THE JOURNAL OF FINANCE • VOL. XXXV, NO. 2 • MAY 1980

Admissible Rate Bases, Fair Rates of Return and the Structure of Regulation

BRUCE C. GREENWALD*

A CENTRAL QUESTION in the theory of regulation is what constitutes a "fair" return on a utility's assets. The current legal answer is that provided by the Supreme Court in its Hope Natural Gas decision. A "fair" return should be "commensurate with returns on investment in other enterprises having corresponding risk" (the comparable return standard) and sufficient to "attract capital"[1] (the capital attraction standard). Unfortunately, in practice these guidelines are ambiguous and, as a result, are not uniformly applied. In one sense, they are definitional. There is an emerging consensus that a "fair" percentage rate of return should equal a utility's market determined "cost of capital."[2] However, the Hope Decision did not specify how the value of a utility's assets, to which the "fair" percentage return could be applied, was to be calculated. This paper addresses the problems raised by this indeterminacy.

The basic difficulty in valuing a utility's assets (i.e., its rate base) is one of circularity. Their value is determined, like those of any assets, by the net income they are capable of producing. But, this in turn is determined by the policies of the relevant regulatory agency and, in particular, by the value such an agency places on the assets of the utility. Thus, valuations by a regulatory authority tend to be self-fulfilling and there is no firmly based principle by means of which this circle can be broken. Attempts to break it have traditionally taken two directions.

The more common is to assert that accounting book value is the only appropriate rate base.[3] The primary justification for using book value is that it represents the cash value of funds paid into a utility by investors net of what is returned to them in the form of depreciation recoveries. Book value (or historical cost) is, thus, the net dollar investment embodied in a utility's assets and this presumably is the "fair" value of those assets on which a "fair" percentage return should be paid.

The second method of breaking the circularity is to appeal to a pseudo-competitive entry and exit argument, which leads to the use of a replacement or current cost rate base. In competitive markets, the threat of potential entry will

* Bell Telephone Laboratories.

[1] Both quotes are from the Supreme Court Decision of FPC et al. v. Hope Natural Gas Co., 320 U.S. 591 (1949), p. 603.

[2] See, for example, Myers (1976).

[3] See, for example, M. J. Gordon (1976) who asserts that Book Value is the only reasonable rate base.

rate of return regulation is chiefly to prevent regulated firms from making "monopoly" profits, limiting the value of a utility's assets to replacement cost will ensure that no such profits are made. Also, since utilities do not, as a rule, have the opportunity to redeploy their assets when the returns they earn fall below what the assets could command elsewhere, "fairness" to a utility's owners demands that their assets be valued at no less than this replacement cost.[4] Unfortunately, since replacement cost usually differs markedly from book value, this argument runs counter to that justifying the use of book value.

What this paper shows is that any arbitrarily chosen rate base valuation principle within a very broad class serves to break the asset valuation circle, provided only that a regulatory authority allows the "fair" percentage rate of return which is appropriate to that valuation principle. This class, which will be referred to as the "admissible" class of rate bases includes historical cost, replacement cost and recently proposed inflation adjusted rate bases. For each member of the admissible class, it is possible to define a "fair" percentage rate of return so that the combination of valuation principle and rate of return guarantee an overall return satisfying both "fairness" criteria of the Hope Decision. Also, under certain conditions, the total expected payment to investors (and thus the total expected payment by consumers) is independent of the choice of rate base within the admissible class. Furthermore, in simple models of rate of return regulation and of lagged regulation, all rate bases within the admissible class satisfy the same conditions which support the choice of an economically efficient level of physical investment by a regulated utility.

The paper goes on to consider circumstances under which some rate bases within the admissible class are superior to others. These conditions are of three sorts. First, there is the problem of risk creation in the regulatory process and the distribution of risk between investors and consumers. Second, in using lagged regulation to produce an efficient allocation of resources by a regulated utility, detailed considerations concerning the lifetimes and distributions of returns from investment projects may also dictate a preference for certain rate bases within the admissible class. Finally, output price patterns may be more desirable using some rate bases than using others.

