Price Cap Regulation of Telecommunications Services: A Long-Run Approach

Ingo Vogelsang

February 1988
The research described in this report was supported by a grant from the John and Mary R. Markle Foundation.
Price Cap Regulation of Telecommunications Services: A Long-Run Approach

Ingo Vogelsang

February 1988

Prepared for The John and Mary R. Markle Foundation
With its recent Notice of Proposed Rulemaking, the U.S. Federal Communications Commission (FCC, 1987) has initiated a discussion about the possible use of price caps to replace rate-of-return regulation for telecommunications services. This would be a major change in the FCC's approach to tariff regulation. In opening the discussion, the FCC's goal "is to determine if the price cap model could be adapted so that it better protects and promotes consumer welfare and the public interest in an efficient and reasonably priced telecommunications network than does cost-of-service regulation." In its analysis the FCC raises a host of issues and asks a number of questions.

Using economic theory, this study seeks to answer some of the questions. This approach contributes a different perspective from that of interested parties such as consumers or competing carriers. The comprehensive Summary has been written in a nontechnical question-and-answer format in an effort to reach a wider audience, which includes academic economists, public utility managers, and policymakers. The Note itself, although more technical, presents its main arguments in fairly nontechnical terms.

The Note was funded under a grant from the John and Mary R. Markle Foundation. It is one of a continuing series of RAND studies dealing with public policy issues in the telecommunications field. Thanks are owed to Leland Johnson and Bridger Mitchell for many insightful suggestions.
Price caps have recently been put forward for discussion by the U.S. Federal Communications Commission (FCC, 1987) as an alternative to rate-of-return regulation for the regulated telecommunications sector. Price capping's intended purpose is to constrain the price increases for a basket of services (of a regulated firm or firms) to the increase either of cost items outside the firm's control or of the Consumer Price Index (CPI), minus a percentage "x" for expected productivity increases.

Price caps bear some similarity to a regulatory adjustment process suggested by this author and J. Finsinger (Vogelsang and Finsinger, 1979; hereinafter called V-F). Under this process the regulated firm is allowed to charge prices for a basket of its services in period t such that it would make no excess profits if these prices were applied to the outputs and costs of the previous period t-1. This pricing constraint clearly is based on profits and therefore on costs, whereas the price caps currently under discussion in the FCC proceeding appear to be completely different from profit (or rate-of-return) regulation. However, price caps cannot be maintained over a long time period without profits and the firm's achieved rate of return resurfacing as arguments for changing the price caps regulation. This study shows that in the long run, price caps lead to a modified version of the V-F regulatory adjustment process. Some of the desirable properties of the original V-F mechanism extend to these long-run price caps.

In particular, if the regulated firm tries to maximize profits it will choose individual prices within a capped average such that total consumer welfare (expressed as consumer surplus) will increase over time. At the limit consumer surplus will converge to the maximum attainable at normal profit levels for the regulated firm. Prices with this property are called "Ramsey" prices. Under the V-F mechanism they can be achieved without the necessity for regulators to assign costs to individual services.
By combining the V-F mechanism with the price cap approach, this study extends the V-F mechanism in ways that reduce or eliminate its two major shortcomings, a lack of incentives by the firm for cost reductions and a lack of adaptability to a changing environment. Both of these improvements are achieved by means of price cap formulas that are sufficiently adaptive to the changing environment; therefore, they need to be formally reviewed only after several years.

By linking price caps to the V-F mechanism, we gain an additional tool and new insights. The previous literature on price caps has at least two shortcomings. First, it concentrates almost solely on the issues of cost efficiency by the regulated firm and equity for certain consumer groups. The issue of allocative efficiency of pricing (that is, correctly dividing up services to the consumers who value them most) has not been raised. The analysis in this Note does precisely that and shows at the same time the implications for the other two issues raised in the literature.

Second, the literature has examined price caps under a very limited time horizon. This would be acceptable if the regulatory agency could and wanted to commit itself to total deregulation within this time span. Otherwise, one has to ask, what will happen when the time comes for a general review of the price caps? This study makes specific suggestions as to how this question can be answered.

By explicitly examining the long-run conditions, this Note provides a new basis for addressing the following questions raised by the FCC in its Notice:

*Question 1:* "At one extreme, the cap requirement could be interpreted to impose a ceiling on the average rates of capped services overall.... At the other extreme the cap requirement could mean a ceiling on the rate associated with each rate element of a service.... In between the two extremes are a myriad of possibilities for defining the concept of price caps." Which "interpretation strikes the best balance between our primary objective of protecting ratepayers against
unreasonable charges for services and our giving carriers both the flexibility to introduce new, innovative services quickly and to provide the most efficient mix of services their networks permit and the incentive to do so"?

Answer: This study shows that imposing a ceiling on the average rates for capped services overall leads to an allocatively efficient price structure and gives carriers flexibility. For purposes of calculating the averages, rates should be weighted by quantities sold in the previous period. These weights should be adjusted each period so that they reflect the actual demand of customers. This averaging leaves room for the carrier to adjust individual rates according to attributable costs and perceived demands. Under assumptions described in the Note, the structure of price caps will then converge to allocatively efficient Ramsey prices. On the other hand, income distributional concerns may require us to hold down specific rates for disadvantaged groups. Usually such individual capping will reduce allocative efficiency. However, this need not be the case if a two-part tariff approach to price caps is used and if, because of "network externalities," other subscribers benefit from disadvantaged groups having telephones.

Question 2: Should "a cap constraint be applied to (a) all services and (b) with effective tariffs on the date that any price cap approach to regulation would be introduced"?

Answer: (a) A cap constraint need not be applied to all services. For example, those offered under competitive conditions could be excluded. However, if the carrier has some market power for the uncapped services, then in the long run the uncapped service may cross-subsidize the capped services. This is likely to be politically more acceptable than cross subsidization from regulated to unregulated services. (b) Without some trial and error by the carrier we do not really know what the best rate structure is. The current rates as initial price caps have three clear advantages in their favor. First, they have been tried. At these rates demand can be met. Second, they are unlikely to arouse opposition from consumers. Third, they currently
provide the carriers with a rate of return that has already been approved. Under a price cap approach, using average rates for capped services overall, any carrier will carefully move individual rates from these initial values toward more rewarding and more efficient levels.

**Question 3:** How should we "treat a dominant carrier's proposal to restructure the rates of a previously capped service"?

**Answer:** Restructuring here refers to the pricing of multiple services that can be combined in different ways. Restructuring should be allowed to the extent that quantitative data on the use of these services in the previous period are available. Then the following approach can be used: The restructured rates should be such that customers, when buying what they bought before, have lower outlays under the restructured tariffs than otherwise. Then they are at least as well off under the restructured rates. Due to the restructuring they will actually change their choice, a fact that will further improve their welfare.

**Question 4:** To what degree does predation remain a threat, and how could we "best prevent predation under a price cap approach to regulating rates"?

**Answer:** The scope of price caps has to be determined by considerations of competition and equity. Competition can force the firm to choose more efficient prices at any time. However, there is a tradeoff. Predation, although unlikely, cannot be fully ruled out. For instance, the regulated carrier might lower its rates for competitive services below the efficient level and charge fully profit-maximizing rates for monopoly services with the average obeying the caps. After successfully driving out its competitors it would reduce the previous monopoly rates and increase the previous predatory rates, again leaving the average within the capped range. The carrier would not gain an overall price advantage from this. However, under economies of scale there might be an advantage in terms of quantities because the carrier would not have to share the market with others. It is theoretically possible that these additional sales represent enough of a motive for
predation. Such cross subsidization as a basis for predation would be even less possible if services supplied in competition with other carriers would not be capped at all.

