Regulating utilities: Have we got the formula right?

Richard Dobbs and Matthew Elson

Although much copied, Britain's RPI–X approach may not be the best way to drive better service and efficiency in the long run

A formula based on the use of rolling historical benchmarks could provide better incentives

Since Margaret Thatcher began selling off state-owned utilities in 1984, the United Kingdom has become something of a role model for privatization. Governments around the world have embraced not only the privatization ideology but also its methods. The RPI–X formula, which is used to set the prices that UK privatized utilities charge their customers, has driven gains in efficiency so successfully that it is becoming the method of choice in countries from Argentina to Australia.

But RPI–X is imperfect, and over the past few years flaws have begun to emerge. It is appropriate for driving efficiency improvements in the early years after privatization, but something different will be needed in the longer term. One possible replacement is a system based on rolling historical benchmarks.

Why regulate?

The productivity of a country's utilities plays a vital part in the productivity of the overall economy for two reasons. First, the price other companies pay for their share of these essential services affects their own performance. Second, utilities in themselves generally make up some 10 percent of a nation's GDP.

It is the importance of the utilities to the national well-being that explains why, at least until the 1980s, most governments chose to own them. But state ownership often undermined efficiency and investment – problems that gave rise to the recent trend toward privatization.

Private ownership also makes it possible to introduce competition, which should in turn force improvements in performance. In reality, however, a significant proportion of the core activities of a utility consists of natural monopolies; either the market is not large enough to support two players operating at the minimum efficient scale or a single system confers major benefits. (Having one gas pipeline down a street, say, is clearly cheaper than having two.) Competition therefore cannot drive efficiency for these activities, although technological advances may change the minimum efficient scale. In any event, natural monopolies,

The authors wish to thank Julia Bede-Cox, Eileen Burnett-Kant, Paul Clarke, Richard Edwards, Nikolaus Frei, Adam Frost, Stephen Hall, David Kausman, Peter Lambert, Keith Leslie, Roy Maslen, John Roberts, and Ashutosh Shastri for their insights and contributions to this article.

Richard Dobbs is a principal and *Matthew Elson* is a consultant in McKinsey's London office.

like monopolies in general, have the power to charge high prices and thus need to be regulated.

In the United States, where utilities have always been privately owned to a large extent, independent regulators cap prices so that profits are confined to the cost of capital return on a utility's asset base. Unfortunately, this "rate-of-return" regulation gives utilities no incentive to improve their performance by reducing operating costs or spending capital more efficiently. If anything, it encourages companies to spend as much money as possible gold-plating their assets to create a bigger asset base on which to earn a return.* So when the first big UK privatization, that of British Telecom, took place in 1984, a new form of regulation was introduced: RPI–X.†

* This gold-plating effect occurs when the allowed return for a utility is above its true cost of capital.

The RPI–X formula was specifically designed to give companies an incentive to improve their efficiency and to pass on the resulting gains to customers. The prices a company may charge are set for what is known as a "control period," typically lasting five years. These prices are supposed to let the company earn what is deemed a fair return on its assets if it operates as efficiently as would a benchmark company.

† Two academics, Michael Beesley and Stephen Littlechild, developed RPI–X. In his 1983 report *The Regulation of British Telecommunications' Profitability*, Littlechild recommended the RPI–X approach to regulate British Telecom. He, however, assumed that not only RPI–X but also the regulation of profits in general would be merely a temporary means of "holding the fort until competition arrives." RPI–X was therefore meant to be a temporary solution.

An independent regulator first assesses the savings the utility achieved by improving its efficiency during the previous control period and translates them into an initial reduction in its prices. Next, the regulator forecasts what a "reasonably efficient" benchmark company would require during its next control period to cover its operating costs, its renewal capital investment (to replace existing assets), and its enhancement capital investment (to deliver such additional outputs as increased capacity, reduced environmental impact, or improved service). With these costs in mind, the regulator stipulates an annual real-price reduction (X) that will give the utility a fair return on its asset base. The utility can then raise prices every year in line with increases in the retail price index (RPI), which measures inflation, but it must first subtract the price reduction factor (X). Hence the price adjustment formula: RPI–X.