I. Fairness

The class of admissible rate bases can be defined most straightforwardly in terms of the "fairness" criteria of the law. This will be done within the context of a very simple but quite general model of regulation. We will assume that;

(A1) There is no depreciation, all earnings are distributed to stockholders and all investment is equity financed.

(A2) Investor earnings are $E_t = \mu_t RB_t$ where μ_t is the allowed percentage rate of return and RB_t is the rate base in period t.

(A3) At the end of period t, before any new investment is installed, the old rate base is revalued according to a rule, $RB'_{t+1} = (1 + \mathrm{h}_t)\mathrm{RB}_t$ where RB'_{t+1} is the

[4] Arguments of this sort are made among others by Leland (1974) and Myers (1972).

valuation of the rate base in period $t + 1$ before taking account of any new investment installed at that time, and h_t is a factor embodying the rules for revaluing "old" assets.

(A4) There is a well-defined, but perhaps unobservable, "cost of capital", r_t^*, that investors require of investments in any given utility.

Given (A1)-(A4), the return to utility investors consists of two parts. They receive distributed earnings E_t plus capital gains which are determined in the market for each utility's common stock. Letting V_t denote the market value of a utility in period t and V'_{t+1} its value before any new equity sales to finance investment at the beginning of period $t + 1$, the expected market return is;

$$\bar{r}_t = \frac{E_t}{V_t} + \frac{V'_{t+1}}{V_t} - 1 \tag{1}$$

where $(V'_{t+1}/V_t) - 1$ is expected capital gains.

Next, for convenience, let;

$$a_t \equiv \frac{RB_t}{V_t} \quad \text{and} \quad a'_{t+1} \equiv \frac{RB'_{t+1}}{V'_{t+1}}$$

Substitution from the definitions of earnings (A2), rate base valuation (A3) and a_t into equation (1) enables the expected return to investors to be rewritten as;

$$\bar{r}_t = \mu_t a_t + (1 + h_t)(a_t/a'_{t+1}) - 1 \tag{2}$$

So far, equation (2) is a matter only of definition. The next two assumptions are substantive;

(A5) V_t adjusts so that the market for a utility's equity is in equilibrium and $\bar{r}_t = r_t^*$.

(A6) Expectations in the market for a utility's equity are stable in the sense that the expected value of a'_{t+1} equals a_t.[5]

Imposing these two requirements on equation (2) yields the equilibrium condition;

$$r_t^* = \mu_t a_t + h_t \tag{3}$$

At this point, it is necessary to define and impose on μ_t and h_t, the criteria set out in the Hope Decision for determining a "fair" rate of return. An obvious testing ground for the "comparable earnings" standard is financial markets. If utility returns exceed those on commensurately risky assets, then investors will shift their funds into utilities, bidding up their market values and eliminating any excess return. Any return on a utility investment which exceeds r_t^* would generate just such a reaction, just as a return below r_t^* would lead investors away from utilities. Thus, r_t^* represents exactly the return which comparably risky assets should command. Formally, therefore, we will interpret the "comparable earnings" standard as:

[5] This assumption is made largely to simplify the exposition: If regulatory authorities obey the fairness condition (D1) then there is always a rational expectations equilibrium in which (A6) holds with $a'_{t+1} = 1$ in every period.

(D1) "Comparable earnings" requires that the total allowed return on the rate base valuation of assets equal r_t^*.

This allowed return consists of percentage earnings, μ_t, plus the notional capital gains, h_t, implicit in the revaluation of the rate base. Formally, therefore, (D1) implies that;

$$\mu_t + h_t = r_t^* \tag{4}$$

Substitution from (4) into equation (3) yields the result that, when earnings are "fair" in this "comparable earnings" sense;

$$\mu_t + h_t = \mu_t a_t + h_t \quad \text{or} \quad a_t = 1 \quad \text{or} \quad RB_t = V_t \tag{5}$$

Equation (5) embodies the basic content of the comparable earnings standard. Whatever the rate base used, if assumptions (A1)-(A6) are satisfied, then the rate of return on a utility's assets will be equal to that available on commensurately risky assets, if and only if, the market value of the utility is equal to its rate base.

