, while the Discos when regulated under price-cap regulation only have an incentive to provide power (and complementary products). Third, competition among Marketcos under retail competition will provide the stimulus on the supply side to create new products in order to attract consumers.

The size of the efficiency benefits resulting from this greater product and service differentiation under retail competition depends on the extent of the differentiation and the value to society of having a broader array of services available. We suggest two approaches to investigate this issue: the development of a conceptual model of the value
of product differentiation, and a review of previous studies of the effects of deregulation on product differentiation in the telecommunications and other recently deregulated industries. Each effort constitutes a separate subtask as described below.

**Subtask 2.1: Development of a Conceptual Model.** In this subtask the focus is on product differentiation that would be expected to arise from competition among Marketcos under retail competition. A theoretical economic model of markets for differentiated products may be developed to describe the extent of product or service differentiation expected to arise as a result of increased competition in electricity markets, and to identify factors such as product-specific fixed costs and the nature of strategic behavior among firms that might affect the extent of differentiation. The model could then be used to identify the determinants of the value of greater product and service differentiation to society. As a part of the model development exercise, it would be useful to review the economics literature on product differentiation to identify relevant findings from previous research and to inform the development of the electricity model.

This exercise should reveal a set of necessary and sufficient conditions for increased product differentiation under retail competition to be welfare-enhancing. It may be expected that these conditions will depend on factors such as the costs to producers of expanding the range of service offerings, the intensity of consumer preferences, consumer demand elasticities, and other determinants of demand. For example, earlier research suggests that free entry could lead to too much or too little product differentiation relative to the social optimum, depending largely on the cost or disutility to consumers of consuming a product that differs substantially from the most preferred product.\(^{39}\)

After the model is developed, simulation techniques may be used to identify ranges of empirical values for the relevant parameters listed above, and to determine the estimated magnitude of efficiency-enhancing product differentiation.

**Subtask 2.2: Analysis of Other Industries.** The implications of deregulation for product and service differentiation in telecommunication markets and other recently deregulated markets such as natural gas may provide insights of relevance to the electricity industry. The telecommunications industry has undergone several regulatory changes that have led to increased competition and an associated proliferation of products and services. From the FCC decision in the early 1970s that allowed the use of customer premise equipment not produced by Western Electric to the 1982 consent decree that lead to the break-up of AT&T and widespread competition in long-distance, the telecommunications industry has seen many regulatory developments that have been accompanied by increased product and service differentiation. Concurrent with these regulatory changes has been the rapid development and adoption of new technologies that have also contributed to the proliferation of new services.

This task should also incorporate a review of existing studies of service differentiation in the telecommunications and natural gas industries that can be attributed

---

to regulatory changes, and the resulting value of product differentiation to society. The
implications of developments in these other industries should be extended to the electric
utility industry for their relevance to the two structural models.

Task 3: Assess the Magnitude of Scale Economies in Distribution in the
Wholesale and Retail Models

We have compared the implications of the different market structures for
transaction costs and for scale economies in the distribution of electricity. With regard to
transaction costs, our discussion suggests that any higher transaction costs resulting from
a higher volume of transactions under the retail competition model will be at least partially
offset by the lower transaction costs due to a narrower scope for utility regulation under
the retail competition model. Marketcos will also serve to reduce the size of the
transaction costs of purchasing electricity for consumers and of selling electricity for
Gencos in the retail model. These observations suggest that it is impossible to predict the
implications of the two different market structures for net transaction costs without more
information about the magnitude of each type of transaction cost identified.

Unfortunately, this type of information does not currently exist, in large part
because neither of these scenarios has come to pass. However, much of this information
would be difficult to obtain even after the switch to a new market structure. For example,
in gas markets where wholesale restructuring has already taken place, it might be possible
to assess the implications for regulatory costs of the change in regulation. However, it
would be much more difficult to estimate the total transaction costs to local distribution
companies associated with gas contracting and spot transactions.

Given these difficulties, we conclude that an attempt to measure transaction costs
is impractical. Instead, we focus on the implications of the two models for scale
economies in distribution and retail sales of electricity. If the importance of scale
economies can be assessed, we can then answer the question, "How large would
transaction costs have to be to offset the efficiency effects of scale economies?"

To measure scale economies in distribution, we suggest using data compiled by
John Kwoka with assistance from the American Public Power Association for 230
municipal utilities that are engaged exclusively in distribution and retail sales of electricity.
With the use of this data it may be possible to estimate a cost function for local
distribution. The estimated cost function may then be used to draw inferences about scale
economies in distribution and to identify the efficient size of a local Disco. In addition, it
may be possible to use the data to separate scale effects in the retail sales function from
scale economies in the wire-service function. This exercise will help to determine if
Linecos are likely to be larger or smaller than Discos, and if the Marketco business is
likely to be highly competitive or more concentrated in local electricity markets. As noted
already, the conclusion will be important in assessing the likely extent of competition in
the retail competition model and the amount of regulation that will be imposed on the
distribution stage of the industry.

39
REFERENCES


### Abstract (Maximum 200 words)

A central issue in the debate over restructuring the electric power industry is the extent to which the market should be open to competition. One aspect of this debate is whether competition ought to be restricted to the wholesale power market or be extended to final retail consumers. This report begins to explore the potential differences in economic efficiency between wholesale and retail competition in the electric power industry. The two market-structure scenarios are defined and the factors responsible for differences in efficiency are described. The report also contains an assessment of the relative importance of the factors and recommendations for pursuing further research.