If the utility can operate more efficiently than its regulator's benchmark (a reasonably efficient company) during the control period, it keeps the benefits of the unanticipated improvements until the end of that period. Thus the utility has an incentive to improve its efficiency. Benefits that persist over time are usually passed on to customers in the form of price cuts at the start of the next

control period, when the utility must raise its efficiency further to go on earning more than its basic allowed return.

Over the past 15 years, privatization in general and RPI–X regulation in particular have proved their worth in the United Kingdom. Operating costs have fallen; in electricity supply and distribution, for example, they fell by 17 percent in real terms from 1992 to 1997, and the number of employees in the electricity industry as a whole simultaneously dropped by more than 26 percent. End users have benefited, as well. In the gas industry, prices paid by UK domestic customers have fallen by 14 percent in real terms since 1992 – a reduction much greater than the decline in the average wholesale contract price of gas for delivery to the United Kingdom. Service standards have in most cases improved. Finally, investors should be happy: the average total annual return to shareholders for UK gas, electricity, and water utilities was 28 percent from 1992 to 1998, as against a 17 percent annual total return to shareholders for the FTSE-100.

The flaws

Despite these benefits, problems with the RPI–X formula are beginning to emerge. For one thing, the use of a forward-looking benchmark to calculate what a reasonably efficient company might require by way of funding distorts the objectives of management. RPI–X also creates inconsistencies in the treatment of operating costs, provides no incentives to improve capital spending efficiency, and generates regulatory uncertainties that help raise the utilities' cost of capital.

The need to manage the regulator

If a company performs more efficiently than the regulator's benchmark level, it is allowed to retain the savings it achieves until the next regulatory review. As a result, the company wants the benchmark for the coming control period to be as soft as possible. The management of the company therefore has as much incentive to manage the regulator's expectations of its future performance as to improve that performance.

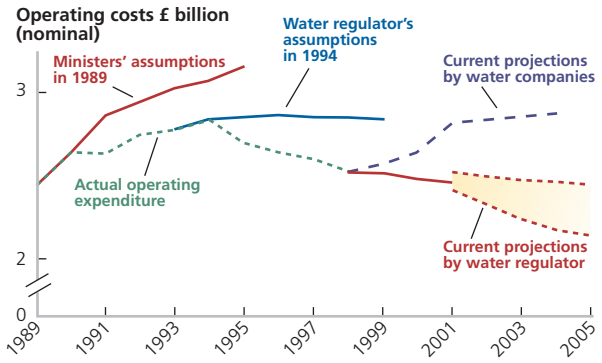
Regulators are at a disadvantage, since they must rely mainly on a company for information about its future costs, volumes, and capital investment needs. The mere fact that a company has an incentive to manage projections,

irrespective of whether it actually does so, means that the regulators have grounds to suspect the accuracy of those projections, thus distorting the price-setting process. This is an example of what John Roberts, an incentive economist at Stanford University, has described as the "barman problem." It is easy, economists reason, for a hypothetical barman to steal a little money from the bar till. Recognizing this, the owner of the bar reduces the barman's wages to reflect the amount of money a dishonest barman might steal. But by behaving in this way, the owner is pushing the barman to steal, as his wages will not be sufficient.

Recent experience in the UK water industry illustrates the problem. In the past two control periods, the water utilities' costs turned out to be much lower than forecast. Having been caught out twice, the regulator has drawn up projections of its own, and they are very different from those of the water utilities (Exhibit 1).

Exhibit 1

UK water industry operating cost projections and outcome



Source: *Prospects for prices*, Ofwat, October 1998

Another consideration is the amount of time and effort that management and regulators must devote to reviews. Because of the barman problem, regulation has become more intrusive as regulators try to understand the efficiency of the company and the decisions it has made, thus departing from the original principles of light-handedness. The five main British regulators, for instance, employ a staff of nearly 900 and have an annual budget of nearly $100 million. The management teams of the utilities devote much effort to managing their regulators – time that cannot be spent improving the performance of the utilities. As the head of regulation at one UK utility admitted, "For three years out of five, the management team is distracted by the upcoming periodic review and the implications of their actions on the outcome."

