CHAPTER 4

Revenue Caps vs. Price Caps: Implications for DSM
(From LBL Report #37577)

Steven Stoft

11 November 1995

4.1 Overview

This chapter analyzes the incentive effects of revenue-cap PBR mechanisms. The importance of this task derives both from the common occurrence of revenue caps among the newly implemented electric industry PBR mechanisms, and from the fact that the revenue cap mechanism has been explicitly proposed as a replacement for the theoretically favored price cap mechanism. This proposed replacement is based on the perception that price caps provide a strong disincentive to utility investment in energy efficiency\(^1\) and has been made explicitly by Moskovitz (1992), and by Marcus and Grueneich (1994, pp.41-42), and has been hinted at by Hamrin et al. (1994, p. 150). Revenue caps are currently in use at Con Ed and SDG&E, and have been proposed for PG&E and SCE.

It is also important to note what this section is not attempting to accomplish. We will be interested on the incentive implications for energy efficiency of price and revenue caps, but we will not attempt answer the question of whether energy efficiency is or should be an agreed objective of state utility policy. We will also not consider the effect of other possible incentive mechanisms on DSM. This chapter focuses solely on the incentives of revenue caps and how they compare with price caps. We will also not be concerned to predict outcomes of the discovered incentives. When we find that a utility has an incentive of \(V\$\) per unit of increase in \(X\), we have solved half the problem of determining how much \(X\) will be increased. To finish the problem one must discover and utilize the cost function for increasing \(X\); this is an entirely separate subject which we will not addressed. It should also be noted that this chapter is not addressed to academic economists, though we believe that some of the ideas in this chapter are new and of interest to that community.

\(^1\) The California PUC, Division of Strategic Planning states: “The price cap model ... provides a strong disincentive to invest in energy efficiency.” (p. 175).
In this chapter we confirm that price caps, coupled with the current pricing structure, create a disincentive to effective energy efficiency programs, and that revenue caps do reverse this disincentive. We also confirm that revenue caps produce exactly the same cost minimizing incentives as do price caps, and that utilities are no more sensitive to growth in customer base with a revenue-per-customer cap than with a price cap.

A recent and potentially devastating critique of revenue caps has been put forward by economists Crew and Kleindorfer (Crew and Kleindorfer 1995). Their critique purports to show that a revenue-capped firm will always set price above the monopoly level. While we acknowledge the possibility of this effect, especially if a firm can engage in successful DSM,\(^2\) we show several ways in which this pricing effect can be inhibited. First, under some circumstances, the Crew-Kleindorfer effect can be inhibited by a price cap used in conjunction with a revenue cap without reversing the revenue cap’s DSM incentives. Second inelastic short-run demand may inhibit the effect. Third, we propose a new hybrid price-revenue cap as the safest way to avoid the Crew-Kleindorfer effect and related price effects.

Lastly we discuss the subtle problem of relative prices. It is well known that price caps, by allowing flexibility in the price of one product or of one class relative to another, will induce the utility to approximate Ramsey pricing.\(^3\) Revenue caps, on the other hand, are shown to motivate large relative price changes in the opposite direction to those of Ramsey pricing. This effect of revenue caps could cause even larger pricing inefficiencies than the Crew-Kleindorfer effect, and should be inhibited either by explicit regulation of relative prices or by the adoption of a hybrid cap that leans towards the price-cap end of the spectrum.

In summary, our analysis of pure revenue caps reveals a number of potential problems:

(1) Incentives to set relative prices inefficiently.

(2) The possibility that a small reduction in the revenue cap will produce a large and unpredictable reduction in price (an effect related to the Crew-Kleindorfer effect).

(3) An incentive to reduce sales regardless of the social benefit.

For those who are concerned with the sales incentives of price caps we recommend, in place of the pure revenue cap, a hybrid price-revenue cap. We also recommend that this be employed in a revenue-per-customer form. Such a cap would only need to replace the energy part of a price cap, and could take a form as simple as the following.

---

\(^2\) For convenience we will often refer to energy efficiency programs simply as DSM (demand-side management) even though the meaning of this term is significantly broader.

\(^3\) This was shown by Vogelsang and Finsinger (1979). Ramsey pricing marks up prices in inverse proportion to a product’s demand elasticity. Thus the price of access, being inelastically demanded, would be marked up more than average. Theoretically, this scheme maximizes consumer welfare.
The existence of fixed costs, causes average cost to decline as production increases. For this reason most firms constantly seek to increase sales. Typically this is done by advertising for new customers and for greater use per customer. Although distribution companies cannot control their customer base, they can, and have been known to, seek to influence their per-customer demand.

Of course the costliness of this encouragement may make it uneconomical, but our first task is to measure the strength of the encouragement.

\[ P_E < \bar{P}_E - \frac{R_E}{q_0 \cdot N} \]  

(4-1)

\( P_E \) is the price of energy, \( \bar{P}_E \) is like a price cap only high enough to compensate for the following revenue term. The subtracted term measures revenue from energy charges divided by initial energy per customer times the number of customers.

This hybrid cap will (1) greatly reduce the incentive to distort relative prices, (2) prevent the uncertain price response caused by a pure revenue cap, and (3) remove the anti-DSM bias of a price cap without causing an incentive to reduce sales without regard to social benefit. The most significant question left unanswered by this chapter is the question of exactly what relative pricing incentive will remain under such a hybrid cap.

4.2 Background

Because price-caps encourage the minimization of average cost, they may also encourage the maximization of sales in order to dilute fixed costs.\(^4\) Unfortunately this behavior often runs counter to the encouragement of energy efficiency which regulators have often promoted. As a remedy, revenue caps have been proposed as a replacement for price caps.

Revenue-cap regulation, in its simplest form, simply limits to a predetermined level the amount of revenue per year that a firm can collect from its customer base. As a consequence the utility has a clear incentive to encourage minimal total demand, and thus minimal demand per customer.\(^5\) One way to do this is to encourage the efficient use of power, but this is not the only way, and there are many complications. Nonetheless, with proper adjustment, and in the right circumstances, a revenue cap might motivate both supply-side cost minimization and demand-side efficiency maximization without imposing too much risk or inducing perverse behavior on the part of the utility.

Revenue-cap regulation is not without precedent. Just as standard ROR regulation is actually a type of price-cap regulation, so ROR with an electric revenue adjustment mechanism (ERAM) is actually a type of revenue-cap regulation. With an ERAM in place, a utility is guaranteed a fixed revenue in place of a fixed price. (Revenue is usually not completely fixed but is made temporarily independent of costs.) Another approach that has been much

\(^4\) The existence of fixed costs, causes average cost to decline as production increases. For this reason most firms constantly seek to increase sales. Typically this is done by advertising for new customers and for greater use per customer. Although distribution companies cannot control their customer base, they can, and have been know to, seek to influence their per-customer demand.

\(^5\) Of course the costliness of this encouragement may make it uneconomical, but our first task is to measure the strength of the encouragement.
Although economists do at times consider the possibility of the firm shifting its own demand curve (Lewis and Sappington, 1992) (Laffont and Tirole, Chapter 4, 199x), there has been little if any academic analysis of price caps from this point of view.

The central assumption behind the advocacy of revenue-cap regulation is that the utility can affect the demand-curve for energy. This assumption is not usually made in the economic analysis of price-cap regulation. This may be why the standard economics literature ignores revenue caps.

Those concerned with energy efficiency point to discrepancies between marginal costs and prices, and to the possibility that utilities can significantly influence demand through DSM programs. If these assumptions are correct, then they need to be accounted for in the analysis of regulatory incentives. Consequently, our analysis first attempts to construct a framework for analyzing the relationships among price structures, cost functions and regulatory mechanisms.

### 4.3 Modeling Industry Costs and Prices and Regulatory Mechanisms

To analyze price and revenue caps, we need to specify the structure of a utility’s costs and prices. Clearly both of these structures are extremely complex, so we use a simplified model that captures the most essential features (more than could be captured by a nonmathematical treatment of this subject). With the necessary simplifying assumptions, we write utility costs as:

\[ C = a + bN + cE + dL \]  

where,  
- \( N \) = Number of customers  
- \( E \) = Total energy  
- \( L \) = Peak load

*Although economists do at times consider the possibility of the firm shifting its own demand curve (Lewis and Sappington, 1992) (Laffont and Tirole, Chapter 4, 199x), there has been little if any academic analysis of price caps from this point of view.*

*One clarification of revenue caps may be needed. Price-cap literature often addresses the problem of capping prices at a multi product firm. An aggregate cap is usually suggested rather than a cap for each individual price. This allows a firm relative-price flexibility, which can improve efficiency. Such an aggregate cap involves a quantity-weighted average of prices that looks similar to a revenue calculation. The difference is that it uses past quantities instead of present quantities as weights. This section is not concerned with average-price caps.*
Later, when it is needed, we will introduce the concept that both $E$ and $L$ depend on $N$. This in no way contradicts equation (4-1), but it does require us to remember that the effect shown by equation (4-1) of changing $N$ is the effect with $E$ and $L$ held constant. This affect has economic significance but it is not the one we are ultimately most interested in. Also note that since our results are differential in nature, nothing would be gained by replacing a local linear approximation by the true cost function (i.e. nothing we say depends on second derivatives.)