**Question 5:** What is the suitability of factors such as changes in access charges, in international accounting rates, in tax laws, in the Separations Manual and in industry-wide costs of operation "as bases for adjusting the rates of some or all capped services"? In particular, should "any price cap plan...require periodic adjustments to price ceilings (or floors) to reflect changes in the purchasing power of money" expressed by the CPI or Producer Price Index? How could "such adjustments...be made without our becoming mired in the cost accounting and rate of return assessments associated with traditional rate-of-return regulation"?

**Answer:** Without any adjustment, price caps will eventually lead to some crisis situation, either one of financial distress or one of unduly high profits for the firm. In response, the main question remains: To what extent should price cap adjustments be automatic and to what extent should they require a formal rate review? Because an automatic mechanism cannot be perfect, there will have to be formal revisions of price caps from time to time. However, automatic adjustments can help stretch the time span between formal reviews. The more an automatic adjustment reflects the actual cost changes of the carrier, the longer this time span can be. At the same time, when actual cost changes are covered, cost accounting problems increase, and the incentive to cut costs weakens. To strike a balance, actual cost changes should be used for large items that cannot be influenced by the carrier, such as taxes and carrier access charges for interexchange carriers. These changes should be given their respective weights in total costs, and the remaining cost changes should be factored in by using industry-wide or nation-wide indices. Among these indices the CPI has two advantages. That it has changed, and by how much, is easily understood, and it reflects the buyers' purchasing power. These advantages could prove to be more important than the fact that the CPI may poorly trace the carrier's actual cost changes.
Question 6: How should an adjustment factor reflecting productivity increases be selected "for our use in revising price caps for domestic and international services"? Should there be "a different productivity factor for AT&T and the LECs [Local Exchange Carriers], and, if so, how...could or should [the factors] be developed"? How should "changes in productivity...be measured with respect to regulated services and, for operations that have some fixed costs, how [may] changes in productivity...be separated accurately from exogenous changes in demand"?

Answer: The basis for an adjustment factor reflecting productivity increases should be the carrier's past performance of total factor productivity over an extended period of five to ten years. Assuming that this past performance can be separated for AT&T and the LECs, these carriers should also have different standards set for the future. The study suggests initially using an overall value for past productivity measures to combine regulated and unregulated services. Exogenous changes in demand cannot be separated accurately from productivity increases so they should also be combined, given the carrier's responsibility for forecasting demand and for installing the correct capacity and correct type of equipment. Productivity increases applied to the price cap formula could be negotiated ex ante, or a percentage of the actual past (total factor) productivity increases could be taken.

Question 7: "How frequently [should] price caps...be reevaluated or adjusted to reflect changes to each of the adjustment factors" discussed above? Should "the general reexamination...be scheduled with prescribed frequency (e.g., every five years) or should [it] occur when the Commission or a carrier perceives a need for such a review"?

Answer: If capped prices on average are expected to follow a downward trend over time, the carrier should be required to apply the adjustment factors and recalculate average price caps once a year. More likely, inflation will outpace productivity increases. In this case the carrier should initiate price changes. The carrier would have to file changes with a demonstration that they comply with the caps. Unless the
FCC or an interested party showed that the caps would be violated, the filing would automatically result in new tariffs after one month. A general reexamination should occur after a prespecified period of five to ten years. The only way an earlier reexamination could be triggered is an emergency situation for the carrier (e.g., Chapter 11 filing). As a result of the reexamination, the average caps should be adjusted once and for all, reflecting the actual profitability of the carrier. At the same time the adjustment factors should be reevaluated for the future adjustments of the capped levels. This can be done in several ways, as discussed in the text.

Question 8: What is "the likelihood that enhanced carrier efficiency or innovation will result from the possibility of the carrier's retaining, in whole or in part, incremental revenues resulting from these factors"?

Answer: The expectation of such improvements is the raison d'etre of price caps. The price cap level with its prescribed adjustments puts pressure on management performance and gives management more freedom to perform. The extent of the incentive for enhanced productivity is positively related to the time span during which the carrier is free from a general reexamination of the cap.

Question 9: What is "the likelihood that projected increases in carrier innovation and/or efficiency that may result from the possibility of higher earnings would lead to lower prices for the tariffed services from which those earnings would be derived"?

Answer: The likelihood that productivity increases would be passed on in the form of lower prices is very high. First, as part of the initial adjustment formula the carriers, favoring the price cap approach, are likely to commit themselves to higher than otherwise projected productivity increases. This will give consumers an immediate advantage. Second, carriers will be reluctant to make full use of allowed tariff increases in times of high productivity growth and high profits. Third, high rates of productivity growth over a longer time will result in lower caps and higher productivity adjustment at the time of a general reexamination of the price caps.
Question 10: Should we "cap rates for services introduced after regulatory reform became effective"?

Answer: Genuinely new services have four potential properties that weigh against initially capping their rates. First, they may represent an innovation that should be rewarded by the opportunity to make high profits on them for some limited time period. Second, their introduction is risky and may require low introductory prices to induce customers to try them out. Third, almost by definition they usually do not represent vital services for customers. Market power for new goods and services is socially acceptable, as illustrated by the patent laws. Fourth, their quantity weight in the averaging process is zero initially because no quantities were previously sold. At the time of general reevaluation of price caps, that is, after five years or so, new services might well be included.

Although genuinely new services pose few problems, a redefinition of old services as new ones has to be avoided. Two assurances have to be made by the regulator. First, old services have to be continued under capped prices until price caps for new services replacing them are well established. Second, service quality has to be monitored, for example, by responding to consumer complaints.

Question 11: Should "a carrier be able to carry forward unused allowable adjustments for use in later years, as the British plan permits"?

Answer: Yes, otherwise the carrier would fully utilize the adjustments earlier; this would lead to an immediate increase in prices with no future advantage for customers.

Question 12: What is the impact of price cap regulation on current Commission procedures, in particular on "rules and regulations directly related to our rate of return prescription and enforcement" and "our Uniform System of Accounts (USOA)"?
Answer: For any long-run general reexamination of the price cap formula the FCC will need to measure the firm's profitability, assess it in its own right, and compare it with other carriers. Thus, there is some need to retain rules and regulations related to rates of return and some USOA. However, this analysis strongly suggests that only the overall profitability and rate of return of the carrier need to be measured, not those of each service.

The theoretical approach taken in this Note provides some guidance to answering most of the questions posed in the FCC docket. Being based on a theoretical economic model, these answers are valid under particular assumptions which may not hold fully in the real world. Subject to this caveat, the Note provides insights that should be useful for the FCC's deliberations.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>iii</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>v</td>
</tr>
<tr>
<td>Section</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. THE V-F MECHANISM</td>
<td>3</td>
</tr>
<tr>
<td>III. MODIFICATIONS OF THE V-F MECHANISM</td>
<td>7</td>
</tr>
<tr>
<td>A Differentiated Lag Period</td>
<td>7</td>
</tr>
<tr>
<td>Adjustments for Input Price Changes</td>
<td>11</td>
</tr>
<tr>
<td>Adjustments for Productivity Improvements</td>
<td>16</td>
</tr>
<tr>
<td>New Services</td>
<td>22</td>
</tr>
<tr>
<td>The Scope of Price Caps</td>
<td>23</td>
</tr>
<tr>
<td>IV. EXTENSIONS</td>
<td>26</td>
</tr>
<tr>
<td>Appendix: OUTLINE OF PROOFS</td>
<td>29</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>33</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

With its recent *Notice of Proposed Rulemaking* the U.S. Federal Communications Commission (FCC, 1987) has opened a discussion about the possible use of price caps to replace rate-of-return regulation for telecommunications services. The basic idea of these price caps is closely related to the RPI-X formula suggested by Littlechild (1983) for British Telecom and implemented there, along with British Telecom's privatization, in 1984. This formula says that over a period of at least five years British Telecom is constrained in adjusting the prices for a basket of its basic services by the condition that the weighted average of these prices increase by no more than the Retail Price Index less X percent. In the prices of its other (nonbasic) services British Telecom is only constrained by the market and by general laws on competition. Littlechild and the FCC argue carefully about the virtues of such a pricing scheme. In short, it should provide incentives for the regulated firm to produce efficiently, and it should protect customers from monopoly exploitation. Also, the administrative burden of price regulation on regulators and firms is likely to be reduced substantially.