Distortion in the timing of improvements
RPI–X aims to encourage utilities to cut their operating costs and provides a mechanism to share the benefits with consumers. But one of the formula's side effects is to make companies focus on achieving cost savings at the start of a control period, since these savings are normally retained only until its end; thereafter, they are passed on to consumers in the form of price cuts. The earlier cost savings can be made, the more the company benefits: it captures about 30 percent of the benefits in perpetuity for savings made in the first year of the control period, but only 7 percent for savings made in the last. As a result, companies have an incentive to delay or conceal improvements until the first year of a new control period. This is not the best way to run a business.

Some regulators have used a so-called glidepath approach to overcome this problem: they carry some of the benefits of improvements made late in a given control period over into the next one. But since it is up to each regulator whether (and how far) to implement this approach, most companies still have less incentive to make improvements toward the end of a control period than toward the beginning.

Lack of incentives to improve capital efficiency
In typical applications of the RPI–X formula, the incentives to improve the efficiency of capital spending for renewal and enhancement are at best unclear. If a company reduces its level of investment, regulators find it hard to determine whether it is investing more efficiently or merely investing less. In consequence, they often prevent it from getting the benefit of unexpected reductions in capital spending, allowing just the actual rather than the planned capital expenditure to enter the regulatory asset base on which the company earns its return. If only the actual capital expenditure is included, the company gains minimal benefit from the reduction. Although exceptions are sometimes made, companies generally do not know in advance how regulators will treat capital spending efficiencies.

In addition, the system gives companies an incentive to make capital expenditures as late in a given control period as possible, so that they enter the regulated asset base at the start of the next control period and begin earning a return quickly.

The fact that companies often reap more benefit from improvements in operating costs than from greater efficiency in capital spending means that there is no incentive to make efficient tradeoffs between the two. This causes particular problems in capital-intensive industries with long-lived assets, where the optimal approach is often to raise spending on maintenance to delay the need for renewal. But if regulators reward reductions in operating costs more generously than greater efficiency in capital spending, the relationship between the two is distorted. This seems to have occurred in the UK electricity supply and distribution industry, where operating expenditures fell by 17 percent from 1992 to 1997, while capital expenditures rose by 35 percent.

This bias toward reductions in operating costs becomes more problematic as the cost performance of a utility improves and additional efficiency has to come from its capital investment program.

Higher cost of capital

Regulators enjoy considerable discretion in determining prices for a forthcoming control period, so the outcome of the review is not easy to predict. Uncertainty produces what the market perceives as regulatory risk (Exhibit 2).

This market perception must explain at least some of the increased cost of equity under the UK system. The beta (a measure of risk) for UK utilities is around 0.9, whereas the corresponding figure for US utilities is around 0.5 when adjusted for similar levels of debt (Exhibit 3). Thus the real allowed return in the United Kingdom must typically be more than 1 percent higher than the return in the United

Exhibit 2

Market perception of regulatory risk

"But Pink [an analyst] warns against assuming all the risks are gone, particularly for the electricity companies, where there are continuing competition and regulatory concerns." *The* Financial Times *on the electricity industry, July 5, 1997*

"Companies with weak noncore operations (e.g., Anglian Water) must accept that they are completely at the whim of a regulator who deems 1–2% real dividend growth to be 'appropriate'."
SG Securities analyst report on Severn Trent, October 1998