The price structure, which is modeled by specifying the revenue equation, has a very similar form:

$$R = P_N N + P_E E + P_L L$$

\[(4-3)\]

where, 
\begin{align*} 
R &= \text{Revenue} \\
P_N &= \text{Access charge} \\
P_E &= \text{Energy charge} \\
P_L &= \text{Demand charge}
\end{align*}

Prices essentially consist of a hookup or access charge, which will be denoted by $P_N$, by an energy charge, $P_E$, and by a demand charge, $P_L$. Distinguishing these components of price is crucial, as is recognized by NMPC in the following statement.

"...the internal price indexes have separate categories for access, demand, and volumetric charges. Accordingly the Company can, via rate redesign, change the relative importance of these charges and still hold its internal indexes constant.” (Lowry 1994, p.7)

Finally, we specify the structure of the two alternative incentive mechanisms: price cap, and revenue cap.

\begin{align*}
\text{Price Cap:} & \quad P_N < \bar{P}_N, \quad P_E < \bar{P}_E \quad \text{and} \quad P_L < \bar{P}_L \\
\text{Revenue Cap:} & \quad R < \bar{R}
\end{align*}

\[(4-4)\]

---

8 This linear approximation is fairly good because the inputs required to support customers (poles and billing), the inputs associated strictly with energy (fuels), and the inputs associated strictly with peak load (wires and generators) do not interact strongly. I.e. if you double peak load and leave energy and customers fixed, fuel use is not changed dramatically, and billing costs don’t change at all.

9 For the utility as a whole, $P_L$ is far from constant, but within customer classes this model is reasonable. Still, it ignores the fact that the peak loads of individual customers are not coincident with the utility’s peak. This could be largely corrected with a proportionality factor. In spite of these deficiencies, the model serves its purpose.
CHAPTER 4

Notice that this is an inflexible price cap, which is simpler than the cap on a Laspeyfer’s price index that is often used in practice. The inflexible mechanism, which caps $P_N$, $P_E$, and $P_L$ separately, differs from its more flexible cousins only in that flexibility allows the utility to choose the relative values of its prices. The implications of this choice will be examined in section 4.9.

4.4 Incentives of Price and Revenue Caps

Incentive mechanisms are designed to induce a firm to optimize its behavior. But the ways in which this optimization is done may be difficult for a regulator to observe and control. Typically, a firm will minimize costs, which in our model corresponds to minimizing the cost parameters $a$, $b$, $c$, and $d$ of equation 4-1. In the current context, we are also interested in the utility’s incentive to modify several variables that are not normally considered to be under its control, namely, $N$, $E$ and $L$. In particular, we are looking for incentives for the utility to reduce $E/N$, energy use per customer, which we will call $q$.

Finding these PBR incentives is now simply a matter of differentiating profit with respect to the utility’s cost parameters and with respect to the customer variables over which it may have control. Profit is given by $\pi = R - C$, and we assume that the regulatory caps are binding constraints. This allows us to compute profit.

\[
\begin{align*}
\text{Price Cap: } \pi &= -a + (P_N - b)N + (P_E - c)E + (P_L - d)L \approx 0 \\
\text{Revenue Cap: } \pi &= \bar{R} - (a + bN + cE + dL) \approx 0
\end{align*}
\]

First consider the cost parameters $a$, $b$, $c$, and $d$. Under either a price or revenue cap, a utility will have a clear and strong incentive to minimize all cost parameters, as can be seen in Equation 4-5 where all of them make negative contributions to profit. These incentives are exactly the same for either mechanism.

Next, consider the customer variables under a revenue cap. The profit equations clearly show that a utility has an incentive to reduce $E$, and $L$, and that these incentives are equal to the marginal cost of each of these variables. ($N$ will be considered shortly.) Thus the incentives are quite substantial. Because a reduction in $q$ (energy use per customer) will reduce $E$, the utility will have a strong incentive to reduce $q$, which is our goal. (Note that just because a firm has an incentive to reduce $q$, it may not choose to do so if it finds it too costly; i.e. cost provides a conflicting incentive. At this time we do not wish to analyze the net incentive.)

\[\text{Footnote: The most common form of price cap, used by NMPC among others, is a cap on a Laspeyfer’s index of price. This index is formed by taking a weighted average of prices with the weightings based on past quantities.}\]

4-6
Finally consider the customer variables \(E\) and \(L\) under a price cap. The incentives of the price cap are more ambiguous than those of the revenue cap, depending on the relative values of the various price and cost components, so we now consider how \(P_E\) compares to \(c\) and how \(P_L\) compares to \(d\). Because there is no transfer payment to the utility corresponding to the fixed cost, \(a\), at least one of the three prices must be higher than its corresponding costs. Fixed cost, \(a\), will be small in a utility that has been run as if it were in a competitive market, but will be large for utilities that have what may become large “stranded assets.”\(^{11}\) We will consider an intermediate case. Incentive problems are typically caused by misalignment between various marginal costs and prices. Specifically, \(P_L\) is typically set to zero for residential customers and the cost of capacity is shifted to the price of energy, inducing \(P_E > c\). This gives the utility a strong incentive to minimize \(L\) and a strong incentive to maximize \(E\) (to the extent that this can be done without increasing \(L\) proportionally). A second pricing bias that is perceived to be widespread is the underpricing of access.\(^{12}\) This shifts costs onto the energy price component, further increasing the incentive to maximize \(E\). However, the incentive to minimize \(N\) is not as great as it would at first appear, because \(N\) plays a role in determining both \(E\) and \(L\).\(^{13}\) This effect is best seen by rewriting the profit equations as follows:

\[
\begin{align*}
\text{Price Cap: } & \quad \pi = -a + [(P_N - b) + (P_E - c)q + (P_L - d)k]N = 0 \\
\text{Revenue Cap: } & \quad \pi = \bar{R} - a - (b + cq + dk)N = 0
\end{align*}
\]  

(4-6)

As before, \(q\) is energy use per customer, and \(k\) is now peak load per customer, so \(qN = E\), and \(kN = L\).

We are first interested in the incentive to maximize \(N\), which can now be found by differentiating either equation in (4-6) by \(N\), and then using the approximation that economic profit is zero.\(^{14}\) For a price cap \(\frac{d\pi}{dN} = [(P_N \ldots )k]\), the value of which is easily solved for from the approximating equation \(\pi = 0\). One step of algebra shows that \([(P_N \ldots )k]\) = \(a/N\), so \(d\pi/dN\) is approximately \(a/N\). Similarly we find that for a revenue cap \(d\pi/dN\) is approximately

\(^{11}\) For a single generation facility the fixed cost may be significant, but a utility does not generate at a single plant. For a utility, capacity is expanded by adding new units. As long as the cost of new units is similar to the cost of existing units, the system fixed cost is small. If some existing units are on the books at values well above their current market price, then we have “stranded assets,” and \(a\) will be substantial.

\(^{12}\) It is not uncommon for customer access charges, especially in the residential classes, to be priced below marginal cost.

\(^{13}\) Note that all we have done is factor the \(N\) out of \(E\) after we describe \(E\) as \(qN\), and the \(N\) out of \(L\) after we describe \(L\) as \(kN\). This does not contradict our original formulation it only reveals more of the structure hidden inside equation 4-1.

\(^{14}\) Note that we do not hold profit equal to zero or even constant when differentiating.
\[(a - \tilde{R})/N.\] Our best estimate is that \(a\) is much smaller than \(R\), so the incentive is to decrease \(N\) under a revenue cap.

In our first round of analysis of price and revenue caps, we have learned that the two mechanisms work equally well to induce reductions in cost parameters. We have also confirmed that pure price caps discourage DSM while pure revenue caps encourage DSM. But we have discovered that revenue caps make profit very sensitive to fluctuations in the number of customers, which can cause excessive uncertainty in profit and undesirable incentives to minimize \(N\). Section 4.5 addresses this problem of the standard revenue cap and then formally analyzes the power of the various incentives. Sections 4.6–4.9 considers the price-setting behavior of firms under these mechanisms, a topic on which we have not yet touched.

### 4.5 The Revenue-per-Customer Cap

The powerful incentive to minimize \(N\) under a revenue cap has two associated problems. If the utility can influence its number of customers, perhaps by deliberately losing customers to alternative fuels, self generation, or retail wheeling, this incentive will induce perverse behavior. The second problem is more common, an excessive unpredictability in profit due to the dependence of profit on \(N\). Fortunately, there is an easy solution to both of these problems: the revenue-per-customer cap mechanism, which is defined as \(R/N < \tilde{R}_N\)

\[
\text{Revenue-per-Customer Cap: } R < \tilde{R}_N \cdot N
\]

Profit: \(\pi = -a + (\tilde{R}_N - b - cq - dk) N \approx 0\)

We can see in the equation that there is now an incentive, of approximately \(a/N\), to maximize \(N\), exactly as with a price cap. (For details, see the previous calculation of \(d\pi/dN\).)

At this point it is useful to summarize the similarities and differences among the three incentive schemes just described. To do this we need to make a few assumptions about the relative magnitudes of marginal costs and prices, and about the level of profit. These assumptions are made to illustrate the above theoretical points but are not intended to be accurate estimates for any particular utility. We will use the assumptions displayed in Table 4-1:
Notice that costs add up to 99% of revenue, so that profit is 1%. This is not a low profit level both because it is expressed as a percent of revenue instead of, as is typical, a percent of assets, and because it is economic profit, which is measured after allowing for the cost of capital. As mentioned previously, $a$ can be either nearly zero or, if the utility would end up with significant stranded assets in a competitive world, it can be quite large. For our example, we will take $a$ to be 10%, which we consider to be an intermediate value. This value mainly affects the incentive on N, which can easily be recalculated by the reader to suit any other value of $a$. The other values are much more difficult to estimate accurately, but their relative magnitudes appear realistic. Note that capacity costs are set at 24%, not because they are known more accurately but to leave room for 1% profit.

