Rate of Return Regulation versus Price Cap Regulation:

A Theoretical and Empirical Comparison

by

Qi Yu

(5184846)

Major Paper presented to the

Department of Economics of the University of Ottawa

in partial fulfillment of the requirements of the M.A. Degree

ECO 7997

Supervisor: Professor Gamal Atallah

Ottawa, Ontario

December, 2010
Acknowledgements

I am heartily grateful to my supervisor Gamal Atallah, and second reader Professor Aggey Semenov for their encouragement, guidance, support and helpful feedback and comments on the drafts of this paper.
# Table of Contents

1. Introduction ......................................................................................................................... 1

2. Rate of Return Regulation .................................................................................................... 3
   2.1. Basic rules and formulas ................................................................................................. 3
   2.2. Estimation of rate base under ROR .............................................................................. 5
   2.3. The Averch-Johnson effect ........................................................................................... 6
   2.4. The causes of ROR ........................................................................................................ 7
   2.5. Advantages ...................................................................................................................... 8
   2.6. Disadvantages, solutions and further discussion ........................................................... 9

3. Price Cap Regulation ............................................................................................................ 10
   3.1. Basic rules and formulas ............................................................................................... 10
   3.2. Regulation effects on investment .................................................................................. 11
   3.3. Price review and its frequency .................................................................................... 12
   3.4. Advantages .................................................................................................................... 14
   3.5. Disadvantages ............................................................................................................... 17
      3.5.1. Quality ................................................................................................................... 17
      3.5.2. Cross-subsidization ............................................................................................... 19
      3.5.3. Insufficient information to estimate parameters ..................................................... 21
      3.5.4. Review and adjustment of the X-factor ................................................................. 22

4. Comparison of ROR and PCR ............................................................................................ 23
   4.1. Theoretical analysis ...................................................................................................... 23
   4.2. Empirical evidence ........................................................................................................ 26
   4.3. Alternative or combined regulations ............................................................................. 27

5. Conclusions .......................................................................................................................... 29
   5.1. Increase in demand and capital ..................................................................................... 29
   5.2. Technology improvement and cost saving ...................................................................... 30
   5.3. Asymmetric information between firms and regulators ............................................... 30
   5.4. Product quality .............................................................................................................. 31
   5.5. Possible topics for future research ............................................................................... 31

6. References ........................................................................................................................... 33
Abstract

This paper reviews the most recent literature about rate of return regulation (ROR) and price cap regulation (PCR), analyzes their features, advantages and disadvantages, and compares these two common regulatory regimes based on both theoretical and empirical evidence. The comparison covers the effects on investment, cost reduction, service quality, review frequency, asymmetric information and parameter estimation under both regulation methods. The empirical studies on the effects of transition from ROR to PCR lead to the conclusion that PCR is better or at least not worse than ROR. Besides pure ROR and PCR, alternate or combined regulations are also discussed. The conclusion provides some criteria and concerns to be considered when a regulator needs to choose a better-matched regime for public utilities: increase in demand and capital, technology improvement and cost saving, asymmetric information between firms and regulators, and product quality.
1. Introduction

Industries such as hydro, water supply, gas, railway, and telecommunications are typically considered as natural monopolies, characterized by economies of scale and/or economies of scope. So the social welfare maximizing government should seek a regulation tool to limit the firm's market power, and keep the balance between the incentives for the firm and benefits for consumers. The purpose of such regulations include inducing investment from public utilities, encouraging productivity improvement, achieving a lower administrative cost, and transferring more benefits to consumers.

This paper describes and compares two popular price regulation methods adopted widely: rate of return regulation (ROR) and price cap regulation (PCR). I discuss the concepts, theories and features of the two regulation methods, as well as their advantages and disadvantages. I also reviews the relevant empirical literature, and points out the main factors to be considered in the implementation. Sections 2 and 3 introduce respectively ROR and PCR theories, discuss its features, strength and weakness, and review the related empirical literatures. Section 4 compares ROR and PCR on certain specific elements, and introduces hybrid and alternative regulations. Section 5 summarizes the factors to be considered when applying regulation to a certain industry or economy.

Rate of return regulation (ROR) is a regime to ensure the firm's earnings cover costs if the firm controls cost in a prudent and economical way. At the same time, it prevents the firm from making profit beyond the fair rate of return. This mechanism practically sets a price limit on the regulated bundle of services. Under ROR, the firm has the freedom of choice of production and investment scale.
ROR was introduced in the 1880s to solve the market power problem in railway corporations in the United States. Since it was adopted in the early 1900s widely, there have been many arguments about how to estimate the firm's cost as well as the relationship between the cost and the regulated price. In fact, the argument of this estimation and relationship between the regulator and a firm (Taussig and Pigou, 1913) still has an effect on the points of view today. Different regulators may have different estimations when they are implementing ROR. These varieties come from different features of industries, the time length of regulation, appraisal intervals and the economy's political system. Despite these differences, the principles of the regulation process are the same in general. Berg and Tschirhart (1988) presented the procedures to set the rate by following three steps:

1. Review firms' costs, and deduct unnecessary costs;
2. Ensure the rate is fair for certain firms;
3. Set prices to let revenues exceed costs, and get a fair rate of return.

Using ROR allows the public to avoid excessive monopoly prices. Although there have been some arguments (see discussion in section 2.6), ROR has been adopted as an effective tool to limit the market power of public utilities in many countries for a long period. Under ROR systems, consumers bear more risks of the cost increase than regulated firms, because the firm can seek and be granted a compensation price rise to cover this cost increase at the next price review, which is usually held within one year (Margues, 2010).

Price cap regulation (PCR) is a regime to regulate the price ceiling of a basket of weighted goods or services to limit the market power in natural monopoly industries. Usually the individual good or service in the basket has no price cap, but the weighted bundle has an allowed price cap in total. Littlechild (1983) introduced PCR to protect the consumer welfare during the
privatization of British Telecommunication. Under this regime, the price cap will be reviewed periodically, and price increase between periods will be limited by RPI - X, where RPI represents the Retail Prices Index or CPI in some countries which measures the inflation rate, and X represents a measure of productivity growth between periods. The value of X is decided by the regulator and is an estimation to assess the efficiency improvement, cost reduction, technology or productivity progress for the past period. It means that the periodic price growth rate cannot exceed the X subtracted from RPI. Furthermore, when RPI - X for the next period is less than 0, the price cap of the next period should decrease from the one of the last period. This regulation limits the price, not profit, so that companies have the opportunity to retain all savings by reducing their costs. The cost of this regulation is also lower than rate of return regulation because of its simple information needs. It does not require the estimation of assets, fair rate and cost assignment, thus also lowers the possibility of regulatory capture.