The previous analysis of price caps has at least two shortcomings. First, it concentrates almost solely on the issues of (a) cost efficiency by the regulated firm and (b) equity for certain consumer groups. The issue of allocative price efficiency (that is, correctly dividing up services to the consumers who value them most) has not been raised. This analysis does precisely that and shows at the same time what the implications are for issues a and b. Second, the current literature has examined price caps under a very limited time horizon. This would be acceptable if the regulatory agency could and wanted to commit itself to total deregulation within this time span. Otherwise, one has to ask, what will happen when the time is up for a general review of the price caps? This study makes specific suggestions for answering that question.
The Note argues that the RPI-X or price cap scheme has to be seen from a long-term perspective that includes adjustments of the formula over time. The question of allocative efficiency of the resulting mechanism can and should be addressed in addition to productive efficiency and fairness. This discussion is substantially facilitated by the fact that the RPI-X formula can be interpreted as a modification of the regulatory adjustment process suggested by Vogelsang and Finsinger (1979), hereinafter called V-F.

Section II briefly describes the V-F mechanism and discusses its virtues and problems. Section III deals with modifying the mechanism to make it equivalent to a long-run version of the RPI-X formula. This modification is done in several steps, each of which highlights a particular problem that so far has not been solved for the V-F mechanism. These problems include adjustments to changes in cost and demand conditions and incentives for cost reduction and productivity increases. The main V-F finding continues to hold: Consumer surplus will increase over time, and in the long run the structure of price caps will resemble Ramsey prices. Section IV shows that, with further refinements, constrained optimal two-part tariffs can be reached as well. By using such price caps, all this can be achieved at substantially lower administrative costs to the regulator than by using rate-of-return regulation.
II. THE V-F MECHANISM

Consider a regulated multi-product enterprise in an intertemporal setting. The firm's objective is to myopically maximize profits $\pi_t = p_tq_t(p_t) - C(q_t)$ for each consecutive time period $t$, $t \in \mathbb{N}$, $N = \{1, \ldots, \infty\}$. $p_t$ and $q_t$ are price and output vectors for the $m$ products that the firm produces. The firm faces a stationary cost function $C(q_t)$ and a stationary demand function $q(p_t)$. There are no intertemporal cost and demand effects. The management of the firm is assumed to know both these functions. The regulator is assumed to know neither of them. However, he or she can observe current posted prices and the last period's output quantities and total cost. The regulator is interested in social surplus $S = \pi + V(p)$, where $V(p)$ is consumer surplus without income effects. The regulator would like the firm at least to break even, $\pi \geq 0$. The heart of the V-F mechanism is a constraint $R_t$ that defines the set of allowed prices for the firm in each consecutive period $t$. This constraint is defined as

$$R_t = q_{t-1}p_t - C(q_{t-1}) \leq 0. \quad (1)$$

Constraint (1) says that in period $t$ the firm may only charge prices that would produce no (excess) profit if applied to the last period's outputs and cost. The regulator requires the firm to fulfill all demand at current prices. Figure 1 provides a two-product example of the constraint. It is drawn in price space. The fact that $p_{t-1}$ is inside the zero-profit contour, $\pi = 0$, indicates that the firm has been making a positive profit in $t-1$. Now the shaded area on and below $R_t$ defines the allowed prices in period $t$. As the figure is drawn, the firm can make a profit in period $t$. Maximizing profits, it will end up at a point like $p_t$. Then a new constraint $R_{t+1}$ will be imposed, which will leave the firm less room for profit maximization. The process

---

1The following description is based on Vogelsang (forthcoming).
Fig. 1—The regulatory constraint $R_t$
defines the allowed prices in period $t$. As the figure is drawn, the firm can make a profit in period $t$. Maximizing profits, it will end up at a point like $p_t$. Then a new constraint $R_{t+1}$ will be imposed, which will leave the firm less room for profit maximization. The process continues in this manner and will eventually converge at a point on the zero-profit contour. The key result is that for this final price vector, consumer surplus attains the highest level possible without the firm incurring losses. Prices with this property are called "Ramsey" prices. Because of the gradient property $\partial V/\partial p = -q$ (Roy's identity without income effects), the demand vector $q_{t-1}$ is perpendicular to the social indifference curve $V_{t-1}$ through $p_{t-1}$. Therefore, by construction the constraint $R_t$ is parallel to the social indifference curve through $p_{t-1}$. The constraint always moves in the direction of the steepest increase in consumer welfare.

A simple interpretation of the constraint $R_t$ is that it forces the firm to reduce its prices in each period on average by the last period's profit margin. Here the average is defined by a Laspeyres chain index; that is, prices are weighted by the last period's quantities. As can be seen from Fig. 1, individual prices may well be increased as long as other prices are sufficiently decreased.

Four assumptions on the cost and demand function are needed for the $V$-$F$ process to converge to Ramsey prices.

**Assumption 1:** The cost function has to exhibit decreasing ray average cost. That means that if all outputs of the firm are increased proportionally, then total cost has to go up less than proportionally.

Decreasing ray average cost is closely related, although not identical, to natural monopoly. This assumption is needed to assure nonnegative profits for the firm. The reason is simple. Although the firm may increase individual prices, the pricing constraint requires the firm to reduce its prices on average by the amount of the last period's profit margin. Hence, on average it will have to increase outputs. If that increase in output leads to a less than proportional increase in cost, then the firm can still make a profit. Otherwise it may face a loss.
Assumption 2: The demand function has to be such that consumer surplus is convex in prices.

This property requires that consumer surplus does not decrease if at current prices consumers can still buy the last period's quantities without paying more in total than before. In other words, we apply the axioms of revealed preference to the aggregate of consumers. This assumption assures convergence. Every time the firm makes a profit the constraint is tightened, and every time the constraint is tightened consumer surplus increases. Hence, we get a monotonically increasing sequence in consumer surplus which is bounded by the maximum consumer surplus compatible with nonnegative profits. It will normally converge to this maximum, which, under the assumption of decreasing ray average cost, coincides with the Ramsey optimum.

In our view, the two crucial assumptions just discussed are not very restrictive. They are unlikely to be violated by regulated telecommunications carriers. However, the following two assumptions are restrictive.

Assumption 3: Cost and demand functions faced by the firm are required to be stationary over time.

This stationarity assumption is unlikely to hold in reality, creating a problem for any lagged mechanism. The mechanism therefore has to be adjusted to changes in the firm's environment.

Assumption 4: The firm maximizes profit myopically for every period t.

Sappington (1980) has shown that strategic behavior violating this assumption could undermine the welfare implications of the V-F mechanism even though it eventually converges. The basic argument behind his derivation is that the regulated firm may inflate its costs in early periods in order to postpone more stringent regulation. While V-F regulation always benefits consumers, it has no built-in mechanism that will always induce cost efficiency.
III. MODIFICATIONS OF THE V-F MECHANISM

A DIFFERENTIATED LAG PERIOD

The V-F mechanism has two major weaknesses which are due to Assumptions 3 and 4 (the requirements of a stationary environment and of myopic profit maximization).