Prior to comparing the three PBR approaches, recall our discussion in Section 1 of the power of an incentive mechanism. Incentive power can be measured by the fraction of each dollar of cost decrease that is ultimately kept by a firm. Thus, if the utility sells another 10kWh, thereby increasing its costs by one dollar, and if this dollar adds fifty cents to the utility’s profits, then the power of the incentive to sell kWhs is said to be $\frac{1}{2}$ or 50%. Mathematically, this is expressed as $\frac{d\pi/dX}{dC/dX}$, where $X$ is the quantity affected by the incentive. It can be noted in advance that with $C = a + bN + cqN + dkN$, $dC/dN = (C-a)/N$, $dC/dq = cN$, $dC/dk = dN$. All of the numerical results in table 4-2 can be computed from the values in table 4-1.

### Table 4-1. Assumptions for Computing Incentive Power

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Magnitude as a % of Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Revenue</td>
<td>$P_N$</td>
</tr>
<tr>
<td>Energy Charges</td>
<td>$P_E$</td>
</tr>
<tr>
<td>Demand Charges</td>
<td>$P_L$</td>
</tr>
<tr>
<td>Fixed Cost</td>
<td>$a$</td>
</tr>
<tr>
<td>Customer Costs</td>
<td>$bN$</td>
</tr>
<tr>
<td>Energy Costs</td>
<td>$cE$</td>
</tr>
<tr>
<td>Capacity Costs</td>
<td>$dL$</td>
</tr>
<tr>
<td>Profit</td>
<td>$\pi$</td>
</tr>
</tbody>
</table>
Actually, the comparison is not so simple. In order to simplify presentation, we have ignored demand elasticity. This is not a problem under price caps, where price is essentially fixed, but it is under revenue caps. For low short-run elasticities, the needed correction is not large, but when near or to the left of the peak of the revenue function shown in Figure 4.1, the effect is dramatic. An exact computation of incentives is shown in Appendix C where we compute the energy-efficiency incentive of a hybrid cap.

In Table 4-2, notice that the three PBR mechanisms behave identically except in the areas noted by the shaded cells. This means all three have the same cost-minimization incentives, and all three treat peak load per customer, \( k \), the same. As noted, the standard revenue cap produces a strong and anomalous negative incentive on \( N \), but the per-customer revenue cap exactly realigns this incentive with that of a price cap. Thus the only difference between price caps and revenue-per-customer caps is in their effect on \( q \), energy use per customer.

The two revenue caps differ dramatically from price caps in their effect on per-customer energy use, \( q \); under the current price structure, they treat DSM far more favorably than do price caps. The particular incentive levels for \( q \) depend on the values of the marginal cost of energy, \( c \), and the energy charge \( P \). However the difference between the price-cap and revenue-cap incentives for DSM is exactly equal to total energy charge divided by total energy cost. In our example, \( P_E \) is 90% of revenue and energy cost is 45% of revenue, so the power of revenue caps to encourage DSM is 200% greater than the power of a price. In order for the two mechanisms to treat DSM the same, \( P_E \) would have to be zero, a circumstance that is neither likely nor desirable.

---

15 Actually, the comparison is not so simple. In order to simplify presentation, we have ignored demand elasticity. This is not a problem under price caps, where price is essentially fixed, but it is under revenue caps. For low short-run elasticities, the needed correction is not large, but when near or to the left of the peak of the revenue function shown in Figure 4.1, the effect is dramatic. An exact computation of incentives is shown in Appendix C where we compute the energy-efficiency incentive of a hybrid cap.
4.6 Crew and Kleindorfer’s Critique of Revenue Caps

So far we have considered the utility’s behavior with respect to costs and customers. Now we must consider pricing. In this regard, a potentially devastating critique of revenue caps was put forward recently by Crew and Kleindorfer (1995) and Kenneth Costello has made a related criticism. Crew and Kleindorfer prove that a revenue cap, if implemented without any other regulatory constraints, will induce a firm to set its price higher than if the firm were a pure monopoly. Because this would be disastrous from a public acceptance viewpoint, and highly inefficient from an economist’s viewpoint, it is necessary to modify the regulatory mechanism in order to avoid this outcome. We will argue that some modifications that accomplish this goal are already in use by PUCs although they may not have been introduced with this intention.

To find a corrective mechanism we must first understand the critique. Crew and Kleindorfer’s argument notes that an unconstrained monopolist will choose a profit maximizing price, $P^*$, which will induce a monopoly level of demand, $Q^*$, and a monopoly revenue, $R^* = P^*Q^*$. If the regulator sets the revenue cap, $\hat{R}$ higher than $R^*$, the monopolist will simply ignore it because a lower revenue maximizes profits. To have any impact at all, the regulator must set $\hat{R} < R^*$.

Assuming $\hat{R}$ is less than $R^*$, the firm will be forced to raise or lower $P$ in order to reduce revenue and satisfy the regulator’s constraint. Generally either strategy is possible: at an extremely high price, sales will fall to such an extent that revenues will decline while as price approaches zero, revenues also decline. So let us consider a high price and a low price, both of which exactly satisfy the constraint on $R$. Because profit is simply revenue minus cost, and revenue is $\hat{R}$ in both cases, the only difference is cost. More electricity will be sold at the lower price, so the cost of generation will be higher at the lower price. Consequently the higher price will be chosen. The higher price is so high that it reduces revenues to $\tilde{R}$, which is lower than $R^*$; therefore, the “high” price must be even higher than the monopolist’s price.

The following figure illustrates this argument. It is important to note that the argument depends on essentially only three assumptions: (1) that at a sufficiently high price, revenue will decline to $\tilde{R}$, (2) that the total cost of generation increases with the quantity generated, and (3) that the dynamics of reaching the equilibrium don’t matter. The first of these could be false though much empirical work points toward a long-run elasticity greater than one at

---

16 In a talk, “Why ‘Yes’ to Price Caps and ‘No’ to Revenue Caps”, given in San Francisco on July 25, 1995, Costello argues that “if price falls and the price elasticity of demand exceeds one, total bills increase.” Reversing this logic we have that the utility can reduce revenue by raising price.

17 If this were not true, the monopolist would set an infinitely high price. In fact the monopolistic price is always in a region where revenues decline with price increases.
current price levels, and elasticity is likely to increase at higher levels.\textsuperscript{18} It is nearly impossible to imagine that the second assumption is false. As will be demonstrated in Section 4.8, the Achilles’ heel of the analysis is the third point. Short-run demand elasticity coupled with regulatory intolerance of long “temporary” violations of the cap will prevent the Crew-Kleindorfer effect if long-run demand elasticity is greater than 1. But as Section 4.8 also demonstrates, related problems remain.

Figure 4-1 depicts both the Crew-Kleindorfer dilemma and a mechanism for avoiding that dilemma. Because a firm prefers a high price to a low price at the same revenue, it is easily seen that when $\bar{R}$ is imposed as a cap, the firm will choose the high-price method of meeting that constraint.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{The Effect of Revenue Caps on Price}
\end{figure}

Fortunately, a simple method exists for prohibiting the economically rational response to a revenue cap, and that is a price cap. It is a bit surprising that both caps can be binding simultaneously; one would think that if the price cap were binding, the revenue cap would not be. This argument would imply that if the price cap had any effect it would simply replace the revenue cap and thus eliminate all of its incentive properties. But Figure 4-1 shows this is not true.

A price cap, placed as shown at $\bar{P}$, will be binding in a global sense while the revenue cap, $\bar{R}$, will continue to bind at the margin. We say that the price cap is binding globally but not marginally because it does not prevent a small (marginal) change in price, but it does prevent

\textsuperscript{18} If this is false then a true electric power monopoly would raise price without limit. This would be a surprising phenomenon.
the desired large change in price from the “lower” price to the “higher” price. Thus the price cap will prevent a large discrete change in price to the high-price profit maximum while the revenue cap will prevent marginal changes in price. The result is a revenue cap that maintains the incentive properties desired by DSM advocates but without any hint of the pricing problem identified by Crew and Kleindorfer.

This technique works only if a firm’s initial price level places it to the left of the revenue maximum as shown above. If the firm starts to the right, on the downward sloping part of the revenue curve, then the firm can only lower revenue by (1) gradually raising price, or (2) lowering price in a large discrete jump. (On the left side of the revenue hill, a small reduction in price leads to an increase in revenue.) A price cap that prevents a firm from satisfying the revenue cap by increasing price will necessarily force a large discrete price reduction; i.e., the firm will have to jump to the left side of the revenue “hill.” Because of this, even the mildest revenue cap (measured by the size of the required revenue reduction) will have a dramatic impact on price and profits. Such a form of regulation is unpredictable in its consequences and thus quite risky.

To recapitulate, if a firm starts on the left side facing a price cap that prevents it from moving to the right of the revenue maximum, then the firm can only reduce revenue by lowering prices and moving to the left, and it can do this with a small change in price. But if a firm starts on the right, a binding price cap will prevent it from lowering revenue by moving to the right, and will force it to move all the way to the left side of the revenue hill by making a discrete and possibly large downward shift in price.