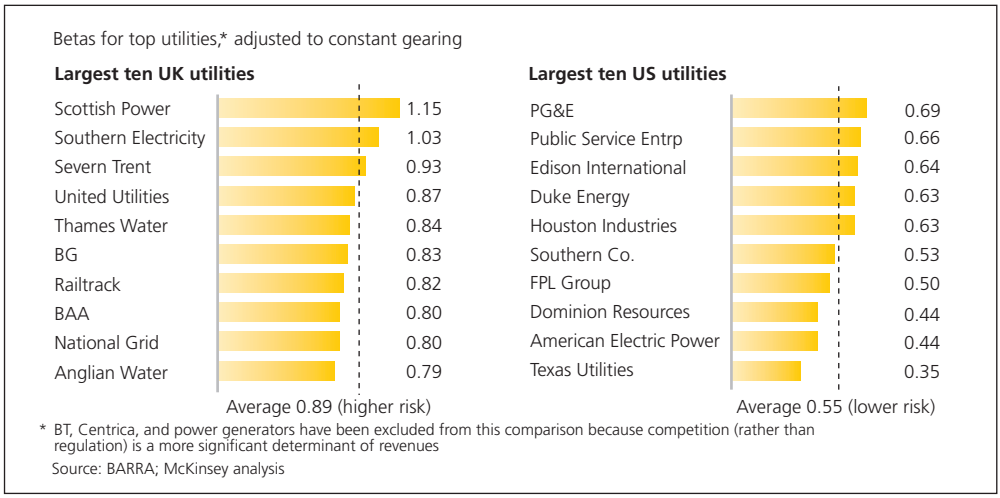
"The only major uncertainty facing the core transmission business is the next regulatory price control which should become effective in April 2001." *Credit Suisse First Boston analyst report on National Grid Group, October 1998*

"History has taught us that the stronger periods of stock market underperformance in the telecoms sector tend to be those of regulatory uncertainty." *ABN-Amro Report on Telecoms, September 1998*

"Regulation and politics will continue to be crucial drivers of the share price, given the group's indirect reliance on government subsidy. The regulatory review began last December and will be drawn out over the next two years. This could be a major cause of volatility in the shares." *SBC Warburg Dillon Read analyst report on Railtrack, March 1998*

Exhibit 3

Riskiness of UK and US utilities



* BT, Centrica, and power generators have been excluded from this comparison because competition (rather than regulation) is a more significant determinant of revenues

Source: BARRA; McKinsey analysis

States. The cost to the United Kingdom's economy is between $2 billion and $3 billion a year in higher utility charges – equivalent to about an extra 1 percent on the basic rate of personal income tax.

Lack of incentives to continue to improve service

Companies have strong incentives to cut their operating costs, so the quality of their services may suffer. Regulators often respond by setting targets for various measures of quality, such as availability of supply. But these targets are rarely based on the real economic value of any improvement in service, and there is rarely a direct link between them and the revenue a company receives. Targets are imposed in a way that gives companies neither an incentive to exceed them nor the freedom to undershoot if the cost of meeting them outweighs the social value of doing so.

A better form of regulation

The shortcomings of RPI–X suggest that a new form of regulation is needed. One possible solution is a model based on a rolling historical benchmark.

With such a model, the regulator would not have to go through the complex process of developing a forward-looking benchmark of what a reasonably efficient company might require in the way of operating costs. Instead, the benchmark would be defined by the actual average oper-

ating costs of the previous, say, five years in real terms. As in the RPI–X model, if a company improved its efficiency, the benefit – namely, the difference between its actual performance and the benchmark – would be shared with customers in the form of price cuts. But since there would be no fixed beginning or end to the control period, the company would receive the same benefit from any savings no matter when it made them.

In effect, the formula would automatically ensure that savings were shared with customers gradually. In the year after the improvement in performance, the rolling five-year benchmark would include just one year of the gain, so the company would retain four-fifths of the benefit while the customer received one-fifth through a cut in prices. In the next year, the company would retain three-fifths while the customer received two-fifths, and so on (Exhibit 4). The company would have an incentive to improve its performance because it would receive three years of the benefit in all. A backlog of "outperformance" would no longer be built up over a five-year period before being abruptly reset by the regulator. Instead, any improvement in performance would steadily be shared with customers.