Let us start with the problem caused by Assumption 4. Sappington's main finding was that under V-F the regulated firm may strategically waste resources. This may happen in particular at the beginning of the process, as the firm attempts to postpone a tightening of the regulatory constraint. Sappington suggested that this was less likely the more the firm discounted the future: Wasteful expenditures have to be borne by the firm now, while the relaxation of the profit constraint occurs with a lag of one period. The firm's discount rate per period obviously depends on the length of the time period involved. A discount rate of 10 percent for a period of one year becomes 61 percent for a period of five years. Thus, by increasing the regulatory lag period, one can solve the problem of potential waste (and lack of incentives for productivity gains). However, this would also reduce the number of iterations of the regulatory process in a given time span. This could slow down the convergence speed, which would therefore increase the weight of Assumption 3.

To mitigate these problems this study suggests differentiating between two types of adjustments. Long-term adjustments by means of a profit-based reexamination should occur only at multi-year intervals. The relevant interval, or period, is designated by subscript \( t \). Such a general reexamination ought to be done after a prespecified period of five to ten years. The only way an earlier reexamination could be triggered would be if an emergency situation arose for the carrier (e.g., Chapter 11 filing). As a result of the reexamination the average caps would be adjusted once and for all reflecting the actual profitability of the carrier. This Note assumes that the length of \( t \) is external to the regulatory process.\(^1\)

\(^1\)The advantage of flexible intervals is that they add a random element to the process which may be desirable. See Bawa and Sibley (1980) and Logan, Masson, and Reynolds (1986).
Short-term adjustments in any subperiod of t (designated by subscript \( \tau \)) occur at short intervals which we call subperiods. Subperiods can span a year or be irregular. In the latter case the carrier should take the initiative for price changes. The carrier would have to file changes with a demonstration that they comply with the caps. The filing would automatically result in new tariffs after one month unless the FCC or an interested party shows that the new tariffs violate the caps. Again, in the following we assume that the length of subperiod \( \tau \) is set externally. The regulatory constraint now becomes

\[
R^1_{t, \tau} = \frac{1}{T} \delta_{t, \tau} \sum_{\theta=1}^{T} \pi_{t-1, \theta} \beta^\theta (p_{t, \tau-1} - p_{t, \tau}) q_{t, \tau-1} \leq 0. \tag{2}
\]

Here \( T \) is the number of subperiods in each period \( t \), and \( \delta \) is the Kronecker symbol with \( \delta_{t, \tau} = 1 \) for \( \tau = 1 \) and \( \delta_{t, \tau} = 0 \) otherwise. Subscripts \( t, \tau \) represent a lexicographical ordering. Dropping subscript \( \tau \) means that we look at the entire period \( t \). Also, the study defines \( t, 0 \equiv t-1, T \). \( \beta \) is the discount factor applied by the firm and assumed to be known to the regulator. Constraint (2) means that in subperiod \( \tau \) of period \( t \) the firm may set prices which, if applied to subperiod \( \tau-1 \)'s outputs, would reduce gross revenues by the average (excess) profit of period \( t-1 \) discounted to the beginning of period \( t-1 \). Whereas the structure of prices will be adjusted in every subperiod, the level of prices is only adjusted once at the beginning of each period.
If we had only one period \( t \) the term

\[
\sum_{\theta=1}^{T} \pi_{t-1, \theta} \theta^\theta
\]

could be interpreted as a lump-sum amount by which the firm's price level would have to be reduced at the beginning of period \( t \). In all subperiods \( t \) of \( t \) a Laspeyres chain index of prices would have to be held constant. If \( t \) is a period of infinite length, if cost and demand functions of the firm do not change over time, and if the firm maximizes the discounted stream of future profits, then this one-period, infinitely-many-subperiods process will converge to Ramsey prices at the profit level defined by the constraint in the steady state.² The firm will also minimize costs in every subperiod after the first.

Figure 2 shows how the firm moves toward more efficient pricing. In the first subperiod \( \tau=1 \) of period \( t \) the firm faces constraint \( R_{t,1}^1 \) which has been constructed just like constraint \( R_t \) in Fig. 1. As can be seen, the social indifference curve \( V_{t,1}^1 \) reached at the point chosen by the firm will usually cut constraint \( R_{t,1}^1 \). The new constraint \( R_{t,2}^1 \) for subperiod \( \tau=2 \) now has the property that it goes through \( p_{t,1} \) and is tangent to \( V_{t,1}^1 \). Under myopic profit maximization we could now use a revealed preference approach to the firm to determine that for subperiod \( \tau=2 \) the firm would choose a point southeast of \( p_{t,1} \) on the new constraint, though for strategic reasons it may choose not to do so. However, in either case, consumers' surplus weakly increases in every subperiod until the process converges at some profit level higher than the initial profit level \( \pi_{t,1} \). This ultimate profit level may be deemed too high from a social point of view. That is one of the reasons why the RPI-X formula has to be adjusted from time to time.

²A sketch of the proof is given in the appendix.
Fig. 2—The regulatory constraint $R_{t,r}$. 
This analysis does precisely that by assuming that there are many
(nevertheless long) periods t. The major change introduced by this
assumption is that π_{t-1} becomes a variable that is influenced by the
firm's behavior. π_{t-1} depends on both revenues and costs during t-1.
Provided that periods t-1 and t are long enough, the firm will have
little or no incentive to incur excessive costs or charge excessive
prices just to manipulate the constraint in period t.³ On the contrary,
the longer the period and thus the lag, the greater the incentive to
engage in innovative and cost reducing activities. Again, this is the
reason why the RPI-X formula for British Telecom will only be reviewed
after five years. With an infinite number of periods t, with a finite
number T of subperiods τ, and under similar assumptions as before, it
can again be shown that prices converge to Ramsey prices.⁴ This time,
however, they would involve zero excess profits.

Constraint (2) stipulates that the firm has to adjust the price
level for its products at the start of every period but not for every
subperiod. From one subperiod to the next only the base of the chain
index changes, not the level. This is neither in line with the RPI-X
formula nor a sufficient adaptation to a changing environment. To
decide on a procedure for adaptation this analysis has to consider which
potential changes to adapt to. Then several indices are suggested which
have the property that they are easily observable.

ADJUSTMENTS FOR INPUT PRICE CHANGES

Without any adjustment, price caps will eventually lead to some
 crisis situation, either one of financial distress or one of unduly high
profits for the regulated firm. Assuming that one wants to avoid such a
 crisis, the main question remains to what extent adjustments should be

³If the analysis did not discount within this constraint, there
would remain incentives for strategic behavior toward the end of period
t. For example, a cost reducing innovation may be profitably postponed
to the beginning of period t+1. The averaging of profits in period t,
however, could mitigate this problem.

⁴A sketch of the proof is given in the appendix.
automatic and to what extent they should require a formal rate review. Since an automatic mechanism cannot be perfect, price caps will have to be formally revised from time to time. However, automatic adjustments can help stretch the time span between formal reviews. The more an automatic adjustment reflects the actual cost changes of the carrier, the longer this time span can be. However, at the same time, cost accounting problems grow and the incentive to cut costs weakens. To strike a balance, actual cost changes should be used for large items that cannot be influenced by the carrier, such as carrier access charges and taxes. These changes should be given their respective weight in total costs, and the remaining cost changes should be factored in by using industry-wide or nation-wide indices.

Changes in the firm's environment can occur with respect to costs and demands. Costs are the sum of products of input prices and input quantities. Costs may therefore change as a result of price changes for inputs or changes in input requirements.\(^5\)

Let us first concentrate on the question of input price changes. These could become relevant during a period or a subperiod.