Thus, for a firm on the right, a revenue cap that forces even a small reduction in revenue, can force a large price reduction. If the firm successfully accommodates this price reduction it will be paid for either out of excess profits or by cost reductions. But the if it cannot find large enough cost reductions this small revenue reduction (and large price reduction) can put the firm out of business or force the regulator to back down.

The auxiliary price cap has unpredictable results when used on a firm to the right of the revenue maximum. Also, as will be demonstrated in Section 4.8, Crew-Kleindorfer style difficulties persist even when short-run dynamics are accounted for. Consequently we are still in need of a safe and predictable mechanism that eliminates the problems of a revenue cap while maintaining its desired incentive properties. The hybrid price-revenue cap presented in the next section satisfies both of these objectives.
4.7 Designing a Hybrid Price-Revenue Cap

We have shown that a price cap discourages DSM while a revenue cap can induce perverse pricing behavior. Fortunately, a hybrid price-revenue cap can be designed to avoid any possibility of the Crew-Kleindorfer dilemma, and to avoid the energy efficiency disincentives of a price cap. As we will see in Section 4.9, a hybrid cap also helps curb the strong distorting effects on relative prices that revenue caps encourage.

A hybrid price-revenue cap is represented by a diagonal line in the revenue-price diagram and is represented algebraically as follows:

\[
\text{In Revenue–Cap Form: } R \quad \bar{R} - b \cdot P \\
\text{In Price–Cap Form: } P \quad \bar{P} - c \cdot R
\]  

Note that the same hybrid cap can be represented in two distinct but equivalent forms: as a variable revenue cap or as a variable price cap. In the revenue-cap form, the cap decreases as the utility increases its price. In the price-cap form, the cap decreases as the utility increases its revenue. Note that in a hybrid cap, \(\bar{R}\) and \(\bar{P}\) are fixed, but they are no longer the limits on \(R\) and \(P\); the limits are now lower and variable. The entire right-hand expression is the cap, and in both cases this is significantly lower than the barred variable. Also note that the utility controls both \(R\) and \(P\), but that it does not control them independently.

Figure 2 Hybrid Caps in the Inelastic Region
We now present a graphical representation of hybrid caps. It turns out that the behavior of these caps is quite different in the upward sloping (inelastic) region of the revenue curve than in the downward sloping (elastic) region. In the latter region it is necessary determine the slope of the hybrid cap in accordance with the maximum elasticity of electricity demand.19

Figure 4-2 shows three examples of hybrid caps in the inelastic region. All three are set to behave exactly like the price cap that is shown with them. That is, they are designed so that they will induce the same price behavior as induced by the price cap. A revenue cap is also shown, and is drawn so that it will produce the same outcome, provided the Crew-Kleindorfer effect is prevented in one of the ways discussed above.

Figure 3Hybrid Caps in the Elastic Region

---

19 To review the connection between elasticities and the shape of the revenue function, consider first an elasticity of one, which by definition is the dividing line between elastic and inelastic. A demand elasticity of one means that when price increases by 1%, quantity demanded will decrease by 1%. Because revenue is the product of price and quantity, it remains unaffected by this particular change in price and quantity. This means that at the divide between elastic and inelastic, the revenue curve is perfectly flat. This indicates the top of the revenue function plotted in Figure 4-1, 2 and 3. If demand is elastic, then a 1% change in price produces a larger decrease in quantity demanded; thus the quantity effect dominates, and revenue declines when price increases. The downward-sloping region on the right-hand side of figures is the elastic part of the demand function. Similarly, the upward-sloping region is the inelastic part of the demand function.
The elastic case, shown in Figure 4-3, presents a different story. Again, the firm begins at point A, but this time a price cap and a revenue cap will have opposite effects. A revenue cap will lower revenue and raise price (point C), while a price cap will lower price and raise revenue (point B). If we use the same revenue cap as in the inelastic example, we generate the unfortunate outcome predicted by Crew and Kleindorfer (and suggested by Costello): the firm raises prices in order to lower revenues. The hybrid cap is designed to cause the same price and revenue effects as the price cap, but without any chance of causing a price increase and with less anti-DSM bias than the price cap.

In order to guarantee that the hybrid cap behaves similarly to a price cap and prevents the Crew-Kleindorfer dilemma, it is only necessary to choose its slope to be more negative than the revenue curve. This should not be difficult; even though the empirical literature is vague, it strongly indicates that long-run elasticity is less than 2. From this elasticity we can compute a safe slope for the hybrid incentive as follows. We begin with a standard constant elasticity demand function in which quantity, $Q$, is considered a function of price, $P$, and in which $\alpha$, and elasticity, $\eta$, are constant:

$$Q = \alpha \cdot P^{-\eta}$$

$$R = P \cdot Q = \alpha \cdot P^{1-\eta}$$

$$\frac{dR}{dP} = \alpha \cdot (1-\eta) \cdot P^{-\eta}$$

$$\frac{dR}{dP} = (1-\eta) \cdot Q = -Q \quad \text{when } \eta = 2$$

This gives us the slope of the revenue curve, and, because we want the slope of the hybrid cap to be less than this slope, the equation also determines the slope of the steepest allowed hybrid cap. From the slope of the hybrid cap, it is easy to write down particular cap formulas. For instance, to cap the utility at its initial price, revenue, and quantity ($P_0$, $R_0$, and $Q_0$), we would use the following hybrid formula:

Revenue Cap Form: $R \quad 2R_0 - PQ_0$

$$P \quad \frac{2R_0 - R}{Q_0} \quad (4-10)$$

Price Cap Form: $P \quad 2P_0 - \frac{R}{Q_0}$

Comparing these to the general forms given by equation 4-8, we find that $\tilde{R} = 2R_0$, and $\tilde{P} = 2P_0$. Note that in the revenue-cap form, the revenue cap decreases with a price increase, while in a price-cap form, the price cap decreases with a revenue increase. Obviously this is only one of a whole family of hybrid caps that can be used. As long as they
are based on an elasticity less than two, they will be safe from the Crew-Kleindorfer dilemma. However, as the elasticity that the cap is based on decreases from two, the hybrid cap becomes more and more like a price cap and thus loses its positive DSM incentive properties.

4.7.1 The Incentives of a Hybrid Revenue-per-Customer Cap

In section 4.5 we argued for a revenue-per-customer cap as a replacement for a pure revenue cap, but then, when we faced the complexities of the Crew-Kleindorfer dilemma, we simplified our analysis by treating only the pure revenue cap. This cost us nothing in terms of insight, but in actual applications we would want to return to a hybrid form of the revenue-per-customer cap. So far we have also avoided the question of exactly what incentives, for or against DSM, will be generated by a hybrid cap. We know only that its behavior will lie somewhere in between a pure price cap and a pure revenue cap. We now remedy both of these shortcomings, and restore some of the pricing detail that has also been left behind.

The simplest hybrid of a price and revenue-per-customer cap uses a hybrid formula only on the energy component of costs and revenues. For the other components a simple rigid price cap is used. This may leave some minor problems with the incentive for load management, but generally, as was seen in Section 4.5, the utility has an incentive towards effective load management even under a price cap. Thus the following simple form should be sufficient, though a more complex form would be needed if price flexibility were desirable.

\[
P_E < \tilde{P}_E - \frac{R_E}{q_0 \cdot N}
\]

(4-11)

Where \( P_N \) is the price of access, \( P_L \) is the demand charge, \( P_E \) is the price of energy, and \( q_0 \) is the initial energy use per customer. This hybrid cap is based on an elasticity of 2. Noting that because of the magnitude of the subtracted revenue-per-customer term it is necessary to set \( \tilde{P}_E \) almost twice as high as \( P_E \).

Turning to the question of incentives for energy efficiency programs (DSM), we are particularly interested in the utility’s incentive to reduce \( q \), the energy use per customer. The calculation of this incentive is quite difficult, but the interested reader may find it in Appendix C-2. Fortunately the calculation has a simple outcome. The utility will have an incentive to reduce \( q \) provided

\[
P_E \cdot E < 2 \cdot c \cdot E.
\]

(4-12)

Returning to Table 4-1, we find that we have estimated \( P_E \cdot E \) at 90% of revenue and \( c \cdot E \) at 45% of revenue. Recall from table 4-1 that \( P_E \cdot E \) is energy charges and \( c \cdot E \) is energy costs. This inequality holds if customers are charged for energy less than twice the cost of producing the energy, taking into account the separate charges and costs for access and power. Thus
inequality 4.11 fails, but would hold as an equality. This indicates the firm will have no incentive to make any change in $q$. If the price of energy had been set at marginal cost, then the utility would have been motivated to reduce $q$. These findings indicate that a hybrid of a price cap and a revenue-per-customer cap can in fact provide protection from Crew-Kleindorfer pricing problems, and retain sufficient incentive properties from its revenue-based side to mitigate the adverse DSM effects of a pure price cap.

4.8 Dynamic Adjustments and the Need for Hybrid Incentives

As is well known, the demand for electricity is quite inelastic in the short run. As was noted at the end of Section 4.6, this implies a revenue curve that slopes only upward, which in turn negates the possibility of the Crew-Kleindorfer effect. This section shows why this short run analysis is inadequate and how to reconcile the short- and long-run views. The results are that, while the Crew-Kleindorfer effect is generally suppressed, it can be enabled by effective DSM, and that short-run elasticities can trigger an opposite but equally problematic effect. These possibilities justify the use of a hybrid mechanism. We make these points by focusing on three specific cases, but perhaps the most important lesson of this section may be learned simply by noting the complexities of the dynamics and gaming possibilities that are introduced by a pure revenue cap.