Following the initial success in British Telecommunication, PCR was widely spread to British gas industry (1986), British airports (1987), water and sewer firms (1989-1990), and electric utilities (1990). Now PCR has been adopted in a wide range of both private and public sectors in energy, telecommunication, transportation and water industries (Baldwin and Cave, 1999) in several countries.

2. Rate of Return Regulation

2.1. Basic rules and formulas

ICT Regulation Toolkit (2000) describes the following two steps to implement ROR: “First, determine the economically appropriate revenue requirement. This is based on prudently
incurred expenses and a fair return on invested capital. Second, set prices for individual services so revenue earned from all the regulated services is not greater than the revenue requirement.”

(ICT Regulation Toolkit, 2000, 5.8 Rate of Return Regulation, p. 1)

Firstly, the total revenue requirement comprises the fixed capital costs and variable costs:

$$TR = rK + TVC$$

(1).

The $r$ here is the allowed rate of return to be determined by the regulator. Such a rate decision may consider the cost of capital from the capital market, and the rate is no less than a competitive rate after the commitment of rate insurance.

The $K$ is the rate base to be included into regulated products or services. The regulator may adopt a valuation approach to estimate the plant from the whole assets of the firm to determine the proportion. The debt costs, net value and replacement value may be taken into account. Section 2.2 will discuss this estimation.

Furthermore, $TVC$ is the total variable cost, and can be derived from the following equation:

$$TVC = Operating\ expenses + Depreciation + Taxes$$

(2).

The regulator can obtain this information from the firm’s accounting records, and assign them to regulated products and services.

Secondly, the regulator sets prices based on the given revenue requirement constraint. This constraint requires that the sum of price weighted outputs from all products and services should not exceed the limit of total revenue requirement. It is expressed as follows: (ICT Regulation Toolkit, 2000)

$$\sum_{i=1}^{N} (P_iQ_i) \leq RR$$

(3).

In this formula, $P_i$ is the price of product or service $i$, $Q_i$ is the output amount of product or service $i$, and $RR$ is the allowed revenue requirement limited by the regulator, which can be
expressed as TR in formula (1) and calculated for all the products and services combined. The firm may determine the price and quantity of a certain product or service, but is constrained by its total revenues. However, the firm has freedom to adjust either sides of this inequality.

2.2. Estimation of rate base under ROR

In order to implement the ROR regime, the following values are estimated by the regulator to determine the total revenue from the regulated services according to the formulas (1) in section 2.1: rate base ($K$), required rate ($r$), operating expense, depreciation and tax. These amounts are mainly determined by auditing accounting information from the firm. Among them the rate base is the most sensitive to the calculation assumptions.

To estimate the rate base ($K$), regulators mainly focus on how to evaluate the assets that are invested for the regulated services, usually under Prudence and Used and useful principles (Jamison, 2007). Prudence here means that the investor is conservative enough to assess the values with cost-minimizing concerns and all necessary information. Used and useful is that the asset or property is only used in the regulated services; otherwise the proportion should be calculated. Bonbright (1988) and Phillips (1993) discuss some approaches to measure the capital to be used for the regulated services under ROR, and regulators may use one of the following methods or their combination: (1) Original (or historical) cost method, which is the most common method used in the US. The regulator gathers the purchase amount related to the regulated services from the firm to determine the current asset value directly. (2) Fair (or economic) valuation, which overcomes the defects of the original cost method during an inflation period. The regulator evaluates the assets by the current and market principle based on the asset’s cash flow or market information. (3) Replacement (or current) cost method, which sets up a
virtual manufacturer to establish the same capacity of the regulated services. This approach can avoid disadvantages of the inflation factor and the difficulties to get accounting information, but also requires the market value of the historical investment, and may need more reconciling procedures between the regulator and the firm (Phillips, 1993).

2.3. The Averch-Johnson effect

Averch and Johnson (1962) found that there is always an over-investment tendency to stimulate inefficient capital increase when the firm is regulated by ROR with the assumption of a greater allowed rate of return than required. This tendency is named the Averch-Johnson (AJ) effect, after the two scholars who introduced the model originally. The AJ effect induces a more capital and less labor intensive technology than optimal balance under ROR.

To extend the research on the AJ effect, Baumol and Klevorick (1970) examined some related propositions. In their study, the following hypotheses about the AJ effect were demonstrated as incorrect: (1) the firm employs an optimal proportion of capital and labor to maximize profits and generates more outputs under ROR than it produces without regulation. (2) The capital amount invested by the firm is a negative function of the gap between the allowed rate of return and the capital’s market rate price. Furthermore, Zhao (2001) clarifies two different ROR types under uncertainty: regulating the expected rate of return where AJ effects always occur, and regulating the rate of return where AJ effects may not occur.

One of the recent proofs of the AJ effect under ROR was presented by Carrera-Gomez (2005) with a case study of ROR in the Spanish port sector. This paper estimated the overinvestment in Spanish ports between 1985 and 1997. Two estimation methods were used and discussed: cost function and input distance function, and the latter one is supposed to be better to explain
technological progress. The results provide empirical evidence of over-capitalization under ROR, suggesting the existence of AJ effects.

2.4. The causes of ROR

ROR is “a development of an incomplete long term contractual relationship that offered utilities a fair rate of return in exchange for the ability to renegotiate the terms of the contract without excessively costly haggling” (Williamson, 1985, p.347). The interval between the rate reviews is usually one year for ROR. The hearings reduce the information asymmetry between the firm and consumers, and limit the monopoly power, because the firm’s cost information is forced to be audited, and the consumers’ price and quality concerns can be considered by the regulator.

Some studies (Baron and Taggart, 1980; Besanko, 1985) showed that asymmetric information was the reason for the emergence of ROR, because ROR may make the regulator overcome the insufficient information about the firm’s labor input, and observe the firm’s capital choices instead. To generate an objective function that maximizes social or consumer surplus, Baron and Taggart (1980) found that the regulator adopted ROR because of the difficulties to get an exact cost function from the firm. To further this observation, Besanko (1985) introduced a model to describe how the regulator develops a regulatory regime when the real labor input is hard to get from the firm, and this model explains that regulators maintain ROR to overcome asymmetric information.