Some might argue that it is inappropriate to use this type of regulation for newly privatized utilities, since it would give the company too much of a windfall gain if performance improvements were easy to make. Society could, however, receive the benefit of the early years of such easily achieved outperformance through a higher flotation price so shareholders would not capture a windfall gain.

Another advantage of a rolling historical benchmark is that it would automate the regulatory process, removing the discretionary element in the regulator's decision making and reducing regulatory risk and hence a utility's cost of capital.

The rolling historical benchmark would reflect easily audited historical performance data, which can be reconciled with the utilities' published accounts. Since regulators would no

Exhibit 4

Using the new five-year rolling historical benchmark formula

Example: 20% one-off reduction in operating costs in year 1
Real £: year 0 operating costs = 100

	Year											
	–3	**–2**	**–1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	
Actual operating costs	100	100	100	100	80	80	80	80	80	80	80	
Five-year rolling benchmark	100	100	100	100	100	96	92	88	84	80	80	
Outperformance against benchmark					20	16	12	8	4	0	0	**Total: 60**

longer have to rely on projections of the utilities' future capital investments and operating costs, the distrust and hostility evident in some relationships between utilities and regulators should decline considerably. Regulators would concentrate on setting the cost of capital and managing the advancement of competition in areas outside the natural monopoly core.

Last but not least, managers would no longer have to spend three out of every five years anticipating how their actions might affect the next regulatory review. Instead, they would be free to focus on improving their business.

The rolling historical benchmark formula can also be applied to industries that have a number of players with different levels of efficiency. In the United Kingdom, this would include the water and electricity distribution monopolies broken up into regional companies at privatization. Before the formula can be applied, however, it will be necessary to resolve a basic philosophical issue: should the returns of a utility be determined by the rate at which it improves its performance or by its level of performance as compared with that of its peers. If the former, the benchmark for each individual company should depend solely on its rolling-average historical level of performance; if the latter, on the whole peer group's rolling-average historical level of performance, appropriately normalized using statistical techniques to reflect differences in the companies' underlying cost drivers, such as household density.

Three additional features can strengthen the system of rolling historical benchmark regulation. First, renewal capital expenditure should be treated in exactly the same way as operating costs so that effective tradeoffs can be made between the two. As with operating expenditure, the allowance for spending on renewal in the automatic price-setting formula should be based on the rolling average of the previous five years.

Second, to prevent deterioration or to promote improvements in the quality of service, the new model should incorporate a mechanism that rewards companies for their outputs. Aspects of quality of service should be valued, their delivery measured, and the prices the utility is allowed to charge adjusted accordingly.

Some regulators are already attempting to implement such a mechanism. In the United Kingdom, British Telecom must compensate users for faults in telephone lines, so it has an incentive to cut the number of days that service is lost. Similarly, companies in the rail industry are penalized for delays (*see* boxed insert, "The UK rail industry punctuality regime"). Thanks to payments for outputs, delays attributed to Railtrack (which owns the UK rail network) have fallen by 40 percent since the establishment of the incentive regime in 1995.

A rolling historical benchmark formula could also be used to promote improvements in quality, as well as reductions in operating and renewal costs. Utilities would be paid or penalized for the value of any rise or fall, respectively, in outputs in relation to the rolling benchmark. This mechanism would encourage utilities to improve the quality of their outputs because they would receive extra revenue as long as the cost of delivering the improvement was lower than its value. The rolling historical benchmark formula would ensure that price cuts eventually pass along the additional income generated by this improvement to the consumer. In the end, the utility would be reimbursed only for the cost of making the improvement, since this cost would be included in full in the rolling-average operating-cost benchmark. For the utility to continue to earn more than its cost of capital, it would have to go on making improvements in outputs. Regulators would play an important role in facilitating economic analysis to determine which outputs matter and the value of changes in them.

Third, the basic formula could be supplemented to encourage companies not only to make efficiencies in

THE UK RAIL INDUSTRY PUNCTUALITY REGIME

When the UK rail infrastructure company, Railtrack, was privatized in 1996, the government feared that operators of train services on the company's network might experience a rising level of infrastructure-related delays as cost reductions were pursued. To prevent this, a punctuality regime was introduced into the price-setting regulatory formula.