If the long-run demand for electricity were inelastic as is the short-run, both short- and long-run revenue curves would slope only upwards, and the Crew-Kleindorfer effect would indeed be impossible. Some may believe this to be the case, and the empirical literature, reviewed in Appendix C, does not refute this possibility. It is interesting to note however, that because the horizontal axis measures price, a revenue curve that slopes only upwards implies that an unregulated monopolist would maximize its profits by raising price without limit. Those who find such a profit maximizing strategy implausible, must believe there is a region of elastic demand, at least for higher prices.

Another crucial observation regarding demand, is that as the industry becomes more competitive and distribution companies lose some of their monopoly power, competition will alter the demand curves they face. Since, firms in an N-firm Cournot oligopoly typically face individual demand curves that are N times more elastic than the industry demand curve, one can expect this effect to be quite dramatic, especially regarding long-run demand. These considerations lead us to examine three particular cases all of which assume that short-run demand is inelastic, but which vary with respect to long-run assumptions and the effectiveness of DSM.
CHAPTER 4

4.8.1 Case 1: Revenue Curve as in Figure 4-1 with Utility Starting on the Left

In this case we assume that the long-run revenue curve behaves as in Figure 4-1: it is inelastic at low prices and elastic at high prices. We also assume that the utility starts in the long-run inelastic region.\textsuperscript{20} Without the constraint of short-run inelasticity, a revenue-capped firm would simply choose the high-price point shown in Figure 4-1. But if demand is short-run inelastic choosing such a high price will send revenue through the roof in the short run. Since a complete adjustment to the long run takes forever (at least in theory), the firm would actually have to overshoot the price target. Assuming the regulators will not tolerate a “temporary” violation of the revenue cap lasting for several years, such a strategy will be disallowed.

Conclusion: The Crew-Kleindorfer dilemma does not affect a firm that faces a demand curve that is both short-run inelastic and, at current prices, long-run inelastic. This is true even if a profitable high-price long-run strategy exists.

4.8.2 Case 2: Long-Run Demand is Elastic at the Initial Price

In this case we analyze point (A) on the down slope of the revenue curve shown in Figure 4-4, or perhaps a point that, due to competition, is on a curve that slopes downward at all prices. We have drawn the entire long-run revenue function, but drawn only small parts of several short-run curves: the upward sloping parts of the zigzag paths.

From point A, there is only one short-term option for the firm: to reduce price and move down its short-run revenue function to B, taking advantage of the short-run inelasticity of demand, and arriving at a lower revenue that complies with the revenue cap. (Ignore the path from B to D until case 3.) After arriving at B, the firm finds its revenue slowly rising as customers move toward their long-run demand curves. This violates the revenue cap, and at some point this violation will be noticed, and the utility will be required to comply with the cap once again. Once again it must decrease price, taking advantage of short-run demand inelasticity. This takes it on the next short diagonal path down and to the left. As this process continues, it is driven to lower and lower prices, and finally to C, even though there is a high price at which the revenue cap would eventually be satisfied.

Conclusion: A firm that is faced with a restrictive revenue cap, short-run inelastic demand and long-run elastic demand, will generally not be able to execute the Crew-Kleindorfer strategy. Instead they will be forced into repeated price cuts possibly leading to bankruptcy.

\textsuperscript{20} Although empirical evidence does not demand this conclusion, a well regulated industry will face inelastic long-run demand.
CHAPTER 4

Figure 4 Using DSM to Escape the Short-Run Elasticity Trap

4.8.3 Case 3: Long-Run Demand is Elastic and Demand Shifts Down

This is identical to case 2, except that the firm’s revenue curve unexpectedly shifts downward by enough to more than satisfy the revenue cap. This could be an autonomous shift caused by weather or a downturn in the economy, or it could be a shift caused by large-scale, utility-sponsored energy efficiency programs. Although DSM seems unlikely to produce a large short-run effect, DSM is of interest because it is the target of the revenue cap, and may for this very reason be unexpectedly successful.\(^{21}\) This shift is depicted by the lower revenue curve in Figure 4-4.\(^{22}\) If such a demand shift occurs, then at point B the firm will be above its new long-run revenue curve, and revenue will fall toward this lower curve. This fall in revenue takes the firm’s revenue below the cap, thereby allowing the firm to raise prices even

---

\(^{21}\) Note that if a regulator relies on a revenue cap, and no other controls, to motivate DSM, the utility will find it advantageous to run a number of programs that reduce demand, but that are not socially beneficial. This may widen the scope of “successful” DSM considerably. Also note that in the particular situation of this example, the payoff from “successful” demand reduction will be far greater than we have computed in our previous examples. These two effects could induce DSM programs quite different than anything seen in the past.

\(^{22}\) Recall that when DSM shifts demand; this does not mean simply moving along the demand curve, a process that involves only changing price and waiting for the response. Instead, DSM shifts the demand curve itself, so that less is demanded at any price, whatever that price may be. When the demand curve shifts down, the revenue curve obviously shifts down too.
though the short-run effect is to increase revenue. This process will continue until the firm reaches the point of maximum monopoly profits, shown as point D.

Conclusion: The firm may use DSM to escape the short-run elasticity constraint, and thereby make its way to the monopolist’s operating point.

Our final conclusion based on these three cases must be that although short-run inelasticity generally prevents the high price response to revenue caps described by Crew and Kleindorfer, case (2) nonetheless demonstrates a related problem, and case (3) shows that successful DSM could make the Crew-Kleindorfer strategy viable. Cases (2) and (3) make a pure revenue cap too risky in most real situations.

In the Section 4.6 we pointed out that an auxiliary price cap could restrain the simple Crew-Kleindorfer dilemma, but can such a mechanism be invoked to remedy the problems of cases (2) and (3)? Unfortunately it cannot be. In case (2) a price cap would have no effect, and in case (3) it would become locally binding if it did have any effect. In this case we end up with higher prices and price-cap regulation. Thus the only useful recourse in cases (2) and (3) is the hybrid price-revenue cap of section 4.7.

### 4.9 Flexible Caps and Relative Prices

In the first part of this chapter we discussed incentives for a firm to affect costs, quantity, and number of customers. In the second part we discussed the overall price response to a revenue cap. We now turn to a third area of consideration, applicable to both price and revenue caps whenever or not they are of the rigid form described in the first section. Because price caps are generally not rigid, and because a revenue cap or a hybrid cap is by its very nature flexible, this topic is essential to a complete understanding. The topic we now turn to is a firm’s strategy for setting relative prices when given price flexibility under PBR.

Both revenue caps and price caps affect relative prices, but the effects are quite different. Price caps are well known for their ability to induce prices similar to Ramsey prices. As we will soon see, revenue caps move prices strongly in the opposite direction. Ramsey prices are designed to maximize consumer welfare, given that a firm must cover costs. Ideally, prices should be set equal to marginal costs, but when there are fixed costs of production, these will not be covered by this “first-best” pricing scheme; instead, it is necessary for the firm to use a markup over marginal costs. The Ramsey problem is to find the set of markups that just covers fixed costs while making the smallest possible reduction in total consumer surplus. The solution to this problem is to use the set of markups that have minimum effect on

---

consumer demand. This is accomplished by marking up low-elasticity products the most and high-elasticity products the least.

Although price caps were not invented with the intention of inducing socially optimal relative prices (Beesley and Littlechild 1989), it was soon discovered that, in a multi-product firm, when the individual price cap is replaced with a Laspeyer’s price index, this will induce markups that are exactly proportional to Ramsey markups. Consider a flexible price cap based on a Laspeyer’s index, which is the typical practice. (The first term in equation 4-13 is a Laspeyer’s price index because it weights present prices (superscript 1) by last period quantities (superscript 0).) This means it is required that:

$$
\sum P_i^1 \cdot Q_i^0 = \sum P_i^0 \cdot Q_i^0 = R^0,
$$

(4-13)

where the sum is over products indexed by $i$. Although this equation looks something like a revenue cap because the cap is revenue at time 0, it does not cap revenue; instead, it caps a weighted sum of prices. This makes it a flexible price cap. In response to such a cap, the utility will set new relative prices, which will in turn induce new quantities. In the next round, the new quantities, $Q_i^1$, can be used as the new weights on price. If we continue to repeat these steps, the quantities will eventually converge to stable values. At these values, the profit maximizing prices will satisfy the following markup equation:

$$
\frac{P_i - C_i}{P_i} = \frac{1}{\epsilon_i} (1 - \lambda), \quad \text{where } \lambda < 1
$$

(4-14)

In other words, markups are proportional to the reciprocal of the elasticities, $\epsilon_i$. The proportionality constant is different than for true Ramsey prices because the firm is allowed to earn a profit, not just to cover costs. $\lambda$ measures the value to the firm of raising $R^0$, and it is known technically as the shadow price of relaxing the constraint.$^{24}$

It is difficult to explain this behavior intuitively because the effect is not a very powerful one, but we examine a simple example in order to see the various forces at work, and to compare this effect to the more dramatic relative price effect induced by a revenue cap.