However, Evans and Garber (1988) found that ROR is the result of interest group equilibrium. They designed an objective function to combine the pressures from interest groups like consumers and firms, and proved the emergence of ROR. According to their model, the
regulator who wants to optimize the objective function will impose ROR to trade-off the low price for consumers and fair rate of return for the firm. They further demonstrated that ROR based on this objective function has AJ effects.

2.5. Advantages

Firstly, ROR provides a means for overcoming monopoly pricing. Although theoretically the regulator monitors the cost and controls the profit, practically determines the price directly. Joskow put it this way: “the regulatory agency sets the prices of services and not rates of return. Once the prices for utility services are set by the regulatory authority, they remain at fixed levels until they are officially increased or decreased by action of that regulatory authority” (Joskow, 1973, p.119), thus the common trend of price increases from the firm is limited.

Secondly, ROR gives the regulated firm an insurance of earning level from its investment, and this can be important in stimulating investment. As a result, it encourages investment in the industries that are short of capital. For example, Hausman and Neufeld (2002) analyzed how the regulator adopts ROR to induce capital entry to the US electric utilities in early 1900s. They observed a panel of data during 1910-1919 from firm bond market records, and developed a model to compare the borrowing costs between the States with and without ROR. The results indicate lower borrowing costs in the States with ROR, which implies the success of ROR by encouraging investments, so this study demonstrated that ROR makes the target capital market more competitive than non-regulated ones by insuring the rate of return. On the other hand, compared with the States without ROR, consumers enjoy higher output in States with ROR.

The third advantage of ROR is the high service quality. Although the AJ effect induces an overcapitalization tendency, it usually brings an extra investment in production technology and
product quality, which ensure higher service quality, maybe even higher than the profit-maximizing level. As Yu (2003) concluded, some industries caring about quality more than price may adopt ROR with this feature. For example, United States Department of Defense uses ROR to ensure quality of the national defense products.

Liston (1993) pointed out some other advantages of ROR. One advantage is that the regulator may attain second-best prices based on marginal cost and demand, and another advantage is that ROR may satisfy some political or social needs to benefit some special consumers by adding universal obligation with the permission of cross-subsidization. He also noticed another potential benefit of ROR, that the yearly review builds a platform to keep balance between different interest groups like consumers and the firm, thus limiting the firm’s market power.

2.6. Disadvantages, solutions and further discussion

The first disadvantage of ROR is the poor incentives for cost reduction. The profit-maximizing firm is encouraged to invest more in capital, not to cut unnecessary costs and become more efficient. However, practically the small gap between rate reviews gives ROR some incentives for cost reduction if the latest change from cost reduction cannot be used for the next review on time (Baumol and Klevorick, 1970).

The second disadvantage of ROR is the AJ effect which induces inefficient overcapitalization when the allowed rate of return exceeds the cost of capital. However, Gilbert and Newbery (1994) reported that ROR combined with a used-and-useful rule can provide an efficient approach to investment. On the contrary, Egert (2009) held that the used-and-useful rule has negative effects on the investment choice because of its uncertainty and unclear review process, which result in distortion of investment. Evans and Guthrie (2006) also indicate the weakness of the optimized
replacement costs method with used-and-useful rule, because the regulator ignores demand uncertainty. This uncertainty exclusion underestimates the potential loss of excess capacity when the demand decreases, and leads to overcapitalization. On the other hand, Evans and Guthrie (2005, 2006) identify Replacement Cost Valuation (a method to estimate the rate base under ROR) as a risk borne by the firm, which will take the loss if the sunk investment was devaluated by the current market but the value difference cannot be included into costs by the regulator. Such a risk limits the incentives for investment. Therefore, Vogelsang (2010) concludes that ROR's effects on investment depend on the way it practically works, such as how the regulator determines the rate base and the review interval.

The third disadvantage of ROR is difficulties in determining the fair rate of return. The firm may "capture" the administrative agency to be allowed to include unnecessary costs, or the regulator may set too tight price constraints below real costs by underestimation of the risks (Liston, 1993).

"Regulatory lag" between hearings may be a disadvantage of ROR, because transferring cost changes to allowed rate changes is always delayed by a rate review process, thus the regulation's commitment target cannot be reached perfectly. However, the regulatory lag can motivate the firm to minimize the costs and weaken the tendency of overinvestment (Kahn, 1988; Johnson, 1989).

3. Price Cap Regulation

3.1. Basic rules and formulas

The following is one of the typical rules and formulas of price cap regulation:

$$P_t = P_{t-1} [1 + (RPI - X)]$$  \hspace{1cm} (4)
The price cap is reviewed periodically here. $P_t$ is the regulated price for the current review period, $P_{t-1}$ is the regulated price for the last review period, RPI is the Retail Prices Index (or CPI in some countries) which measures the inflation rate from period $t-1$ to $t$, and the value of $X$ decided by the regulator is an estimation to assess the efficiency improvement, cost reduction, technology or productivity progress from period $t-1$ to $t$. For example, if the RPI increased by 4% in the last review period, and productivity increased 2% in the same period, the company can increase price by no more than 2% in the next review period. For the regulated firms who provide multiple products, the price cap is usually the maximum price of a “basket” weighted by a set of products and services.

3.2. Regulation effects on investment

PCR has both positive and negative effects on incentives for investment. As an example of a positive effect, Carbral and Riordan (1989) illustrated that PCR promote cost-reducing investment if the price ceiling is not set too low. Buehler, Burger and Ferstl (2010) also proved the incentives of increasing capacity investment by more aggregate output under PCR. The most negative concerns of investment decisions under PCR are from the uncertainties, because the firms bear the risks of demand decrease, sunk costs, technology revolution and regulator’s commitment. Theoretically the same price level as the marginal cost of capital is not enough to compensate for such risks when the firm makes decision of investments under PCR, therefore the regulator may need a higher level of price caps to cover an insurance premium.