The obvious next step is to investigate the implications of the capital attraction standard and here the primary problem is its apparent redundancy. Utilities with positive earnings should always be able to raise some capital, if they are willing to sell shares at low enough prices, even though the allowed percentage rate of return may fail to meet that required by the comparable earnings standard. In effect, they do this by transferring existing earnings from current to new stockholders in order to provide these new investors with the returns necessary to attract them.

However, interpreting the capital attraction standard as merely the ability to raise some capital is impractically loose. First, this interpretation does not speak to the underlying issue of "confiscation", since it does not preclude effectively confiscating the earnings of existing stockholders in order to finance new investment. Second, if returns are "fair", utilities should not only be able but willing to finance new investment. And, under the circumstances described above (i.e., "low" allowed rates of return), new investment will reduce the equity value of existing shareholders and should, therefore, be resisted by management.

A more appropriate interpretation of the capital attraction standard is that;

(D2) If returns satisfy the "capital attraction" standard, then utilities should be able to raise capital without either gain or loss to the equity of existing shareholders.

Mathematically, this means that;

$$\Delta V_t = I_t \tag{6}$$

where, during period t, ΔV_t is the change in the market value of a utility that results from an investment I_t. Together with the "comparable earnings" requirement of equation (6) this implies that;

$$\Delta RB_t = I_t \tag{7}$$

where ΔRB_t is the increase in the value of the rate base that results from an investment I_t.

Thus, the two "fairness" criteria of the Hope Decision (as interpreted here) are compatible if and only if the rate base valuation rule satisfies equation (7). All rate bases within this group are, therefore, acceptable on fairness grounds and they will be referred to as Admissible Rate Bases. Formally the class of Admissible Rate Bases can be defined as follows;

(D3) An Admissible Rate Base is one calculated according to the rule;

$$RB_{t+1} = (1 + h_t)RB_t + I_t.$$

Most rate bases either used or commonly proposed satisfy this definition. Book value rate bases correspond to an h_t function which embodies the conventional accounting rules of depreciation. Replacement cost rate bases have an h_t which embodies rules for recalculating the replacement cost of the investment previously installed. Inflation adjusted rate bases can be defined by having h_t be an appropriate function of the rate of inflation (appropriately defined) as well as the accounting condition of a utility's assets. Indeed, even the much maligned "fair value" rate bases satisfy the admissibility standard with h_t as book value plus some fudge factor. In all these cases, the value in the rate base of new investment at the time it is installed is equal to the dollar expenditure involved. Thus, they all satisfy the basic admissibility criterion.

II. Rate Base Selection and the Cost of Capital

What has been demonstrated so far is that under a fairly general set of assumptions, regulatory authorities can satisfy the fairness conditions of the law and still retain the freedom to choose among a broad class of rate bases. A question that then naturally arises is how their choice will affect the overall earnings of utility investors and hence the overall cost of capital to consumers.

This question will not be answered completely here. However, the following two theorems, which are proved formally elsewhere,[6] delineate the circumstances under which the choice of a rate base can influence the overall cost of capital.

THEOREM 1. *If there is no uncertainty, frictionless capital markets, and the present discounted value of utility investments in physical capital is finite, then all rate bases within the Admissible Class lead to the same present discounted value of investor earnings, provided only that terminal and initial values of the rate base are zero.*

THEOREM 2. *If there exists either a complete set of contingent commodity markets or a set of securities which spans all contingencies, then the implicit expected cost of investor earnings is independent of the rate base chosen.*

The proof of Theorem 1 is straightforward and instructive. With no uncertainty, all assets must yield equal market returns in each period. This return may be paid either as capital gains or as current income. The choice of a rate base merely alters the proportions between these two. Any change in h_t, the growth rate of the rate base, leads to an equal but opposite change in μ_t, the percentage of current earnings. Furthermore, when rates of return are "fair", h_t also represents

[6] See Greenwald (1979).

the increase in the market value of the utility. In a world of certainty, this increase in market value is in turn equal to the increase in the present discounted value of future earnings. Therefore, any decrease in present earnings leads to an exactly equal increase in the present discounted value of future earnings.

Similarly Theorem 2 states that shifting between rate bases in the Admissible class merely changes the pattern of payments across contingencies without altering their current implicit expected cost.