To solve this problem for the period, this study suggests allowing cost increases computed for the previous period's input quantities to fully pass through. This would not change Constraint (2), but would mean that, for instance, \(\pi_{t-1} < 0\) would result in a higher allowed price level in period \(t\). Because the regulatory lag (on average) is quite long, the firm would have an incentive to hold input prices down. However, in the interim the firm would not be shielded against input price increases outside its control. To allow for this pass-through of input price increases during subperiods, the regulatory constraint has to be changed. Assuming that there are \(n\) inputs with quantities denoted by the quantity vector \(v_t\) and prices denoted by the price vector \(w_t\), the constraint would be

---

\(^5\)Clearly, changes in the price of inputs trigger changes in input requirements for a cost-minimizing firm.
\[
R^2_{t,r} = \frac{1}{T} \sum_{t=1}^{T} \pi_{t-1} \delta_{t,r} \beta^{\beta} - (p_{t,r-1} - p_{t,r})q_{t,r-1} \\
+ (w_{t,r-1} - w_{t,r})v_{t,r-1} \leq 0.
\]  

(3)

The last term on the left-hand side is a generalization of the notion of fuel adjustment clauses used in regulated energy utilities. It has the advantage over these clauses that no bias is introduced because all inputs in principle are treated symmetrically. It also contains \( R^1_{t,t} \) as the special case where \( v \) is constant over time.

Compared with Constraint (2), the additional information requirement for Constraint (3) is quite substantial. However, regulators normally have access to the firm's accounting data, which should reveal input prices and quantities. Aside from the large amount of data to be handled, the problem of intertemporal cost effects also remains. While inputs of labor and raw materials are usually consumed within one period, capital inputs by definition span several periods. Thus, measuring the correct amount of capital inputs may pose a problem.

The obvious disadvantage of a full input price adjustment is that it substantially lessens incentives for the firm to keep input prices down. That is why the FCC wants to restrict the use of adjustment clauses to just those inputs whose prices the firm cannot influence. If the contribution of these inputs to total costs is high and if they are few, then including them in such an adjustment clause may make both administrative and economic sense. The administrative burden for the firm and the regulator would be small and the flexibility gained would be substantial. If the contribution of these inputs to total costs is small, then additional adjustment indices are needed. These could relate to individual inputs, such as a wage index for construction workers, or to a large composite of inputs, such as a producer price

\footnote{As Eq. (8) below shows, these data may have to be collected and used anyhow to measure productivity increases of the firm.}
index. In these cases again the administrative burden would be low. At the same time, however, there would be the danger that the firm's individual input prices on average would deviate substantially from the index. The more this is likely to be the case the shorter the period $t$ to be chosen.

A realistic adjustment to changing input prices requires a combination of pass-through and indexing in such a way that all or almost all inputs are covered. The resulting price constraint reads

$$R^3_{t,r} = \frac{1}{T} \delta_{t,r} \sum_{\theta=0}^{T} \pi_{t-1,\theta} \beta^\theta - (p_{t,r-1} - p_{t,r}) q_{t,r-1}$$

$$+ (w^*_{t,r-1} - w^*_{t,r}) v_{t,r-1} \leq 0. \quad (4)$$

Here

$$w^*_{t,r} = (w^1_{t,r}, \ldots, w^k_{t,r}, w^{k+1}_{t,r}, \ldots, w^n_{t,r}, \text{PI}_{t,r})$$

and $\text{PI}^j_{t,r}$ is the price index for input $j$ in subperiod $r-1$. $w^j_{t-1}$ is the price of input $j$ in the last subperiod of period $t$. The indices PI could be aggregated for a number of inputs and would be weighted accordingly. Assuming that one of the firm's aggregated inputs is capital we can see that our constraint can also be formulated as a rate-of-return constraint. This can be done by substituting the allowed rate of return, $s_t$, for the cost of capital $r_{t-1}$. Since $r$ is hard to observe we generally will have $s_t$ unequal to $r_{t-1}$. This would make regulation asymmetric with respect to input prices. That is why $s_t$ should only be adjusted once for every period $t$.

Instead of adjusting an initial profit constraint for input price changes, the RPI-X formula used for British Telecom adjusts the allowed price level by the change in consumer price index (RPI in the United Kingdom or CPI in the United States). This procedure holds two basic
advantages. First, the RPI is extremely simple to apply, and its movements are widely understood by the public. Second, it is fair to consumers in the sense that it neutralizes real price changes. At the same time RPI movements are likely to be somewhat in line with movements of input price indices. So, the RPI adjustment also provides some flexibility to the supplying firm. In many regulated industries, however, input price changes have been substantially different from changes in the RPI. Examples include electric utilities and transportation industries heavily dependent on energy inputs.

Similarly, telecommunication costs could be highly sensitive to changes in interest rates, which may move in quite different ways than does the RPI. Somewhat closer to the carrier's input price movements would be the Gross Domestic Product (GDP) deflator and the Producer Price Index (PPI), but again these would not take care of peculiarities of a particular regulated industry. Assuming that one chooses the RPI to allow for price adjustment, this could, as in the case of British Telecom, completely replace any input price adjustment. The relevant constraint in this case would be

\[ R^4_{t, \tau} = \frac{1}{T} \sum_{\tau = 1}^{T} \left( \pi_{t-1, \tau} \beta^\theta \left( \text{RPI}_{t, \tau-1} \text{P}_{t, \tau-1} - \text{P}_{t, \tau} \right) q_{t, \tau-1} \right) \leq 0, \]  

where \( \text{RPI}_{t, \tau-1} \) is the consumer price index for subperiod \( \tau-1 \) based in the last subperiod of period \( t-1 \). We cannot use the current consumer price index because that is only available after the current subperiod is over. Thus, the process is always going to contain some lag that cannot be completely adjusted for.

What can we do if the index used deviates significantly from the actual movement of weighted input prices? One correction is built into the mechanism already. It is the adjustment at the end of each period for excess (otherwise insufficient) profit over that period. This, however, does not undo the index problems of the past. It only shifts the base for the index. Furthermore, if the index has a built-in bias,
the same index problem can surface again and again. To wit, if the RPI systematically understates the average increase in input prices, the firm will see its profits eroded during the course of each period. We therefore have to correct the index itself after each period. To do this, this study suggests one of the following three methods:

First, one may perform a counterfactual experiment by simulating the past with other indices. Then the index (or combination of indices) that best reflects the actual input price movement of the firm is selected.

Second, one can regress the firm's actual weighted input price development during the last period against the index used during that period. If this regression yields strong results, the resulting parameters can be used to adjust the index formula.

The third method combines the first two: Use regression analysis both to identify the best index (or combination of indices) and to project into the future the difference between this index and the actual input price movement experienced by the firm.

These methods are technically straightforward but may arouse public discussion, particularly because they are unfamiliar tools for policy making.

ADJUSTMENTS FOR PRODUCTIVITY IMPROVEMENTS

We now come to the "X" in the RPI-X formula. Aside from input price changes, the firm's costs are influenced by technical changes in the input quantity requirements. Generally, we only expect technical progress to occur. That is, over time we expect to see a reduction in input quantities for a given quantity of outputs. We nevertheless sometimes observe that certain outputs seem to require more inputs now than in the past; there can be three reasons for this. First, the quality of inputs, such as the purity of an ore or the education or motivation of a worker, can have deteriorated. Second, the quality of the output, such as the brakes, acceleration, or durability of a car, can have improved. Third, unmeasured inputs or outputs (externalities) can vary over time. None of these factors can be excluded a priori for the telecommunications sector. However, current indications are that
technical progress, in the form of productivity increases and new outputs, is most likely to continue in this industry. Then the questions arise: How should we adjust our constraint for productivity increases, and how should we adjust for new products?