Some studies on the relationship between regulation and investment focus on these uncertainties. Lehman and Weisman (2000) revealed that in comparison with their ROR counterparts, PCR regulators tend to attract more firms to enter the industry to induce
competition. This tendency may weaken the firm's confidence and the performance of PCR accordingly, and reduce the incentives for investment. They suggested that "the performance of future price cap plans will depend not only on limiting the market power of the regulated firm, but on limiting the regulator's incentives for excessive market entry as well" (Lehman and Weisman, 2000, p.355). By considering a "discrete poisson arrival process" (Biglaiser and Riordan, 2000, p. 750) and the "option value of waiting to invest" (Biglaiser and Riordan, 2000, p. 751), Biglaiser and Riordan (2000) studied how the optimal output price is affected by uncertainty of technological progress and demand uncertainty, and concluded that the optimal price to maximize the social welfare depends much on these two variables.

Rogues and Savva (2009) observed the impact of demand uncertainty on investment under price caps, and found that a price ceiling can reduce incentives for underinvestment because the firm may decide to produce more than the monopoly level. On the other hand, they also concluded that PCR makes the firm overly careful to unnecessarily postpone some valuable projects, and that should have a negative effect on investment. According to their paper, PCR effects on investment heavily depend on the optimal price cap estimation by the regulator: "a relatively high price cap can speed up investment compared to an unregulated industry, while a stringent price cap will act as a disincentive for investment" (Rogues and Savva, 2009, p.520). Furthermore, the failure of this estimation has asymmetric consequences, with tighter price caps having more severe effects than laxer ones.

3.3. Price review and its frequency

Periodic reviews usually take place to transfer more cost reduction benefits from firms to customers under PCR. When necessary, not only price caps but also the regulation itself can be
adjusted. Although theoretically the regulator proposes an X factor update from the estimation of the whole industry’s efficiency improvement, it is hard to avoid using the firm’s cost information.

Some essays analyzed the firm’s behavior by the periodic pattern under PCR regimes, and found that the regulated company spent most efforts to cut costs at the beginning of a regulatory period after a recent price review, and these cost reduction activities were rare when it was close to the next review. This negative trend is one form of the ratchet effect (Laffont and Tirole, 1993). Baltagi and Griffin (1988) introduce a model (Generalized Index of Technical Change) to measure and analyze time series of the cost reduction activity. Applying Baltagi and Griffin’s model to a sample of data from English and Welsh Water and Sewerage companies during their two regulatory cycles (1995-1999 and 2000-2004), Bottasso and Conti (2009) estimated a cost function to provide evidence that the firm has increasing tendency to save cost in the early phases of the regulatory period, while the incentives are lower in the latter period. Tella and Dyck (2008) also found evidence consistent with this strategic behavior by firms. According to their paper, the cost function within the four-year regulatory cycle is U-shaped, which indicates early cost reductions and then reversal in the second half of the period. This result explains how firms try to revise the information provided to the regulator to gain a better commitment for the next regulatory period under PCR, because the next round of X-Factor will be affected by the estimation of recent productivity.

Another important feature of PCR review is how frequently the value X should be reassessed. The firm is promised to keep all extra profits from cost reduction between two reviews, thus creating a “regulatory lag”. While a longer lag gives the firms a room to gain extra profits from their productivity enhancement soon after the commitment is effective, consumers have no chance to benefit from this improvement before the next review. On the other hand, a shorter lag
raises consumer welfare but degrades the incentive for the firm to be more efficient because it may induce a higher X shortly in the next review and decrease future profits. Evans and Guthrie (2006) found that the price review interval affects the estimation of output price and cost of capital, so PCR causes more risks about irreversible investment concerns than ROR when more frequent reviews occur. Besides the benefit allocation trade-off, the administrative cost of price reviews is one more concern for the regulator to determine how often the rate hearing should be scheduled. Empirically the interval is usually set as four to five years in most countries where PCR is adopted.

3.4. Advantages

Price cap regulation has stimulating effects on firms in the following ways. The first advantage of PCR is that firms are encouraged to reduce costs and keep all they can save from efficient behavior. This is why PCR was introduced and regarded as “incentive regulation”, and such a regime has the potential of enhancing productivity. Empirical evidence was provided by Mathios and Rogers (1989), who reviewed the price of long distance services which were provided by AT&T, and found that the price level in the States adopting PCR was lower than that in the states adopting ROR. Schmalensee and Rohlfs (1992) concluded that AT&T gained an extra benefit of 1.8 billion dollars from PCR during the period of 1989-1991 compared with ROR in 1986-1988, while consumer surplus increased by 20 billion dollars during the same period, thus indicating that over 90% of gains from PCR were transferred to consumers. Erbetta and Cave (2007), using DEA (Data Envelopment Analysis) also found improvements in technical and allocative efficiency within the sample water and sewage utilities under PCR.
supervised by Office of Water Services from 1993 to 2005, and identified a significant improvement since 1999 price review.

Coco and De Vincenti (2008) provided a theoretical demonstration that PCR may induce cost-reducing efforts by its periodic rate base reviews. A dynamic model was set up to explain the relationship between productive efficiency improvement and the length of the concession period. The longer the concession period is, the more the benefits that price reviews can generate to increase productive efficiency. Furthermore, this paper explores the calculation of the optimal frequency of price reviews to maximize the pure cost-efficiency outcome, and concluded that the optimal solution negatively depends on the following two variables: the slope of the demand curve and the intensity of the disutility of effort.

However, Aubert and Reynaud (2005) challenged the assumption of PCR's advantage in cost reduction by observing a sample of 211 water utilities in the same State (Wisconsin) and for the same period of 1998-2000 but under regimes of ROR, PCR and hybrid schemes respectively. The interesting result is that the most efficient group is ROR. On the contrary, the hybrid mechanism which combines ROR and less information requirements gets the worst efficiency assessment by their stochastic cost frontier approach. Bottasso and Conti (2003) found cost reduction from the water industry in English and Welsh area during the emergence of PCR between 1995 and 2001, but argued that the reason was mainly because of capital market competition since 1994 rather than PCR itself.