---

$^{24}$ The derivation of this result is given in Einhorn (1991, p. 36).
CHAPTER 4

Table 4-3. Price Caps Affect Relative Prices Only Weakly

<table>
<thead>
<tr>
<th></th>
<th>Equal Markups</th>
<th>Flexible Markups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Cap:</td>
<td>2.20</td>
<td>2.20</td>
</tr>
<tr>
<td>Product Number</td>
<td>1 2</td>
<td>1 2</td>
</tr>
<tr>
<td>Elasticity</td>
<td>0.30 0.70</td>
<td>0.30 0.70</td>
</tr>
<tr>
<td>Marg. Cost</td>
<td>1.00 1.00</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>Price</td>
<td>1.10 1.10</td>
<td>1.14 1.06</td>
</tr>
<tr>
<td>Quantity</td>
<td>0.97 0.94</td>
<td>0.96 0.96</td>
</tr>
<tr>
<td>Profit</td>
<td>0.10 0.09</td>
<td>0.14 0.05</td>
</tr>
<tr>
<td>Total Profit</td>
<td></td>
<td>0.191 0.192</td>
</tr>
<tr>
<td>Change in Profit: (1)</td>
<td></td>
<td>0.0014</td>
</tr>
<tr>
<td>Change in Consumer Surplus: (2)</td>
<td></td>
<td>−0.0007</td>
</tr>
<tr>
<td>Change in Social Welfare: (1)+(2)</td>
<td></td>
<td>0.0007</td>
</tr>
</tbody>
</table>

We can see from the table that when the firm is allowed to price flexibly, it chooses to increase the markup on the inelastic product (number 1). This reduces the sales of product 1, but not by as much as the sales of product 2 are increased when the price of product 2 is lowered. Since the price cap allows the price of one product to go up as much as the other goes down, there is a net gain in sales. Sales are profitable, so there is a small net gain in profit. This gain in profit is less than 1% of the initial profit level (this is very much less than 1% of equity). Thus the effect on profit is rather small, so the firm does not have a very strong motive to make such a change.

We now turn to the effect of a flexible revenue cap on relative prices. This revenue cap can be expressed algebraically as follows:

\[ \sum P_i^1 \cdot Q_i^1 - \sum P_i^0 \cdot Q_i^0 = R^0 \]  \hspace{1cm} (4-15)

Although this expression looks very similar to that of the flexible price cap, notice that current prices are now weighted by current quantities instead of past quantities, so revenue is actually being capped in this case. When this formula is analyzed just as the price-cap formula was, a new markup equation is found. Both of these derivations may be found in appendix C-4.

\[ \frac{P_i - C_i}{P_i} = \frac{1}{\epsilon_i} (1 - \lambda) + \lambda, \quad \text{where } \lambda > 1 \]  \hspace{1cm} (4-16)

---

25 The products, for example, could be capacity and energy, or energy for customer class A and for customer class B.

26 Iso-elastic demand curves are used throughout these examples.
Table 4-4. Revenue Caps Affect Relative Prices Strongly

<table>
<thead>
<tr>
<th></th>
<th>Equal Markups</th>
<th>Flexible Markups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rev-Cap</td>
<td>2.10</td>
<td>2.10</td>
</tr>
<tr>
<td>Product Number</td>
<td>1 2</td>
<td>1 2</td>
</tr>
<tr>
<td>Elasticity</td>
<td>0.30 0.70</td>
<td>0.30 0.70</td>
</tr>
<tr>
<td>Markup</td>
<td>0.09 0.09</td>
<td>-0.72 0.68</td>
</tr>
<tr>
<td>Marg. Cost</td>
<td>1.00 1.00</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>Price</td>
<td>1.10 1.10</td>
<td>0.58 3.17</td>
</tr>
<tr>
<td>Quantity</td>
<td>0.97 0.94</td>
<td>1.18 0.45</td>
</tr>
<tr>
<td>Profit</td>
<td>0.10 0.09</td>
<td>-0.49 0.97</td>
</tr>
<tr>
<td>Total Profit</td>
<td>0.19</td>
<td>0.48</td>
</tr>
<tr>
<td>Change in Profit: (1)</td>
<td></td>
<td>0.280</td>
</tr>
<tr>
<td>Change in Consumer Surplus: (2)</td>
<td></td>
<td>-0.732</td>
</tr>
<tr>
<td>Change in Social Welfare: (1)+(2)</td>
<td></td>
<td>-0.447</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.74</td>
<td></td>
</tr>
</tbody>
</table>

This formula differs from the price-cap markup formula in two ways. First, an extra $\lambda$ is added to the inverse elasticity term. Second, $\lambda$ is greater than 1, which causes the inverse elasticity term to be negative. This reversal of sign means that the firm will decrease rather than increase the markup on low elasticity products, as shown in Table 4-4.

Notice that all of the effects are more dramatic for the revenue cap than for the price cap. Price increases not by 4% but by 300%, and decreases not by 4% but by 42%. Profit increases not by less than 1% but by nearly 150%, and the changes in consumer surplus and social welfare are also hundreds of times larger. As the elasticity of a product approaches 100%, the markup on the product tends toward infinity.

Surprisingly, the reason for these effects is also subtle. For both products, an increase in price causes a strong increase in profit. In fact, the effects are similar because they are primarily caused by the increasing gap between price and cost. The advantage of increasing the elastic product’s price is that this will cause less increase in revenue than will be caused by an increase in the price of the less elastic product. For this reason, the tradeoff favors a high price on the elastic product and a low price on the inelastic product.

These effects are exacerbated when one product is elastic and the other inelastic. In this case, the elastic product will be priced as high as possible, resulting in essentially no sales, which uses up a minimum of the allowed revenue. The firm can then collect the remainder of the allowed revenues from the inelastic product, which will be very profitable. These conclusions are the result of numerical analysis, so they may not apply universally. However, they are certainly true in many cases.
The above conclusions concerning overpricing may be mitigated when short-run elasticities are small, but dynamic pricing behavior will be very complex, and may well be problematic.

Because price-caps cause the inelastic product to be marked up more, and revenue caps cause it to be marked up less, we can assume that a hybrid cap could have a very minor effect on relative markups. Revenue caps produce the more powerful effect on relative prices, so it will probably be necessary to use a hybrid cap that leans more heavily in the direction of a price cap, in order to achieve neutrality.

A complete analysis of the effect of hybrid caps on relative prices is quite complex and beyond the scope of this paper. However, until such analysis is done, hybrid caps that allow price flexibility should be located near the price-cap end of the spectrum.

A simpler way to manage the strong relative-price incentives of revenue caps is to fix relative prices. This must be done as a separate regulatory measure, rather than being built into the revenue cap. But it is a simple matter to require a firm to keep relative prices fixed or to change them by less than a certain percent per year. A defacto constraint on relative prices is probably in effect wherever revenue caps are in use. Relative prices could be fixed using the normal rate review procedure, which is still in place. In such a setting, regulators can simply refuse to allow relative price shifts they are uncomfortable with, even those that do not violate the revenue cap. This process may result in less efficient pricing than would be achieved under a price cap with pricing flexibility, but it does effectively prevent pricing anomalies that would otherwise be caused by a revenue cap.

4.10 Summary and Conclusions

Revenue caps were proposed as substitutes for price caps in order to eliminate the anti-DSM bias of price caps while maintaining the incentives to minimize costs. We have shown that this idea is basically sound, although the use of revenue caps presents a new set of problems. Hybrid caps simply allow a utility to compromise between a price cap and a revenue cap. Fortunately it is possible (under some price/cost conditions) for such a compromise to mitigate all of the revenue cap’s problems without restoring any of the anti-DSM bias found in a price cap.

The use of revenue caps poses three potential problems:

1. The utility may dramatically alter relative prices.
2A. The utility may respond by setting price at or above the monopoly level.
2B. The possibility that a small reduction in the revenue cap will produce a large and unpredictable reduction in price (an effect related to the Crew-Kleindorfer effect).
3. An incentive to reduce sales regardless of the social benefit.
The first problem can be solved by regulating relative prices, a process that is going on de facto at all utilities that are currently using revenue caps. However, the use of a hybrid price-revenue cap would significantly mitigate this problem, and a correctly designed hybrid may eliminate it completely, which would allow the utility price flexibility; a desirable step in the direction of competition.

The second two problems are both related to the shape of the revenue function. Typically revenue increases with price for low prices and decreases with price for high prices. If a firm is in the low-price region, a revenue cap with no other constraints, will allow the firm to meet the revenue cap by setting price in the high price region. This is the problem pointed out by Crew and Kleindorfer, and it is easily solved by three different techniques: (1) short-run price inelasticity coupled with restrictions on temporary violations of the revenue cap, (2) a non-binding price cap, or (3) a hybrid price-revenue cap. Currently it is probably being solved in practice by the first technique, which occurs through standard implementation procedures. The only real possibility of the Crew-Kleindorfer effect to take effect is in a market with long-run elastic demand, and the possibility of an autonomous (or DSM caused) drop in demand that happens after the mechanism is set in motion. This seems unlikely but our recommended hybrid cap would prevent it.

Problem (2B) is much more likely to occur. It requires only that long-run demand have an elasticity near 1 or greater. In this case, since the Crew-Kleindorfer effect will be prevented by short-run demand elasticity, the firm will be forced to meet its revenue cap by a price reduction. But since reducing price by a modest amount does little or no good in reducing long-run revenue, the firm will eventually be forced into a drastic price cut. This would probably force the abandonment of the revenue cap in order to avoid putting the firm out of business. Again, problem 2B would be prevented by the use of a hybrid cap.