It has to be noted that under Constraints (2) through (5) the firm is allowed to keep all productivity increases until the beginning of the next period. This lag can generate some excess profit, the expectation of which should boost the incentive to engage in productivity-enhancing activities. After the lag consumers should benefit from the firm's productivity increases. Rather than relying on this lagged adjustment, the RPI-X formula forces the firm ex ante to agree to give away productivity increases to the consumers in the form of price reductions in real terms amounting to X percent per subperiod. This procedure has some drawbacks but also a number of advantages. The main drawback is that the firm carries all the risks of productivity increases. If it cannot reduce its real costs by (approximately) X percent, then it may incur losses. It may also find difficulty engaging in types of R&D that will pay off in terms of cost reductions only much later. The firm will therefore want to negotiate conservatively in setting X. Also, the initial price and profit level of the firm will have to be higher than if no X factor is included. On the other hand, as part of the initial adjustment formula, the carriers, wishing to implement the rate-setting reform, are likely to commit themselves to higher productivity increases than they can safely predict.

This study sees four major advantages in the X percent adjustment of the ceiling for every subperiod. The first is that consumers immediately enjoy increasing benefits for some time (that is, until t+1). Second, carriers will be reluctant to use their allowed tariff increases in times of high productivity growth and high profits.\(^7\) So consumers lose little if X is set too low. Third, high rates of productivity growth over a longer time period will result in lower caps and higher productivity adjustment at the time of a general

\(^7\)The carrier should be able to carry forward such unused allowable adjustments to later years; otherwise the carrier would fully utilize the adjustments earlier with no advantage for customers.
reexamination of the price caps. Fourth, with some good judgment regarding X, period t can be made quite long, mainly because the rate of return achieved by the firm will be within some reasonable range.

There is reason to believe that a predetermined X percent will be too generous initially and may be too tight later. Under the RPI-X approach, at least four potential forces can increase the profitability of the firm. The first is general productivity increases attributed to innovations. The time pattern of such increases is hard to predict. On one hand, the pool of potential innovations decreases over time. On the other, the RPI-X scheme may initially help open up the pool. The second potential force is improvements under current technology that have not been realized because of regulatory distortions or lack of cost minimizing incentives. A key intention behind the RPI-X is to initiate a one-time improvement here. The third force comes from economies of scale that can be realized through growth of demand and lower prices. The RPI-X scheme would, in particular, have the effect of moving the firm down the demand curve. Economies of scale are increasingly exhausted at larger output quantities. So their effect diminishes over time. The fourth force comes from changes in the price structure. To the extent that cross subsidization is reduced (or, more subtly, consumer welfare is increased), the firm can make additional profits which reduce the current tightness of the constraint. Again, these gains are more easily achieved at the beginning of the adjustment process than later. On the other hand, to avoid overshooting, the firm may want to explore demand reactions of consumers in small steps.

There are various possibilities for including projected or negotiated productivity increases in the regulatory constraint. The simplest way within our framework is to insert RPI-X instead of RPI in Constraint (5):

\[
R^S_{t,r} = \frac{1}{T} \delta_{t,r} \sum_{\theta=1}^{T} \pi_{t-1,\theta}^{\theta} - [(RPI_{t,r-1} - X)p_{t,r-1} - p_{t,r}]q_{t,r-1} \leq 0.
\]
This essentially makes us arrive at the Littlechild approach. Rewriting Constraint (6) yields:

\[
\frac{p_{t,r} q_{t,r-1}}{p_{t,r-1} q_{t,r-1}} \leq RPI_{t,r-1} - X - \frac{\delta_{t,r} \sum_{\theta=1}^{T} \pi_{t-1,\theta} \beta^{\theta}}{T \, p_{t,r-1} q_{t,r-1}}. \tag{6'}
\]

Compared with Constraint (6') the RPI-X constraint used by Littlechild would read

\[
\frac{p_{t,r} q_{t,r-1}}{p_{t,r-1} q_{t,r-1}} \leq RPI_{t,r-1} - X. \tag{7}
\]

The difference between Constraints (6') and (7) is that, as a result of our long-term approach, we introduce an adjustment factor as the third term on the right-hand side of Constraint (6'). This adjusts for excess profits (otherwise for insufficient profits) from period t-1 at the beginning of each period t. Such an adjustment is not necessarily required at the time the process is introduced because under the previous cost-of-service regulation the firm may have earned only a normal rate of return. Also note that on the left-hand side, Constraints (6') and (7) use Laspeyres chain index with changing weights for every subperiod t.\(^8\) The reason is that the chain index reflects

\(^8\)In practice revenues seem to be used as weights for this price index. As OFTEL (1986, p. 10) states: "The average is to be calculated by using as weights the revenues that BT reasonably believes to have been received from each relevant class of business." Using revenues as weights for a price index is definitely incorrect from a theoretical point of view, since revenues already contain prices. Revenues and quantities can move in the same or opposite directions. Price indices with revenue weights therefore can have perverse properties. On the other hand, quantities of services may be difficult to measure.
the changed consumption patterns over time. At the same time it shares
with the simple Laspeyres index the characteristic property that it
underestimates (overestimates) welfare improvements (deterioration).
This is an advantage over the Paasche index, which has the additional
drawback that it can only be measured after the period is over.9

As discussed above, the factor "X" could contain various cost and
demand effects such as economies of scale, technical change, increase
in demand, and reductions in cross subsidies.10 The most important
of these influences could be captured in the expected relative change
in total factor productivity (TFP): \( X = \frac{dTFP}{TFP} \). If we use a Laspeyres
index to measure total factor productivity, this can be expressed
conveniently as

\[
X = \frac{P_{t-1}(q_t - q_{t-1})}{P_{t-1}q_{t-1}} - \frac{w_{t-1}(v_t - v_{t-1})}{w_{t-1}v_{t-1}}, \quad \text{or}
\]

\[
X = \frac{P_{t-1}q_t}{P_{t-1}q_{t-1}} - \frac{w_{t-1}v_t}{w_{t-1}v_{t-1}}.
\]

The basis for an adjustment factor reflecting productivity
increases should be the carrier's past performance of total factor
productivity over an extended period of five to ten years. Assuming
that this past performance can be separated for AT&T and the Local
Exchange Carriers (LECs), different future standards should also be set
for the different carriers. Initially, productivity measures should not
be separated into regulated and unregulated services, but rather an
overall value should be used. Exogenous changes in demand cannot be

9 Independently, the problem of these weights is also discussed by
10 Changes in demand are treated in Brennan (1987).
separated accurately from productivity increases, so they should also be combined, given the carrier's responsibility of forecasting demand and installing the correct capacity and correct type of equipment.