The regime sets a value on delays by estimating the cost of the inconvenience to the expected number of passengers who travel on a given train. It measures punctuality at around 500 points across the 10,000 route miles of network and attributes blame for the delays. Payments made by the train operator to Railtrack for access to the rail network are adjusted to reflect the value of delays caused by the infrastructure company, which can then determine the optimal level of improvement in punctuality by trading off the value of fewer such delays against the cost of achieving these reductions.

enhancement capital spending but also to spend the money only on outputs that deliver value. This would entail the use of a rolling-average regulatory asset base in setting the return element of the allowed revenue. Although all enhancement investments would automatically go into the regulatory asset base, they would not earn the full rate of return for five years, because of the use of the rolling-average asset base in setting prices. Companies would recover the cost of capital return on their investment during this period only by ensuring that it actually delivered outputs that were valued by the output part of the regime described earlier. Again, they could earn more than their cost of capital in the longer term only if they continually made investments that delivered outputs with a value greater than the cost of capital return on the investment. An additional advantage is that the properties of this formula would tend to make companies largely indifferent to whether an investment was classified as an enhancement or a renewal investment, so they would not have to argue with regulators about the allocation of investment between the two.*

Is a stick needed?

One possible drawback of the rolling historical benchmark approach is that managers would not necessarily strive continually to improve the performance of their utilities. Without the spur of periodic reviews that set prices at levels forcing managers to raise the bar, they might choose to sit back and earn the cost of capital on their utilities' current levels of performance.

Since companies would still have the carrot of the extra return they could gain by exceeding the benchmarks, this may not be a real issue. In fact, the UK experience has shown that companies have consistently beaten the benchmarks set by regulators. In any case, if there is a problem it could be overcome in several ways:

- A regulator could tell a company's board of directors and shareholders what performance improvements could be implemented. The standard corporate governance processes would then duly bear upon the company's management team to deliver them. A management team that consistently failed to do this would be exposed to pressures from investors or even to a hostile takeover bid.

* This indifference can be illustrated by considering the economic cost of an investment to a company under both allocations. If the investment is treated as an enhancement, the company receives a cost of capital return on the investment as it enters the regulatory asset base. But since the company only gets its reward in full after five years (because the rolling average of the regulatory asset base is used in determining prices), the investment costs the company about three years of the cost of capital return on the investment. If the investment is treated as a renewal investment, it enters the rolling average of renewal expenditure, and the company receives payback in full over the next five years. The investment has cost the company the time value of money between making the investment and receiving the repayment: around three years of the cost of capital return on the investment. Since these costs are similar, the allocation of the investment between renewal and enhancement affects the company very little.

- There could be a cap on prices in addition to the one provided by the rolling historical benchmark formula, so that management teams that did not improve performance would eventually face diminishing returns. For high-performing companies such a cap would be irrelevant, so they would have little need to influence its level and thus would not have to deal with the barman problem mentioned earlier. For more poorly performing companies, managing the level of the cap would become an important value driver, with all the adverse consequences of RPI–X.

- Assumptions about long-term productivity improvements – 1 or 2 percent a year, say – could be built into the benchmark permanently, forcing utilities to raise productivity at a higher rate if they wanted to earn more than their cost of capital. To preserve the benefits of rolling historical benchmark regulation, these assumptions would have to be set on a long-run basis and not be open to regulatory discretion. Ideally, they should be determined before privatization to give investors a clear idea of the value of a business.

Without doubt, the RPI–X formula gives companies an incentive to improve operating-cost efficiency, particularly in the early years after privatization. It therefore represents a considerable improvement on rate-of-return regulatory formulas. But as opportunities to boost cost efficiency start dwindling, companies need a formula that encourages them to focus on improving all aspects of their performance over the long term. A formula based on rolling historical benchmarks could be the answer.