The third problem is that a revenue cap, while producing an incentive to reduce sales, does not target that incentive towards economically justified energy efficiency improvements. The incentive is too encompassing, and so encourages non-socially beneficial as well as beneficial reductions in sales. Again, a hybrid cap would greatly reduce or even eliminate this problem.

This leads us to recommend the hybrid price-revenue cap as a replacement for a pure revenue cap if one is concerned with the incentives for the utility to manipulate demand. More specifically the cap should take a revenue-per-customer approach and should be based on an elasticity of 2 or less. A correctly designed hybrid cap will not allow a price increase (except due to DSM) and will eliminate most if not all of the anti-DSM bias associated with a price cap. Such a cap would only need to replace the energy part of a price cap, and could take a form as simple as the following.

\[ P_E < \bar{P}_E - \frac{R_E}{q_0 \cdot N} \]  

(4-17)
$P_E$ is the price of energy, $\tilde{P}_E$ is like a price cap only high to compensate for the following revenue term. The subtracted term measures revenue from energy charges divided by initial energy per customer times the number of customers. Probably the most important remaining question is what changes in relative prices will be induced by a hybrid cap. Until this is answered utilities using the hybrid cap may have to maintain the tradition of implicitly regulating relative prices.

In spite of its limitations, the subject matter of this chapter covers much new ground that should be of interest to those considering or already using revenue caps. Also in spite of the simplifications that are necessary in this chapter, it should be presumed that the basic results are true of the more complex revenue caps found in practice. Some have already made claims to the contrary, asserting for example that particular revenue caps are not subject to Crew-Kleindorfer type effects. On the face of it, such claims seem quite unlikely unless the particular caps have additional mechanisms designed to reverse the effects described in this chapter. Thus any claims for exemption from these conclusions should be documented by careful calculation before they are accepted. In particular these calculations must account for the price elasticity of demand. Although revenue-cap incentives have not been directly covered in this depth by either the academic or the policy literature, this chapter is intended only as an introductory intuitive treatment of the subject and not as a final and definitive treatment; many questions remain unanswered.
APPENDIX C

Incentive Properties of a Hybrid Cap, and Long-Run Demand Elasticity

C.1 Overview

This appendix further addresses two issues raised in Chapter 4: the incentive effect of a hybrid price cap and the probably value of the elasticity of long-run electricity demand. Section C.2 focuses on the incentive of primary interest in Chapter 4, the incentive to implement energy efficiency programs. This is analyzed for a hybrid cap composed of a mixture of a price cap and a revenue-per-customer cap. The results confirm equation (4-11) of section 4.7. Section C.3 focuses on demand elasticity with particular emphasis on the empirical literature.

C.2 Incentives Under a Hybrid Price/Revenue-per-Customer Cap

The goal of this section is to evaluate the incentive to engage in effective energy-efficiency programs under a hybrid cap that combines a price cap with a revenue-per-customer cap. We begin by specifying such a cap.

The simplest hybrid revenue-per-customer cap uses a hybrid formula only on the energy component of costs and revenues. For the other components a simple rigid price cap is used. This may leave some minor problems with the incentive for load management, but generally, as was seen in Section 4.5, the utility has an incentive towards effective load management even under a price cap. Thus the following simple form should be sufficient, though a more complex form would be needed if price flexibility were desirable.

\[ P_N < ar{P}_N, \quad P_L < ar{P}_L, \quad \text{and} \]

\[ R_E < \bar{R}_N \cdot N - (\epsilon - 1) P_E \cdot m_0 \cdot q_0 \cdot N \]  

Where \( P_N \) is the price of access, \( P_L \) is the demand charge, \( R_E \) is the revenue from the energy charge, \( \bar{R} \) is fixed, \( P_E \) is the price of energy, \( q_0 \) is initial energy use per customer, and \( N \) is the number of customers and is assumed fixed. The mixture of this hybrid cap is based on an elasticity of \( \epsilon \). The new variable, \( m \), is the DSM control parameter. This actually modifies the meaning of \( q \), so that \( m \cdot q \) is now the true energy use per customer. \( m_0 \) is the initial value of this variable. The variable \( m \) is needed because we wish to differentiate profit (\( \pi \)) with respect to \( m \) in order to evaluate the incentive to promote energy efficiency.

We will be interested only in the hybrid energy revenue cap, and until the end of our calculations we will not need to distinguish between the various constants that multiply \( P_E \). Because we are dealing only with the energy part of the cap we will simply drop the subscript...
APPENDIX C

Note that we have now introduced the fact that energy use is a function of the price of energy. Because we omitted this fact in Section 4.5, we mis-estimated the power of revenue cap incentives in that section. Here we will correct that simplification. Substituting (3) into (2) and solving for \( \tilde{R} \) we have:

\[
\tilde{R} = P \cdot m \cdot q(P) \cdot N + \alpha P \tag{4-4}
\]

Because the total differential of \( \tilde{R} \) is zero, we have:

\[
(P \cdot m \cdot q' + m \cdot q + \alpha) dP + P \cdot q \cdot dm = 0. \tag{4-5}
\]

Using the assumption that demand elasticity, \( (dq/dP)(P/q) \), equals \(-\eta\) allows us to find:

\[
\frac{dP}{dm} = \frac{P}{(\eta - 1) m - \alpha / (q \cdot N)} \tag{4-6}
\]

We now expand \( \pi = R - C \), the definition of profit, to find:

\[
\pi = \tilde{R} - \alpha P - c \cdot m \cdot q(P) \cdot N \tag{4-7}
\]

Differentiating this with respect to profit gives

\[
\frac{d\pi}{dm} = -\alpha \frac{dP}{dm} - c \cdot N \cdot \left( q + m \frac{dq}{dP} \frac{dP}{dm} \right) \tag{4-8}
\]

Substituting for \( dP/dm \) and \( dq/dP \) in equation (8) gives
\[
\frac{d\pi}{dm} = \frac{-\alpha P}{(\eta - 1)m - \alpha / (q \cdot N)} - c \cdot N \left[ q + m \cdot \left( \frac{-\eta q}{P} \right) \left( \frac{P}{(\eta - 1)m - \alpha / (q \cdot N)} \right) \right]. \tag{4-9}
\]

This which to
\[
\frac{d\pi}{dm} = \frac{c \cdot N \cdot [m + \alpha / (q \cdot N)] q - \alpha P}{(\eta - 1)m - \alpha / (q \cdot N)}. \tag{4-10}
\]

In the initial state of regulation, we have set \(m_0 = m\), and \(q_0 = q\), so we can make these substitutions now as we substitute for \(\alpha\) in equation (10). This gives us
\[
\frac{d\pi}{dm} = \frac{c \cdot N \cdot [m + (\epsilon - 1)m] q - (\epsilon - 1)m \cdot q \cdot N \cdot P}{(\eta - 1)m - (\epsilon - 1)m}. \tag{4-11}
\]

This simplifies to
\[
\frac{d\pi}{dm} = \frac{(\epsilon - 1)R - c \cdot N \cdot [\epsilon \cdot m] q}{(\epsilon - \eta)m}. \tag{4-12}
\]

Note that the sign of both numerator and denominator have been reversed because the hybrid cap should be based on an elasticity, \(\epsilon\), that is greater than the actual elasticity of demand, \(\eta\). So we may now assume that the denominator is positive. Note that \(c \cdot m \cdot q \cdot N\) is just the cost \((C)\) of producing energy. Thus
\[
\frac{d\pi}{dm} < 0 \quad \text{if and only if} \quad (\epsilon - 1) \cdot R < \epsilon \cdot C \tag{4-13}
\]

This indicates the utility will have an incentive to promote energy efficiency provided the revenue-cost inequality holds. If energy were priced at it’s marginal cost this inequality would certainly hold. In the case of the example in section 4.5, \(R = 2 \cdot C\), so with \(\epsilon = 2\), as assumed in section 4.7 the inequality becomes an equality, and the utility is neutral towards energy efficiency.

As a final step we transform the hybrid cap from its revenue-cap form to its price-cap form. Solving equation C-1 for \(P\) yields:
C.3 Evaluation of Long-Run Elasticities for Electricity Demand

Below, we consider two theoretical approaches to determining the long-run elasticity of electricity demand. First, the Averch-Johnson model is shown to predict elastic demand. Second, a model based on the assumption that the firm is optimally regulated (with full information) is shown to produce the same result. However challenges to both theoretical lines of reasoning exist and are put forward. This leads to section C.3.2 which considers the empirical evidence.

C.3.1 Are Utilities Currently Operating in the Inelastic Portion of their Demand Curves?

As was demonstrated in Chapter 4, unless the utility is operating in the long-run inelastic portion of its demand curve, using a revenue cap is rather dicey business. In this section we investigate the question: is the typical utility facing a demand curve that is inelastic in the long run? We will not be able to make a definitive answer because both the theoretical and empirical literature is inconclusive. In particular we will find the following:

- The Averch-Johnson model predicts long-run elastic demand;
- Because standard COS regulation is not pure ROR, A-J may not apply;
- The empirical evidence is ambiguous; and
- The firm regulated for the social optimum does operate in the long-run inelastic region.

We begin with the Averch-Johnson model of rate-of-return (ROR) regulation. It assumes that at every instant the firm is forced to set its price so that it earns exactly some allowed rate of return \( R^* \) on its invested capital. This allowed rate is over and above the cost of capital. A second, and less restrictive assumption, is also made: that the utility’s output is an increasing function of both capital and labor.