To estimate the X factor ex post, the regulator would need to know the same data required to calculate the input price adjustment in the form of Constraint (3) above. This leaves us with several possibilities to set the X in Constraint (6'). First, we could take from Eq. (8) the average relative change in TFP from the last period and set this as the standard for the current period. Second, we could regress actual relative changes in TFP against time in order to establish a trend that is then extrapolated to the future. Third, productivity increases applied to the price cap formula could be negotiated ex ante. Fourth, one could establish a sharing rule for the actual relative change in TFP from each last subperiod $t_{i}, t-1$ to establish $X_{t_{i}, t}$. Such a sharing rule has been used successfully by French public enterprises (Marchand, Pesty, and Tulkens, 1984). Because there is a lag, one may also assume that the full relative change in TFP is passed on to the consumers. To preserve incentives, the lag should then be length $t$ rather than length $t$. In this case substituting Eq. (8) into Constraint (6') and using a Laspeyres input price index instead of the RPI gives us

$$\frac{P_{t_{i}, t} q_{t_{i}, t-1}}{P_{t_{i}, t-1} q_{t_{i}, t-1}} \leq \frac{w_{t_{i}, t} v_{t_{i}, t-1}}{w_{t_{i}, t-1} v_{t_{i}, t-1}} - \frac{P_{t_{i}-2} q_{t_{i}-2}}{P_{t_{i}-2} q_{t_{i}-2}} - \frac{w_{t_{i}-2} v_{t_{i}-2}}{w_{t_{i}-2} v_{t_{i}-2}} - \frac{\delta_{t_{i}, t}^{T} \sum_{t_{i}-1}^{T} \pi_{t_{i}-1, \theta} \delta_{\theta \bar{\theta}}}{T P_{t_{i}, t-1} q_{t_{i}, t-1}}.$$  \hspace{1cm} (9)

<table>
<thead>
<tr>
<th>Output</th>
<th>Input</th>
<th>Output</th>
<th>Input</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Price</td>
<td>Quantity</td>
<td>Quantity</td>
<td>Adjustment</td>
</tr>
<tr>
<td>Index</td>
<td>Index</td>
<td>Index</td>
<td>Index</td>
<td></td>
</tr>
</tbody>
</table>

This constraint can be readily interpreted. The term on the very right gives the average price adjustment done every $T$ years to account for excessive or insufficient profits over this time span.
Leaving out this term, the remainder of Constraint (9) consists only of price and quantity indices for inputs and outputs. It can be interpreted in two ways: Either the sum of an index of output prices and an index of output quantities (lagged) has to be kept smaller than the sum of an index of input prices and an index of input quantities (lagged); or the difference between the index of output prices and the index of input prices has to be smaller than the lagged difference between the index of input quantities and the index of output quantities.

Using Constraint (9) would determine the long-run price caps quasi-automatically. This would be administratively easy. However, it might leave the regulated firms with some risks arising from lags. In an uncertain environment the probability distribution of profits is truncated from above by the price caps. If this leads to losses, the regulated firm can only make up for them through a relaxation of the price caps at a later time. In some way the firm may have to be compensated ex ante for this truncation.

NEW SERVICES

New services potentially have four properties that weigh against capping their rates initially. First, conceptually, genuinely new products are related to productivity increases. Both may require innovative activity represented by R&D. This means that they may have to be rewarded by the opportunity to make supernormal profits on them for some limited time. Second, their introduction is risky and may require low introductory prices which would induce customers to try them out. Third, almost by definition they usually do not represent vital services for customers. Fourth, without any change in the formulas our approach would give new products no quantity weight in the subperiod of

\footnote{If capped prices on average are expected to follow a downward trend over time, the carrier should be required to apply the adjustment factors and recalculate average price caps once a year. More likely is that average caps can be expected to go up in nominal terms. Then the carrier has a self-interest in making the allowed price increases that it needs.}
their introduction: They were not produced in the subperiod before. They could therefore initially be priced at the firm's discretion. In the next subperiod their pricing would be constrained, however. Even if R&D may be treated as a current expense, a subperiod of unconstrained pricing may be too short in general to give enough incentives to the firm for product innovations. Hence, genuinely new products should go uncapped for at least one period.

A danger inherent in this policy toward new products is that the regulated firm might simply repackage old products and present them as new, at the same time abandoning the old service that was capped before. This would not be allowed under the current approach because the firm would still have to serve the demand for the old service at the capped price. To rid itself of an old service, the regulated firm would have to set its price so high that no one would buy it any more. This, however, would usually be prevented by the quantity weight previously attached to the old service. This does not imply an inability of the regulated firm to fool its customers about the nature and quality of its services. Quality control and quality information by the regulators remain essential.

THE SCOPE OF PRICE CAPS

We have so far assumed a multi-product monopoly firm facing a regulatory price constraint on all its outputs. The price cap approach has, however, also been suggested for only a subset of the firm's outputs and for firms operating in a more competitive environment. These suggestions may be complementary to each other. A firm facing competition in some services but being a monopolist in others may need price caps only in its monopolistic services. The author has argued elsewhere (Vogelsang, 1984) that an adjustment process like the one described by Constraint (6) can also be fruitfully applied to single services. This can still lead to an allocatively efficient outcome provided the other (noncapped) services are supplied competitively. In case the regulated firm has market power for the other services, price caps for regulated services may lead to cross-subsidizing the capped services in the long run. The reason is that in the absence of cost
separation among the services, a long-run adjustment of price caps will occur under an overall rate-of-return constraint for the firm. Assuming that the firm can generate above normal returns for its other services, it may (with a lag) be forced to charge noncompensatory rates for the capped services. Politically, this kind of cross subsidization may be quite acceptable. If not, in the absence of fierce competition for the other services some cost allocation will have to be made to separate costs of the capped services. This means that the administrative regulatory burden is likely to be greater if not all services are price capped under the same constraint.

Now, what happens if some of the services under the constraint are offered in competition to other firms while some services are not? The scope of price caps has to be determined by considerations of competition and equity. Competition can force the firm to choose more efficient prices at any time. Under normal circumstances the constraint should give the regulated firm enough flexibility to handle both competition and price caps. Three problems could arise, though.

First, the constraint may not be sufficiently adjusted to the risks of a competitive industry. In part this is a problem of measuring the cost of capital including risk. In part it is the problem of ups and downs in a competitive environment. The constraint may truncate profits on the upside but not on the downside.

Second, in the short and medium term, competition will mean lower prices for competitive services but this may mean higher prices for the monopolized services. This may be unfortunate for the captive customers. There is no long-run alternative short of entry barriers or subsidies, though.

Third, although predation is unlikely to occur under a price cap approach, it cannot be fully ruled out. Using average rates for capped services overall, the regulated carrier might lower its rates in competitive areas below the efficient level and charge fully profit-maximizing rates in monopoly areas with the average complying with the caps. After successfully driving out its competitors, it would reduce the previous monopoly rates and increase the previous predatory rates, again leaving the average within the capped range. The carrier would
not gain an overall price advantage from this. However, there might be an advantage in terms of larger sales because the carrier would not have to share the market with others. The theoretical possibility exists that these additional sales represent enough of a motive for predation. By not including in the overall average the rates of services deemed to be supplied under competition, such an incentive for predation by means of cross subsidization might be reduced. There still remains the problem that the carrier may raise uncapped (or even capped) rates after driving out competitors. However, the initial profit sacrifice will put the firm in a situation of low overall earnings that neither management nor shareholders will like. To be fully on the safe side, predatory acts by regulated carriers may have to be watched just like predatory acts by other firms.
IV. EXTENSIONS

Price caps imposed jointly on all or most of the services of a public utility may pose problems of distributional equity. Under the averaging allowed by joint price caps the firm may choose to increase tariffs for the poor by more than it does tariffs for the rich. There are two ways out of this problem. One is simply to impose prespecified caps on individual services or services to particular consumer groups. This is likely to lead to some reduction in allocative efficiency. Although this may be the price society has to pay for distributional equity, in the telecommunications sector there is a more elegant way out. It has to be noted that two-part tariffs are quite compatible with the price cap approach (Vogelsang, 1987). The way to cap two-part tariffs is to split up a service into the components access and use and to equate quantity of access demanded with the number of customers for the service. In other words, by getting a telephone line a customer would buy one unit of access. This access creates a positive externality for others who can now call this person. Then the fixed part of a two-part tariff can be used to create a lifeline rate for needy customers. To the extent that these customers create positive access (or network) externalities and to the extent that the other customers are inframarginal, this would even enhance allocative efficiency.

Thus, by extending the notion of output quantity from usage to access we are able to move from simple tariffs to two-part tariffs. When doing so the initial price of access would be zero. By splitting up usage into blocks we can generate any type of nonlinear tariff compatible with average price caps. The only condition is that the blocks as quantity weights add up to the original weights before defining the blocks.