Together the two theorems indicate that appropriately chosen rate bases can alter the cost of capital only in a world of uncertainty with incomplete markets. However, in such an imperfect world, lowering the cost of capital is not unambiguously desirable. In most cases, reducing investor risk increases the risk bourne by consumers. Thus, in choosing an optimal rate base any reduction in the cost of capital due to lower investor risk must be balanced against the added burden of risk bourne by consumers.

Another point that should be made is that while costs of capital are invariant with either no uncertainty or complete markets, the choice of a rate base is still not an academic exercise. Shifting payment streams has two important consequences. First, it determines who pays. For example, lowering current returns and increasing h_t, shifts the burden from present to future consumers. Second, since current earnings are tied to current price levels, the choice of a rate base determines the temporal and cross-contingency distribution of prices and hence the welfare loss imposed by having to satisfy a utility's budget constraint. Even without uncertainty it is, therefore, important to examine the question of choosing a socially optimal rate base from within the admissible class.

III. Efficiency Without Uncertainty

With no uncertainty, economically efficient regulation must satisfy two basic criteria. First, for any given level of output, regulated firms must be lead to choose efficient combinations of inputs. This requires that they equate marginal rates of technical substitution (e.g. between labor and capital) to those which prevail in the unregulated sector, assuming, of course, that resources in the latter are properly allocated. Without such a condition, rearrangements of factors exist which could, in principle, increase outputs in both sectors.

Second, regulated firms must produce optimal levels of output. Since in practice outputs are determined by prices, this requires that a regulatory authority choose the "right" price levels. In a competitive environment, efficient prices equate the marginal rates of substitution of consumers on the one hand to the marginal costs of firms on the other. However, in a regulated environment, this is seldom feasible. Regulated firms should, in theory at least, be characterized by increasing returns to scale. Marginal cost pricing will usually, therefore, produce revenues which fail to cover total costs. In the absence of some external means of financing these deficits, regulated prices must be set in order to satisfy a budget constraint. Efficient regulation must ensure that the prices which satisfy these budget constraints minimize any deadweight losses imposed on consumers.

The more interesting problem is that of resource allocation, since in simple models a definitive conclusion can be stated. The main result which is again proved elsewhere,[7] is;

THEOREM 3. *All admissible rate bases support an efficient choice of inputs, provided that h_t is independent of current operating decisions and new investment is included at its depreciated*[8] *cost at the time it is incorporated into the rate base.*

To illustrate why the theorem is true and what is meant by depreciated cost, consider the situation of a regulated firm faced with a sequence of inelastic demands, $Q_1, Q_2, \cdots, Q_T$. Assuming that the utility is small and that unregulated markets are in competitive equilibrium, efficient input choices will;

$$\min_{It,Lt} \sum_{t=1}^{T} \rho^t[w_tL_t + p_t^I i_t]$$

subject to;

$$F_t(L_t,k_t) = Q_t \qquad t = 1, \cdots, T$$

and

$$k_t = k_{t-1}(1 - \delta) + i_t$$

where L_t is the amount of labor input, w_t is the level of wages, p_t^I is the price of investment goods, k_t is the capital stock, i_t is real investment, ρ is a discount factor, δ is a depreciation rate and F is a production function with the usual properties.

If an interior maximum exists, the first order conditions which determine the optimal levels of i_t and L_t take the form;

$$\frac{F_t^k}{F_t^L} = \frac{p_t^I - \rho(1 - \delta)p_{t+1}^I}{w_t} \qquad t = 1, \cdots, T \tag{8}$$

where F_t^k and F_t^L are the marginal products of capital and labor respectively. Together with the output constraints, these equations completely characterize the set of efficient input choices.

In contrast, a utility which maximizes its market value will solve the following problem in each period.

$$\begin{gathered} \max p_tQ_t - w_tL_t - p_t^I i_t + \rho V'_{t+1} \\ \text{subject to } F_t(k_t,L_t) = Q_t \\ \text{and } k_t = (1 - \delta)k_{t-1} + i_t \end{gathered} \tag{9}$$

where p_t is the price of output set by the regulatory authority at the beginning of period t.

[7] See Greenwald (1979).