The argument for demand elasticity under ROR proceeds by contradiction. Assume that demand is inelastic at the firm’s equilibrium, so that a price increase (quantity decrease) increases revenue. The firm would decrease output by decreasing its labor input, thereby not changing its rate base or the amount of profit it is allowed. Decreasing labor decreases costs, while decreasing output allows a price increase that increases revenue (by the assumption of inelasticity). Thus revenue is increased while cost is decreased, so the net effect is an increase...
in profit. Thus the firm was not at equilibrium as assumed, which is a contradiction. This shows that the firm’s equilibrium must be in the elastic region.

The above proof can best be understood through the following dynamic. As long as the firm is in the inelastic region of its demand curve it can increase revenue by cutting output. As long as it cuts output by cutting labor this will not affect it’s allowed profit level, but will reduce costs. So it just keeps cutting output, and earning more revenue for less cost until output is so low that it finds itself in the elastic portion of the demand curve. If there is no inelastic portion, then the firm will continue to lower output and raise price without limit.

For a firm having a normal production function and under ideal ROR regulation, this argument is conclusive. However actual cost-of-service regulation is a bit more complex than pure ROR. First ROR fixes prices periodically and lets actual return differ from allowed return during these periods. Second it maintains a standard of “used and useful” for all capital investments. This latter restriction is relevant in that it may, at some point, prevent the reduction of labor that was hypothesized in the above argument. If labor is reduce too far and output falls too much below what the capital is capable of producing, the regulator may find the capital no longer “used and useful.” This may thwart the above described strategy of the firm to move to the elastic region of demand. We cannot say for sure that it will, because we do not know at what output level demand becomes long-run elastic.

C.3.2 Empirical Estimates of the Long-Run Elasticity of Electricity Demand

Since theory fails to answer this question we turn next to empirical work on demand elasticities. This summary of empirical work on demand elasticities is drawn from Chapter 7 of E. R. Berndt’s *The Practice of Econometrics: Classic and Contemporary* (1991).

Econometric analysis of elasticities for electricity are complicated by several factors. Primary among these is the derived nature of the demand for electricity. Electricity is consumed not by people directly but by a stock of electric appliances. In the short-run, people can only vary their electric demand by changing their demand for the services these appliances provide. In the long-run, however, the stock of appliances can be changed. While this may not seem that complex, it does create a daunting data requirement. Franklin M. Fisher and Carl Kaysen, among the first analysts to tackle such a modeling approach directly concluded “to estimate [the stock of appliances] by states and years with any kind of reliability is simply out of the question.”

To fully model the demand function, data on utilization of an existing stock at a given price is needed as well as data on the changes in the stock resulting from a given price. Furthermore, since choices of appliances as assumed to be based on expectations of future prices, consumers expectations about prices must be modeled. Although Fisher and Kaysen did try to work around direct estimation, in the end they cautioned that “it is worth reiterating how poor our data really are.”

---

Even when these data requirements can be worked around, the relationship between demand and the multi-part tariffs common in the electric industry makes the choice of price indicators nettlesome. The basic problem is that the amount a customer demand affects the price they are charged because of the multi-part tariffs and price affects demand because the downward slope of the demand curve. The question then arises whether to use the average price or the marginal price? Generally an average price can not adequately capture demand reaction under a complex rate-structure and so using it as a regressor can result in a serious bias. Using a single number for the marginal price, however, fails to capture changes in price that may induce a customer to change their demand such that they end up on a different tariff-block.

The complexity of choosing real world indicators aside, econometricians are always faced with a fundamental question of which functional form to use when simplifying the real world into a model. The most common to date largely because of it’s simplicity is log-linear forms. The resulting constant elasticity of price and income from a simple model defies reality in the extremes. Adding layers of complexity can result in result in elasticities that vary with price and income, but these risk running afoul of underlying economic theory. This has lead to interest in more flexible forms such as the translog and the generalized Leontief which of course come with their own host of complexities.

Not surprisingly, the empirical estimates of long-run price elasticity of demand vary greatly based on the assumptions used. Efforts to directly include appliance stocks vary from $-1.1$ to $-1.3$ when average price is used as a regressor and from $-0.4$ to $-0.7$ when marginal price is used. Indirect efforts that avoid using measures of the stock of appliances result in mean long-run elasticities of $-0.8$ when marginal price is used and $-1.0$ with average price is used.

Since the inelastic region of demand is represented by an elasticity of greater than negative one, these numbers paint a very ambiguous picture. On the whole though, they seem to suggest that firms are at best only producing slightly into the inelastic region of demand.

Only in the area of socially optimal theory can we find some certainty about what part of the demand curve the firm should be operating in. The easiest way to see why an optimally regulated firm will produce in the inelastic region of the demand curve is too look at a graph of profits against quantity. The profit curve takes the form of a hill. The top of this hill is the profit maximizing point where marginal revenue equal marginal cost. This is where a
monopolist would choose to operate. We know that the inelastic region of the demand curve begins when marginal revenue equal zero, so we know that this region will start to the right of the peak of the hill. We also know that the optimally regulated firm will produce at the Pareto-optimal point where price equals marginal cost and profits are zero. This is the right most point on the profit hill. Putting what we know together, we can see that the optimally regulated firm will produce in the inelastic region of the demand curve. Figure C-1 depicts this intuitive proof (Train, 1991).
C.4 Appendix to Section 4.9: Relative Prices Under Price and Revenue Caps

C.4.1 Relative Prices under a Price Cap

Begin by writing down the price-cap mechanism. Superscript 1 denotes this period while superscript 0 denotes last period. Both price, $P$, and quantity, $Q$, are vectors.

$$ P^1 \cdot Q^0 = R^0 $$

The firm maximizes profit, $\pi = R - C$ under the price-cap constraint, so we write down the Lagrangian for this maximization:

$$ = P^1 Q^1 - C(Q^1) - \lambda \cdot (P^1 Q^0 - R^0) $$

We assume that cross elasticities are zero, and this allows us to differentiate each price with respect to its quantity. The partial of the Lagrangian with respect to each particular quantity equals zero.

$$ \frac{dP_i^1}{dQ_i^1} Q_i^1 + P_i^1 - C_i'/P_i - \lambda \frac{dP_i^1}{dQ_i^1} Q_i^0 = 0 $$

Now divide through by $P_i^1$ and multiply the last term by $Q/Q$.

$$ \frac{dP_i^1}{dQ_i^1} \frac{Q_i^1}{P_i^1} + \frac{P_i^1}{P_i^1} - \frac{C_i'}{P_i^1} - \lambda \frac{dP_i^1}{dQ_i^1} \frac{Q_i^0}{P_i} \frac{Q_i^1}{Q_i} = 0 $$

Now use the fact that elasticity, $\epsilon$, is given by $-(dP/dQ)(Q/P)$.

$$ \frac{1}{\epsilon} + 1 - \frac{C_i'}{P_i} + \frac{\lambda}{\epsilon} \frac{Q_i^0}{Q_i} = 0 $$

Now substituting markup, $\mu$, for $1-MC/P$ and rearranging we have:
\[ \mu_i = \frac{1}{\epsilon_i} \left( 1 - \lambda \frac{Q^0_i}{Q^1_i} \right) \]

If \( Q^0 \) is now replaced by \( Q^1 \), and the firm re-optimizes, and this sequence is repeated the quantities will quickly converge to stable values. At this equilibrium we will have \( Q^0 = Q^1 \), which gives the result we were seeking.

\[ \mu_i = \frac{1 - \lambda}{\epsilon_i} \]

C.4.2 Relative Prices under a Revenue Cap

Begin by writing down the revenue-cap mechanism. Superscript 1 denotes this period while superscript 0 denotes last period. Both price, \( P \), and quantity, \( Q \), are vectors. Note that it differs from a price-cap mechanism in that the quantities are current, thus making the index being capped the current revenue and not just a fixed weighting of prices.

\[ P^1 \cdot Q^1 = R^0 \]

The firm maximizes profit, \( \pi = R - C \) under the price-cap constraint, so we write down the Lagrangian for this maximization:

\[ = P^1Q^1 - C(Q^1) - \lambda \cdot (P^1Q^1 - R^0) \]

We assume that cross elasticities are zero, and this allows us to differentiate each price with respect to its quantity. The partial of the Lagrangian with respect to each particular quantity equals zero.

\[ \frac{dP^1_i}{dQ^1_i} Q^1_i + P^1_i \cdot C'_i - \lambda \left( \frac{dP^1_i}{dQ^1_i} Q^1_i + P^1_i \right) = 0 \]

Now divide through by \( P^1_i \).
\[
\frac{dP_i}{dQ_i} \cdot \frac{Q_i}{P_i} + \frac{P_i}{P_i} - \frac{C_i'}{P_i} - \lambda \left( \frac{dP_i}{dQ_i} \cdot \frac{Q_i}{P_i} + \frac{P_i}{P_i} \right) = 0
\]

Now use the fact that elasticity, \( \epsilon \), is given by \( -(dP/dQ)(Q/P) \).

\[
-\frac{1}{\epsilon} + 1 - \frac{C_i'}{P_i} + \frac{\lambda}{\epsilon} - \lambda = 0
\]

Now substituting markup, \( \mu \), for \( 1 - MC/P \) and rearranging we have:

\[
\mu_i = \frac{1 - \lambda}{\epsilon_i} + \lambda
\]

which is the formula given in section 4-9.