The second advantage of PCR is its incentives for investment efficiency. Regulators assure the price ceilings, but do not guarantee the return on investment, thus firms should investigate the market risks, technical capacity and cost of capital to decide the investment amount. The firm should consider the uncertainties of demand and cost expectation together with the price ceiling
promise, and make investment choices carefully. This concern reduces the incentive of overinvestment under ROR (AJ effect) and results in relatively efficient investments. On the other hand, the potential underinvestment strategy for the oligopolistic firm to raise price is also limited by price caps. Rogues and Savva (2009) compared the model of investment under uncertainty with and without PCR, and concluded for the existence of an optimal price ceiling for the regulator to maximize the investment incentives when the demand is not changing too much. Buehler, Burger and Ferstl (2010) also find that PCR may increase aggregate investment when the price ceiling is more than the marginal cost but less than the peak price if the firm is efficient enough, despite that the price level is lower than the monopoly level. As discussed above, the benefit of inducing efficient investment depends on appropriate price ceiling settings under PCR.

Liston (1993) also identifies another advantage of PCR: that the scheme reduces the monitoring cost of regulators compared with ROR. The regulator may have the ability to gain the information required by a price review, rather than the private knowledge of firm’s cost information, thus avoiding ROR’s distortions and extra administrative cost. Consequently it is easier and cheaper to estimate an industry’s productivity increase and CPI for PCR than to investigate every company’s accounting records for ROR. The longer frequency setting is another reason why PCR is superior to ROR on the aspect of administrative costs (usually four to five years vs. one year).
3.5. Disadvantages

3.5.1. Quality

Under PCR, in order to pursue higher profits, a firm may try to reduce costs through service quality reduction. This is in line with some of the empirical studies. For example, Armstrong, Cowan, and Vickers (1994) observed the service quality of British Telecom during 1985-1989, and found that service quality was declining. Alexander (2001) noted that reduced service quality does exist for Ameritech. In fact, service quality degradation occurs in particular when the delivered service quality cannot be easily assessed by consumers.

To explore an optimal combination of price and quality under PCR, Currier (2006) introduced a price cap system which generates the optimal choice of price and quality using a Laspeyres index approach and related graphical analysis, and concluded for the possibility of loss of consumer surplus. According to the study, the quality-adjusted Laspeyres price caps result in lower welfare than a non-regulated regime, but this negative effect could be reduced by encouraging market access by low-cost firms or discouraging market access by high-cost firms.

However, empirical evidence suggests that lower quality is not an inevitable consequence of PCR. Tardiff and Taylor (1993) studied the former Bell companies and observed that the performance was not affected by PCR, so their study did not support the hypothesis of reduced service quality under PCR. Furthermore, using 12 subscales to measure telephone service performance under PCR, Banerjee (2003) found not worse or even better service quality via a Granger-causality test with multi-dimensional data on 49 local exchange carriers during the period 1991 – 1999.
Although there is conflicting evidence supporting the opposite result, degradation of service quality - as a result of incentives for cost reduction - is still a potential disadvantage of PCR (Sheshinski, 1976; Sappington, 2003, 2005; Weisman, 2005).

In order to solve this problem, price regulation must be implemented with additional quality standards. As Kahn (1988, p.22) noted, “Price has no meaning except in terms of an assumed quality of service … price regulation alone is economically meaningless”. Reviewing certain quality requirement regulations together with price regulation, Milne (2003) concluded that there are two approaches to incorporate service quality into PCR: encouragement and enforcement. Encouragement could include publication of performance statistics and reports about completion of performance targets. Enforcement usually includes penalties or cash payments to customers when the firm fails to achieve the required performance or quality standard. Weisman’s analysis (2005) revealed that the lower price caps result in more incentives to reduce investment in quality, and these incentives may be adjusted by multi-market participation of the regulated firms. To consider a penalty regime, while revenue-share penalties (the firm gets fined for poor quality based on its revenue-share amount) can degrade the incentives for investment in service quality, profit-share penalties (the firm gets fined for poor service quality based on its profit-share amount) generate incentives for increased investment in service quality, and more requirements to expose the regulated firm’s quality standard also encourage investment in service quality.

One solution is to incorporate a Q-Factor (Currier, 2007), which is a measurement of service quality, into the price cap constraints. The Q-Factor can adjust the allowed price to fall when the service quality is reduced. That is to say, the firm is selling higher quality services at higher prices, and alternatively, buying off consumers to have the opportunity to obtain lower prices by
accepting lower quality services. So the Q-Factor should correctly represent consumers' tradeoff between price and service.

Currier (2007) discussed the Q-Factor as well as its selection process. In addition, although the Q-Factor indicates consumers' tradeoff between price and service, policy makers cannot actually use this tradeoff in policy making because they lack sufficient theoretical guidance. To address this issue, he proposed a modified Laspeyres price cap for regulators to better capitalize the tradeoff. It is a quality corrected price, and doesn't require a large amount of information.

3.5.2. Cross-subsidization

Under PCR, firms producing multi-products or dealing with different product markets may be able to cross-subsidize their products. Under ROR, the regulator can set regulatory prices based on the individually calculated cost of each product or service that the regulated firm provides. Thus to a certain degree, cross subsidization still exists under ROR, but it is the result of a balancing equilibrium (we obtain a second best result due to information costs and transaction costs). With PCR, however, because the regulated firm has more freedom in determining the price for each of its products or services, it can develop cross-subsidizing pricing strategies by setting high prices in the market where it has monopoly power, and low prices in competitive markets (Baumol and Blinder, 2008). In general, there are two types of cross-subsidies: cross-subsidization across different products or services, and cross-subsidization among different consumers.

Cross-subsidization across different products or services can be reflected in the rate of return as well as in its variation for different products or services. According to table 1, when the Federal Communications Commission (FCC) first introduced PCR in their interstate service in
1990, the average rate of return was 12.9% across the seven telecommunication companies. The individual rate of return, which is between 11% and 14%, is not significantly different among the different services of common line, switching, transport and special access. However, this changes a lot in 1993 after the implementation of PCR for four years. Although the total rate of return of the seven companies did not change much (increasing from 12.9% to 14.0%), the individual rate of return varies a lot across different services. Services in the competitive market such as common line, transporting and special access have a slower growth or even negative growth, while the monopoly service of switching has on average a rapid growth in its rate of return from 12.9% to 30.8%. At NYNEX, their rate of return of the switching service increases from 13.1% to 47.5%, as compared to a drop in the special access service from 9.1% to 5.7%. Apparently, these variations cannot be explained by factors such as external cost variation or productivity growth differences. The only reasonable explanation is that the telecommunication companies take the advantage of the PCR over a basket of services to carry out the cross-subsidization strategy.