[8] What is meant by "depreciated" cost should become clear below. Instantaneously "depreciated" and current cost are equal. In a broader economic sense "depreciated" cost is similar to theoretical "replacement cost". See Greenwald (1979) for a more extended discussion of these issues.

$\partial V'_t$

Assuming that regulators always set p_t to allow the utility a "fair" rate of return;

$$V'_{t+1} = RB_{t+1} = (1 + h_t)RB_t + p^I_{t+1}(1 - \delta)i_t \tag{10}$$

where $p^I_{t+1}(1 - \delta)$ is the "depreciated cost" of Theorem 3. The righthand equality in equation (10) defines the set of Admissible Rate Bases in this lagged regulatory context.[9]

Substituting from equation (10) into equation (9), the maximand of the firm becomes;

$$\max p_tQ_t - w_tL_t - p^I_t i_t + \rho[(1 + h_t)RB_t + (1 - \delta)p^I_t i_t]$$

And, the first order maximum condition is;

$$\frac{F^k_t}{F^L_t} = \frac{p^I_t - \rho\, p^I_{t+1}(1 - \delta)}{w_t} \tag{11}$$

Since this is identical to the efficiency conditions of equation (8), a utility which is regulated in this way should always choose an efficient combination of inputs for any rate base within the Admissible Class (i.e. for any sequence of h_t's).

The fact that the combination of a "fair" percentage rate of return with an Admissible Rate Base supports optimal input and investment decisions, also argues in favor of the traditional rate-of-return-rate-base approach to regulation. It has been suggested that the division of earnings determination into separate rate of return and rate base phases is essentially artificial. However, setting earnings as a lump sum provides no signals to utilities which might support optimal input decisions. Within the traditional context, this can be done using relatively accessible information on rates of depreciation and market returns.

The second problem of optimal price determination leads to the standard Ramsey-Baumol-Bradford-Boiteux analysis with commodities differentiated by time of delivery rather than by type. Desirable rate bases in this regard are ones which lead to high current earnings in periods of low demand elasticity. Capital gains (i.e. rate base increases) will be low or even negative in such periods and high in periods of high demand elasticity.

IV. Efficiency Under Uncertainty

With complete contingent commodity markets uncertainty has no fundamental impact on the analysis of the last section. Prices and rate base values need only be redefined over contingencies as well as time periods. Inherently, therefore, the interesting case is that of an economy with incomplete markets.

In this context, uncertainty affects the analysis in two significant ways. First the distribution of risk between consumers and investors now becomes an important efficiency consideration. Unfortunately, it is difficult to say anything definitive on the subject. Random revaluations of the rate base increase the risk

[9] It is straightforward to show that when regulators set a "fair" level of prices at time t, then $\frac{\partial V'_t}{\partial i_t} = 0$ at the optimum level of investment. Also as the lag tends toward zero $(1 - \delta)p^I_{t+1}$ tends toward p^I_t and $p^I_t i_t = I_t$, so this reduces to the definition of the last section.

borne by both investor and consumers and should, therefore, be ruled out. If there is no money illusion, this means that random real revaluations should be avoided and, thus, that an inflation adjusted rate base is the proper point of departure in considering risk distribution. However, beyond these two points, the only positive theoretical guidance is qualitative. If consumers are strongly risk averse and investors are not or if utility risks are largely diversifiable through asset markets, then investors should bear most of the risk associated with utility operations. On the other hand, if the market for utility stocks is a limited one and large numbers of customers pay only a small portion of their incomes for utility services, then there would be advantages to spreading the risks of earnings fluctuations among the broad base of consumers (who would be compensated with lower average prices).

The second impact of uncertainty is that it seems likely to complicate the problem of inducing optimal input choices. As noted in Section II different rate bases may lead, in the absence of complete contingent markets, to different expected costs of capital. There is an advantage (assuming that consumers are risk neutral) to using rate bases with lower costs of capital, since they reduce the deadweight loss generated by the budget constraint. But, there is a danger in using such a rate base. If the cost of capital is reduced by an appropriate choice of rate base, then the implied subsidy on capital could lead to over-investment.