A similar technique can also be used to treat a dominant carrier's proposal to restructure the rates of a previously capped service. Restructuring here refers to the pricing of multiple services that can
be combined in different ways. This can be done to the extent that quantitative data on the use of these services in the previous period are available. The tariffs suggested by the carrier for the restructured services would have to be applied to the quantities actually chosen by the carrier's customers in the previous period. Then one calculates the total outlay of customers under these tariffs and compares the result with the outlay under unrestructured tariffs that comply with the price caps. If customers have lower outlays under the restructured tariffs they would actually be better off buying the old quantities. Restructured tariffs having this property should be allowed. Because of the restructuring, consumers will actually choose a different set of quantities than before. By a revealed preference argument, they will only do so because it will make them still better off.

For any long-run general reexamination of the price cap formula the FCC will need to measure a firm's profitability, assess the profitability in its own right, and compare it with that of other carriers. This only has to be done once every five to ten years. Thus, there is some need to retain rules and regulations related to rates of return and some Uniform System of Accounts (USOA). However, according to this analysis only the overall profitability and rate of return of the carrier have to be measured, not those of each service. This means that costs do not have to be assigned in more or less arbitrary ways to individual services. This should substantially reduce dissension over accounting conventions and procedures.
Appendix

OUTLINE OF PROOFS

In this appendix two propositions are stated more formally and their proofs outlined.

Assumptions 1, 2, and 3 are stated in the main part of the Note.

Assumption 4A: The regulated monopoly firm knows its cost and demand functions and maximizes the discounted stream of profits over an infinite horizon and subject to the regulatory constraint defined by Constraint (2). At the beginning of the process profit is positive.

Proposition 1: Under Assumptions 1, 2, 3 and 4A the regulatory adjustment process defined by Constraint (2) will converge for \( t = 1 \) and \( T = \infty \) to Ramsey prices at the profit level achieved in the steady state.

Sketch of proof: The proof is very similar to the one given more formally in Vogelsang (1987). It goes as follows:

Step 1: The discrete-time Euler equations are derived for the firm's constrained maximization problem. It is shown that in the steady state, if it exists, they imply the structure of Ramsey prices.

The firm maximizes

\[
L = \pi_{1,1} + \lambda_{1,1}[\pi_0 - (q_0 - q_{1,1})p_0] \\
+ \sum_{r=2}^{\infty} \left( \pi_{1,r} + \lambda_{1,r}(p_{1,r} - p_{1,r-1})q_{1,r-1} \right) \beta^{r-1}.
\]  

(10)

The first order conditions for this problem are

\[
\frac{\partial L}{\partial p_{1,r}} = \left( p_{1,r} - \frac{\partial C}{\partial q_{1,r}} \right) \frac{\partial q_{1,r}}{\partial p_{1,r}} + q_{1,r}(1 + \lambda_{1,r} - \lambda_{1,r+1}\beta) \\
+ \lambda_{1,r}\beta(p_{1,r+1} - p_{1,r}) \frac{\partial q_{1,r}}{\partial p_{1,r}} = 0
\]

for \( r = 1, \ldots, \infty \)  

(11)
In the steady state, if it exists, all the relevant variables are stationary. Then we can drop time subscripts and Eq. (11) becomes

$$\frac{\partial L}{\partial p} = (p - \frac{\partial c}{\partial q} \frac{\partial q}{\partial p} + q(1 + \lambda - \lambda \beta) = 0 \quad (12)$$

This implies

$$\langle p - \frac{\partial c}{\partial q} \frac{\partial q}{\partial p} = - q(1 + \lambda - \lambda \beta) \quad (13)$$

Noting that $\frac{\partial q}{\partial p}$ is an $n \times n$ matrix, while $p$, $\frac{\partial c}{\partial q}$, and $q$ are $n \times 1$ vectors, we see that Eq. (13) describes the Ramsey price structure.

**Step 2:** By not changing prices at all the firm can always make a profit if it made a profit at the beginning of the process.

**Step 3:** Because of convexity of consumer surplus, consumers are better off after each period in which the firm made no loss: They can still buy what they bought last period without paying more. The firm can gain nothing strategically by making a loss in any period because that does not affect the level of allowed prices for the future. This implies a monotonically increasing sequence of consumer surplus levels. This sequence is bounded by the Ramsey optimum. Hence, it must converge. Step 1 then establishes the result.
Proposition 2: Under Assumptions 1, 2, 3, and 4A the regulatory adjustment process defined by Constraint (2) will converge for $t=1,\ldots,\infty$ and finite $T$ to Ramsey prices with no excess profits ($\pi = 0$) in the steady state.

Sketch of proof: The proof is quite similar to that of Proposition 1.

Step 1: The discrete-time Euler equations to the firm's maximization problem are shown to imply the structure of Ramsey prices in a steady state.

The firm maximizes

$$L = \sum_{t=1}^{\infty} \sum_{r=1}^{T} \left\{ \pi_{t,r} + \lambda_{t,r} \left[ \frac{1}{T} \delta_{t,r} \sum_{\theta=1}^{T} \pi_{t-1,\theta} \right] - (p_{t,r-1} - p_{t,r}) q_{t,r-1} \right\} \beta^{tT+r}.$$  \hfill (14)

The first order conditions to this problem are

$$\frac{\partial L}{\partial p_{t,r}} = \frac{\partial \pi_{t,r}}{\partial p_{t,r}} \left(1 + \frac{\beta^{T+1}}{T}\right) + \lambda_{t,r} q_{t,r-1}$$

$$+ \beta \lambda_{t,r+1} \left\{ \left[p_{t,r+1} - p_{t,r}\right] \frac{\partial q_{t,r}}{\partial p_{t,r}} - q_{t,r} \right\} = 0$$

for all $t = 1,\ldots,\infty$ and $r = 1,\ldots,T$.  \hfill (15)

If the steady state exists Eq. (15) becomes

$$\frac{\partial L}{\partial p} = (p - \frac{\partial \pi_{t,r}}{\partial q} \left[1 + \frac{\beta^{T+1}}{T}\right]) + q \left[\lambda - \beta \lambda + 1 + \frac{\beta^{T+1}}{T}\right] = 0.$$  \hfill (16)
This implies

\[
(p - \frac{\partial C}{\partial q} \frac{\partial q}{\partial p}) = -\frac{\lambda(1 - \beta)T}{T + \beta^{T+1}}.
\]  

(17)

A steady state in this case can only exist if profits vanish because otherwise the constraint could not be stationary. Hence Eq. (17) would imply Ramsey pricing.

**Step 2:** Assuming that the firm made a profit over the previous period, it will have to reduce its prices for the current period in such a way that it would just break even if it sold the old quantities. Actually it will have to sell more on average because of the (on average) lower prices. Because of decreasing average cost this will be profitable to do.

**Step 3:** This step is almost the same as Step 2 above. The difference is that now the firm can gain a relaxation of the profit constraint in the next period by making a cumulative loss in the current period. We cannot totally dismiss this possibility if the firm discounts the future very little. However, the firm would only do this if such a strategy enabled it to make larger profits in the future than it forgoes now. Hence, finite subsequences of profit levels always exist that add up to positive sums. These finite subsequences themselves form an infinite sequence of positive profit levels associated with a monotonically increasing level of consumer surplus. Since positive profit levels are associated with a tighter constraint in the next period, the constraint tightens more and more over time. Hence, profits will eventually go to zero. Again consumer surplus converges to the Ramsey optimum.
REFERENCES


Brennan, T.J., "Capping 'Average' Prices of Regulated Monopoly Firms," Graduate School of Arts and Sciences, George Washington University, Washington, D.C., October 1987.