**Table 1 - Selected Percent Returns on Interstate Services**

<table>
<thead>
<tr>
<th></th>
<th>Interstate Total</th>
<th>Common Line</th>
<th>Switching</th>
<th>Transport</th>
<th>Special Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ameritech</td>
<td>14.5 15.4</td>
<td>12.4 12.7</td>
<td>12.0 21.7</td>
<td>11.7 14.0</td>
<td>16.3 12.1</td>
</tr>
<tr>
<td>Bell Atlantic</td>
<td>14.4 15.1</td>
<td>10.5 13.2</td>
<td>13.0 21.9</td>
<td>12.7 10.5</td>
<td>12.3 10.1</td>
</tr>
<tr>
<td>Bell South</td>
<td>13.2 13.9</td>
<td>11.4 10.1</td>
<td>14.0 34.0</td>
<td>13.3 11.8</td>
<td>11.2 16.2</td>
</tr>
<tr>
<td>NYNEX</td>
<td>10.8 13.6</td>
<td>8.2 8.1</td>
<td>13.1 47.5</td>
<td>11.3 12.6</td>
<td>9.1 5.7</td>
</tr>
<tr>
<td>Pacific Bell</td>
<td>12.7 12.8</td>
<td>11.1 11.0</td>
<td>13.6 32.7</td>
<td>13.8 6.4</td>
<td>12.6 11.9</td>
</tr>
<tr>
<td>Southwestern Bell</td>
<td>11.2 13.1</td>
<td>11.2 9.5</td>
<td>12.5 33.8</td>
<td>7.2 14.4</td>
<td>11.1 12.4</td>
</tr>
<tr>
<td>US West</td>
<td>13.7 14.2</td>
<td>12.5 9.9</td>
<td>12.4 24.1</td>
<td>13.2 14.5</td>
<td>10.1 12.1</td>
</tr>
</tbody>
</table>

Source: Loube (1995)
With asymmetric cost information, cross subsidization among different consumers can be reflected in the actual price changes. Take the telecommunication industry in Britain for example, the varying actual X value across different consumer groups in 1991-1996 clearly indicates this. In that industry, the average X value of different consumer groups is 6.6%. However, the X value is 9.3% for commercial consumers, and only 4.2% for residential consumers. In addition, among all the residential consumers, the X value is 5.7% for those 20% with higher usage, but 2.7% for those 80% with lower usage (Wang, 1998). Therefore, under PCR on a bundle of products or services, the residential consumers cannot be effectively protected. In the long run, if a service has a lower initial price index than that of the other services, as time goes by, the price variance across different services will become larger and larger, and deviate more from their optimal values.

3.5.3. Insufficient information to estimate parameters

Both underestimation and overestimation of the X factor have negative effects on PCR, but it is difficult to obtain an optimal price ceiling. Liston (1993) pointed out that the regulator faces a trade-off between firms and consumers: lax price cap constraints give a firm great incentive to reduce costs, but weaken the welfare and benefit of the regulation. In contrast, low price caps do not give a firm enough room to compensate their uncertainty concerns, thus induce underinvestment due to possible high risks about cost fluctuations. Rogues and Savva (2009) extended the trade-off analysis and found that underestimation of X-factor is more dangerous than overestimation.
Insufficient information is another problem when selecting the value X. Schmalensee (1989) outlined that reliance on actual cost information weakens the incentive effects of PCR. Kiss (1991) used econometric method to estimate the coefficients of the production or cost function. Based on the estimated production or cost function, he then estimated the X-factor. Liston (1993) found it difficult to establish a production or cost function for developing economies or a fast growing industry to apply the econometrics method discussed by Kiss (1991). The indexing approach is the most popular method to measure the productivity and its improvement by comparing growth rates of outputs and inputs. For example, Roycroft (1999) used total factor productivity (TFP) to evaluate efficiency enhancement for Ameritech companies, and Bernstein (2006) suggested an X-factor for the Peruvian telecommunication industry by TFP calculation. More or less, the X-factor’s dependence on cost information (private knowledge of the firm) limits the accuracy of estimation, thus weakens the incentives or welfare targets of PCR.

3.5.4. Review and adjustment of the X-factor

As discussed in Section 3.3, the value X is reviewed and assessed mainly on the basis of the efficiency information from the firm. So the firm has an incentive to influence the price review by providing distorted cost information. Although the firm has incentive to reduce cost immediately at the beginning of a period, the expectation of next review results in cost increase and efficiency reduction toward the end of the period as mentioned before. Pint (1992) proved the previous year’s inefficiencies in his study of the UK utilities. Therefore, firms are discouraged to lower the current costs which will be used as the base for the next assessment, and this discounts the benefits of PCR. Longer lag between price reviews might be the most important strategy that makes PCR superior to ROR, and might be the main reason that results in
PCR’s incentive effects. Kahn (1988) argued that ROR can also have the same incentive features as PCR if the review interval is set long enough. Ergas and Small (2001) summarized the relationship in this way: “The longer is the tenure of the cap, the greater incentive the firm has to cut costs immediately”. (Ergas and Small, 2001, p.13)

4. Comparison of ROR and PCR

4.1. Theoretical analysis

Based on the discussion in Sections 2 and 3, the following table summarizes the theoretical differences between the two regulatory approaches:

<table>
<thead>
<tr>
<th>Table 2 – Comparison of ROR and PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROR</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Investment</td>
</tr>
<tr>
<td>Cost reduction</td>
</tr>
<tr>
<td>Service quality</td>
</tr>
<tr>
<td>Information required and estimated</td>
</tr>
<tr>
<td>Review frequency</td>
</tr>
</tbody>
</table>

Ergas and Small (2001) summarized the theoretical relationships between ROR and PC in Figure 1. The horizontal axis indicates the extent of insured earnings ranging from zero to 100%, which is the amount of return guaranteed by the regulating authority. In practice, ROR is
regarded as a high-level insurance regime, and PCR is regarded as an incentive regulation with low-level insurance.

**Figure 1**
Comparing Regulatory Regimes

![Diagram showing the comparison between Pure Incentive Regulation and Pure Rate of Return](image)

Source: Ergas and Small (2001)

However, the practicalities of PCR usually share the essential factors of ROR, as Johnson (1992, p. 40) observed: “price-cap regulation can best be regarded as a loose form of rate-of-return regulation with a formal time lag. Price-cap regimes typically include a periodic review of performance (including the historic rate of return) and an adjustment in the formula to bring the projected rate of return in line with what regulators would regard as just and reasonable.” Bergman (1998) also pointed out that in practice, ROR and PCR are not too dissimilar because the X factor decision must consider the regulated firm’s rate of return as an important factor.