In fact, this concern is misplaced. Admissible rate bases continue to support optimal resource decisions. The fact that creation of any "new" security is tied inextricably to utility investment means that investment serves two beneficial functions. First, it reduces the utility's costs and helps provide capacity for new output. Second, it creates "new", more desirable patterns of risk-spreading. An efficient level of investment is one which balances the sum of these two benefits against the opportunity cost of the capital employed, and this is precisely the level that a value maximizing utility will undertake. Admittedly, such a level of investment could not be justified solely in terms of the benefits of the new capital in production. However, the added benefits in terms of risk spreading exactly counter-balance the subsidy implicit in the reduced cost of utility capital.[10]

V. Depreciation

The principal limitation of this analysis is the rudimentary structure of depreciation and investment which characterizes the models described so far. Two aspects of this structure have especially significant consequences for the generality of the conclusions.

First, while Theorem 3 implicitly allows physical capital already in place to be valued at will, it rests on a distinction between "new" and "old" capital that is largely artificial. In the Theorem, new capital must be properly valued, because the marginal decisions of firms, which affect only investment, are influenced by

[10] In the matter of output prices, the complication introduced by uncertainty is purely formal. The Ramsey price analysis of the previous section can be carried out over the time periods and contingencies covered by existing markets.

this valuation. The valuation of old capital, which can no longer be controlled by firms, need not be governed by efficiency considerations. This accounts for the freedom which regulatory authorities have in assigning a value to old capital.

In practice, firms can control the rate of investment (usually disinvestment) even in old physical capital via their maintenance and usage policies. Thus, efficiency considerations will, in reality place some limitations on the valuation of old capital as well. Nevertheless, while this does lead to an increased emphasis on replacement cost, the fact that efficiency is influenced by marginal valuations means that there remains a broad class of Admissible Rate Bases (defined essentially by different constants of integration), and risk and pricing considerations should still enter into choices among them.

Second, while replacement cost is easily determined when depreciation rates are constant and proportional, in the real world replacement cost determination is not easy. In fact, a problem arises in theory as to whether replacement cost can be defined exogenously or whether it depends on the investment and other policies of regulated firms. If the latter is true, then, in Theorem 3, there might be no unambiguously defined replacement value which could be used for incorporating new investment into the rate base. In fact, as is shown elsewhere,[11] as long as demand is inelastic, "replacement" cost can be defined unambiguously. However, the information needed to do so is usually unmanageably extensive. Therefore regulatory authorities may be limited to relatively simple rate base calculations and in these cases the analysis in Sections III and IV can be applied with only slight alterations.

REFERENCES

Baumol, W. J. (1967), "Reasonable Rules for Rate Regulation: Policies for an Imperfect World," in A. Phillips and O. E. Williamson, eds. *Prices: Issues in Theory, Practice and Public Policy*, Philadelphia: Univ. of Penn. Press.

Baumol, W. J. (1971), "Optimal Depreciation Policy: Pricing the Products of Durable Assets," *Bell Journal of Economics*, (2), pp. 638–656.

Baumol, W. J. and Bradford, D. F. (1970), "Optimal Departures from Marginal Cost Pricing," *Amer. Econ. Rev.*, (60), pp. 265–283.

Boiteux, M. (1971), "On the Management of Public Monopolies Subject to Budgetary Constraint," *Journal of Economic Theory*, (3), pp. 219–240.

Gordon, M. J. (1976), "The Current State of Rate of Return Regulation," Paper presented at the FCC Conference on Future Planning, Washington, D.C., July 12, 1976.

Gordon, M. J. (1977), "Comparison of Historical Cost and General Price Level Adjusted Cost Rate Base Regulation," *Journal of Finance*, (32), pp.

Greenwald, B. C. (1979), "Admissible Rate Bases and Fair Rates of Return," unpublished paper, Bell Laboratories.

Leland, H. E. (1974), "Regulation of Natural Monopolies and the Fair Rate of Return," *Bell Journal of Economics*, (5), pp. 3–15.

Myers, S. C. (1972), "The Application of Finance Theory to Public Utility Rate Cases," *Bell Journal of Economics*, (3), pp. 58–97.

Myers, S. C. (1976), "Rate of Return Regulation—A Critical Appraisal," Paper presented at the FCC Conference on Future Planning, Washington, D.C., July 12, 1976.

[11] See Greenwald (1979).