To compare the weakness of ROR and PCR in privatized utility in developing and transition economies, Kirkpatrick, Parker and Zhang (2005) conducted a questionnaire study on a sample
of developing and transition countries. Their results reported the difficulties that regulators faced under each regime. The issues are illustrated in Table 3.

**Table 3 - Comparison of the reported difficulties faced when operating price caps and rate of return regulation in developing and transition economies**

<table>
<thead>
<tr>
<th>Difficulties</th>
<th>PCR</th>
<th>ROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information asymmetries or inadequate information on the firm’s cost and revenues</td>
<td>23 (96%)</td>
<td>10 (59%)</td>
</tr>
<tr>
<td>Enterprises providing misleading information</td>
<td>14 (58%)</td>
<td>8 (47%)</td>
</tr>
<tr>
<td>Serious levels of customer complaints about rising prices</td>
<td>17 (71%)</td>
<td>8 (47%)</td>
</tr>
<tr>
<td>Enterprises earning excessive profits</td>
<td>4 (17%)</td>
<td>4 (24%)</td>
</tr>
<tr>
<td>Enterprises over-recruiting labor</td>
<td>4 (17%)</td>
<td>3 (18%)</td>
</tr>
<tr>
<td>Enterprises under-recruiting labor</td>
<td>2 (8%)</td>
<td>3 (18%)</td>
</tr>
<tr>
<td>Enterprises over investing in capital equipment</td>
<td>4 (17%)</td>
<td>9 (53%)</td>
</tr>
<tr>
<td>Enterprises under investing in capital equipment</td>
<td>10 (42%)</td>
<td>2 (12%)</td>
</tr>
<tr>
<td>Excessive rises in the pay of senior management</td>
<td>3 (13%)</td>
<td>6 (35%)</td>
</tr>
<tr>
<td>Problems with quality of service</td>
<td>12 (50%)</td>
<td>6 (35%)</td>
</tr>
<tr>
<td>Inability to recruit staff skilled in the management of regulation</td>
<td>4 (17%)</td>
<td>3 (18%)</td>
</tr>
<tr>
<td>Political pressures, for example ministerial intervention in setting prices</td>
<td>15 (63%)</td>
<td>7 (41%)</td>
</tr>
<tr>
<td><strong>Total number of difficulties reported</strong></td>
<td>112</td>
<td>69</td>
</tr>
</tbody>
</table>

Source: Kirkpatrick, Parker and Zhang (2005)

The collected questionnaire was consistent with traditional views of ROR and PCR. ROR has problems with over-investment, complaints from customers about price, and asymmetric information between the regulator and firms, while PCR has problems with under-investment, complaints from customers about prices rising, and lower service quality. Moreover, the survey results indicate that PCR regulators in developing and transition economies encounter more problems due to their fast changing institutional environments and inadequate administrative experiences.
4.2. Empirical evidence

The empirical studies on the effects of transition from ROR to PCR lead to the conclusion that PCR is better or at least not worse than ROR based on efficiency criteria. Roycroft (1999) used an approach of total factor productivity (TFP) growth introduced by FCC to calculate the X-factor, and obtained about 4.5% productivity growth during the ROR to PCR transition in the Ameritech operating companies through 1997, while historical productivity growth under ROR has been typically estimated as 1% or less by a regulatory commission. The regression analysis results indicated that the introduction of PCR leads to statistically significant increases in productivity compared to ROR.

Resende (2000) assessed the efficiency scores of ROR and PCR for US local telephony using the Data Envelopment Analysis (DEA) approach. Their study provided empirical evidence that PCR is superior to ROR in production efficiency.

Estache and Rossi (2005) studied labor productivity data of a sample of privatized and public electricity distribution firms in Latin America. Their analysis was in line with theoretical expectations that firms operating under PCR and hybrid schemes are more efficient than firms operating under ROR. Another interesting finding is that labor productivity appears similar in private and public firms, and this result is not consistent with traditional hypothesis and studies.

Bernstein (2006) observed an annual TFP growth of 1.66% from Telefonica del Peru during the period 1996-2003, and generated an X-factor of 4.06% accordingly. The efficiency improvement can be demonstrated by the following two results: 1. During the transition to PCR, the telecommunication industry grows faster than the general industries in Peru. 2. At the same time, input price of the sample telecommunication industry grows less rapidly than the general industries in Peru.
4.3. Alternative or combined regulations

To avoid the pitfalls of pure ROR or PCR, regulators may adopt a variety of alternative or combined regulations that possess the features of both regimes and other regulatory forms. Given the fact that the firm's preference and capability of cost reduction are private knowledge and are usually asymmetric information to the regulators', Lewis and Sappington (1989) introduced a model that contains a selectable "menu" of two regulatory regimes. In this model, two regimes are offered to be selected: PCR or surplus sharing (can be seen as ROR here). Firms will be automatically grouped into high or low productivity according to their choices of the regime. This self-selection mechanism auto-identifies the high-productivity firms which choose PCR and pursue cost-reducing investments, and retains the guarantee of firms choosing ROR to protect their incentives for investment.

Pint (1992) introduced a stochastic-cost model to compare ROR and PCR, and showed that both regimes can induce overinvestment. While ROR is usually associated with stochastic intervals between hearings and the adoption of average costs since the previous hearing, PCR is associated with fixed intervals and test-year costs approach adopted by regulators. A numerical example then illustrated that combining the use of average-cost data, the fixed intervals bring more welfare and transfer more gains to consumers. So this result suggested PCR with the use of average-costs as a better solution, which combines elements of both ROR and PCR.

Baake (2002) analyzed a stylized network industry under two hybrid sets of regimes: PCR combined with ROR versus PCR combined with a universal service obligation (USO). The result illustrated that both regimes improve social welfare compared to pure PCR. Among these two enhancements, PCR/USO is a better choice from a welfare perspective when firm's profits and the number of customers is regarded as fixed.
Parker and Kirkpatrick (2005) argued against the assumption that PCR is better than ROR for low-income economies, and found the following difficulties for regulators to implement PCR in developing countries: inadequate information, lack of professional knowledge, and institutional defects. They further introduced sliding-scale regulation as a third and useful approach, and compared the three regimes, as summarized in Table 4:

**Table 4 - Summary of the relative advantages of ROR, PCR and sliding-scale regulation in low-income economies**

<table>
<thead>
<tr>
<th></th>
<th>ROR</th>
<th>PCR</th>
<th>Sliding-scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency Incentives</td>
<td>Low: incentives to inflate opex(^1) and capex(^2)</td>
<td>High provided that price reviews are infrequent; efficiency benefits retained until the next price review</td>
<td>Medium: suppliers can earn excess profits by being more efficient but must share these profits with consumers</td>
</tr>
<tr>
<td>Difficulty of administration</td>
<td>Low: requires monitoring of current revenue and cost data to prevent inefficient expenditures, but the process is similar to that which occurs under state ownership</td>
<td>High: requires information on future financial and economic trends that may be beyond the ability of a regulatory office in a low income economy to collect</td>
<td>Medium: particularly needs regular and reliable current profit data</td>
</tr>
<tr>
<td>Threat of regulatory gaming</td>
<td>Low: rate of return can be reset to cover the cost of capital annually, or even more frequently if necessary</td>
<td>High: inflating of cost of capital and opex and capex needs when the cap is set. Difficult to correct quickly later</td>
<td>Medium: risk of hiding profits</td>
</tr>
<tr>
<td>Threat of regulatory capture</td>
<td>High: frequent rate reviews may encourage capture</td>
<td>Low: but benefits obtainable over a lengthy period of the price cap are set too generously</td>
<td>Medium: suppliers may apply pressure to influence agreed profit levels and sharing rates</td>
</tr>
<tr>
<td>Risk of political and social rejection</td>
<td>Low: prices set according to costs and therefore more likely to seem fair</td>
<td>High: excess profits or losses leading to closure are both risks</td>
<td>Medium: share higher profits</td>
</tr>
</tbody>
</table>

Notes: The terms low, medium and high are “relative to the other two methods”;
1. Opex = operating expenditure;
2. Capex = capital expenditure

Source: Parker and Kirkpatrick (2005)
Blank and Mayo (2009) constructed a regulator's choice model to explore a regulatory optimum by maximizing total political support, which aggregates the firms' and customers' support. The former can be expressed by economic profit (a form of increasing function of price level) and the latter can be expressed by consumers' surplus (a form of decreasing function of price level), so the optimal rate-setting is possible. Analyzing the profit sharing regulation, their study concluded for welfare improvement and suggested a hybrid regime to combine ROR and PCR with a profit sharing plan including side payment arrangements.

5. Conclusions

In practice, rate of return regulation is more suitable for countries with inadequate investment in natural monopoly industries, while price cap regulation is more applicable for countries with more focus on costs and prices. Choosing between the two types of regulation should take into account key factors below.

5.1. Increase in demand and capital

The application of price regulation requires consideration of demand and supply in specific monopoly industries. A potential deficiency of price cap regulation is its possible lack of incentives for investment. Under rate of return regulation, firms do not hesitate to make investments. However, with price cap regulation, the regulator no longer assures investors of their returns on investment, which creates uncertainty for investors. Therefore, relatively speaking, it is less likely to adopt the price cap regulation when dealing with industries that have rapidly increasing demand and require large-scale investment. To the contrary, it is more sound
and safe to apply price cap regulation to industries that have supply exceeding demand and need to improve efficiency.

5.2. Technology improvement and cost saving

Technological improvement affects the efficiency of the price cap regulation more. For industries with faster technology development, such as telecommunications, adopting price cap regulation will induce firms to swiftly reduce costs, so price cap regulation is more efficient in this case. However, the application of price cap regulation is less efficient for industries with slower technology development, such as water supply, hydro and gas industries.

5.3. Asymmetric information between firms and regulators

The degree of information asymmetry affects the administrative cost of price regulations. The presence of information asymmetry can be detected in the following ways:

1) Asymmetric cost information. When there is only one single cross-regional monopolist in the industry, the cost information is relatively highly asymmetric due to the absence of efficient conditions for comparative competition across regions.

2) Asymmetric profit information. For example, in hydro retailing business, because firms usually set prices with complicated non-linear pricing strategies, it is difficult to obtain accurate information about the actual prices as well as profits.

3) Asymmetric quality information. In many cases, there are difficulties to monitor the quality of the monopolistic products and services at a low cost, which results in a high level of asymmetric quality information.
(4) Asymmetric demand information. When the demand for monopolistic products is changing dramatically and is not predictable in the future, the information of demand is highly asymmetric.

When the firm enjoys more private knowledge which it is not willing / not able to share with the regulator, the administrative cost of applying price cap regulation can be relatively high, and this information asymmetry makes it not much different from rate of return regulation. When the regulator finds a reliable way to gain the productivity information easily, PCR is supposed to be superior to ROR (Liston, 1993).

5.4. Product quality

Under rate of return regulation, according to the A-J effect, there is a tendency for firms to overcapitalize. This usually signals overinvestment, which probably induces improvement in technology and service quality. When adopting price cap regulation, additional rules on quality requirement standards must be in place to prevent firms from lowering product quality as a result of their pursuit of cost reduction without related regulations. When products or services for which detailed quality requirement standards are difficult to formulate (or quality requirement standards are difficult to be implemented or monitored), rate of return regulation is recommended, such as at United States Department of Defense (Yu, 2003).

5.5. Possible topics for future research

In future, more studies can be focused on PCR and hybrid regulations due to more adoption of PCR. One of such topics can be related to a trade-off between the commitment to the firm and the consumer's welfare. When there is asymmetric information, the regulator has less
information than the firm does. But if he observes excessive entry during the last period, he can claim an underestimation of X for the last period, and thus increase X in the next period. So firms may consider this uncertainty and make too conservative decisions to enter. This anti-competitive effect will limit enough investment and result in negative welfare consequences. On the other hand, if the regulator declares less variation in X, sets higher price cap, or makes other deregulating efforts, the price level can be lower as a result of new entries and more competition. But for monopolists, fewer restrictions against the firm’s market power may not achieve competitive price which is the best price for social welfare. It would be interesting if a new study can discuss this game and provide a better solution under the condition of asymmetric information.
6. References


