

AN EX ANTE PERSPECTIVE ON
DEREGULATION, VIEWED EX POST*

Louis Kaplow

Discussion Paper No. 32

6/87

Program in Law and Economics
Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by
a grant from the John M. Olin Foundation.

*This paper was presented at a conference on Competitive Issues in Electric Power, held at the Kellogg School of Management, Northwestern University, in March 1987. It is derived from two larger works on transition policy. See Kaplow (1986, 1987). Many formal details, elaborations, and qualifications concerning the model can be found there. The remarks contained here concentrate on and extend those portions most relevant to the context of deregulation.

AN EX ANTE PERSPECTIVE ON
DEREGULATION, VIEWED EX POST

Louis Kaplow

AN EX ANTE PERSPECTIVE ON
DEREGULATION, VIEWED EX POST

Louis Kaplow

Abstract

The transition from one legal environment to another--such as that involved with electric utility deregulation--raises concerns related to the efficiency of the adjustment process and questions of whether gains and losses caused by the change warrant mitigation, such as through compensation and windfall taxation. This investigation addresses the latter issue from an ex ante perspective that analyzes the effect of the anticipation of transition relief on incentives and risk bearing. It extends the work presented in Discussion Paper Nos. 7 and 20 by applying the basic model of transitions to the prospect of deregulation when some or all of the relevant ex ante investment decisions are in the past.

Introduction

The problem of transition from one legal environment to another raises two rather separate issues. One concerns the adjustment process, by which agents are informed and begin to act upon different rules and capital, operations, and contracting begin to move toward a new equilibrium. This issue is forward looking: although such a process begins from the point that exists at the time the process of change commences, the problems raised concern how quickly the trajectory of a system will change, in what direction, and at what cost. Much of the analysis of electric utility deregulation has been devoted to these issues.

A second aspect of the transition to a regime of deregulation concerns the bearing of gains and losses by those who made investments under the old rules but now find themselves operating in a new environment. For example, some capacity constructed (or under construction) will substantially decrease in value if it is faced with new competition and is no longer guaranteed the return that existed under regulation. In other instances, capacity that may have been excessive, or only marginally profitable, under the old regime may increase in value when permitted to be used in a way more openly competitive with other, less efficient capacity. Such "windfall" losses and gains are often viewed as an unfortunate, unfair, and perhaps even inefficient side-effect of legal change, and thus deserving of some sort of mitigation. In some contexts, such mitigation may consist of compensation for losses or windfall taxation of gains. In others, grandfather provisions, delayed implementation, or phased-in implementation may be advocated to mitigate these effects of transition, in whole or in part.

This paper is addressed to this latter aspect of the transition to a deregulated environment.¹ As a result, the analysis will take as given the effects of deregulation. In particular, it will consider the economic justification, if any, for departing from what otherwise would be the optimal deregulatory path -- i.e., the path that would be efficient to pursue in the absence of this latter set of transition considerations. Since the arguments presented are general, they will abstract from the details of the adjustment process, while making note of how the likely features of electric utility deregulation bear on the relative importance of various conclusions.

The analysis will be presented in two stages. The first section offers an ex ante analysis of transition relief. The basic model evaluates the impact of the prospect of deregulation -- combined with the prospect of transition relief -- on ex ante incentives and risk bearing. In a simple, two-period model, insurance and investment decisions are made in the first period, where there is a given probability that deregulation (or, more generally, any relevant legal change) will occur in period two. To evaluate welfare effects, the model assumes that reform of the existing regime occurs in period two if and only if the information and circumstances at that time indicate that reform is desirable. Moreover, whatever transition relief (compensation, windfall taxation, phased-in implementation) is to be provided is assumed to be fully anticipated in period one.

The basic result of this model is that any transition relief -- whether involving the mitigation of losses or extraction of "windfall" gains -- is inefficient. Even when markets are unable to achieve first best risk

¹ This investigation also does not consider legal complications that might accompany, e.g., divestiture proposals that would arise due to bond covenants and other legal commitments. [See Golub and Hyman (1983).]

mitigation, transition relief designed to further mitigate risk imposes a greater cost in distorting ex ante incentives than the benefits achieved by reducing the incidence of risk. In concrete terms, letting both gains and losses fall unmitigated will provide ex ante incentives for investment in period one to adjust optimally to the prospect of deregulation in period two. These incentives will be blunted to the extent investors anticipate that transition relief will mitigate those gains and losses that would otherwise arise in the event of deregulation. The argument suggests that, in terms of investors' incentives and risk bearing, the problem should be viewed as analogous to that concerning whether the government should mitigate market risks -- as from uncertain demand and technological change -- where it is generally thought that no relief is appropriate.

Section 2 addresses the applicability of the ex ante perspective of section 1 in light of the fact that the immediate prospect of many transitions one might evaluate, including that involving electric utility deregulation, is ex post in the sense of the model. Basically, period 1 in the model might be said to correspond to the past few decades, while period 2 corresponds, roughly speaking, to the present. Reflection on the analysis of the ex ante model will reveal that transition relief -- in particular, full transition relief -- is desirable from a static, ex post perspective. This model is one of many examples of the basic time consistency problem with much policymaking. In the long run, the government cannot hope to achieve the benefit from repeated ex post deviations from announced intentions without incurring the accompanying ex ante costs. Since the repeat nature of the game -- taking into account the broad realm of regulation, and legal rules more generally, rather than the particular context of electric utility deregulation -- enables the government to establish and maintain a reputation for consistent action,

the question presented by a particular instance of legal change is whether transition relief based on ex ante considerations is necessary either to maintain an existing positive reputation or to begin the process of establishing one.

Another aspect in which section 1's analysis may be relevant despite the time frame in which this issue is presented concerns the fact that, in many respects, electric utility deregulation may best be thought of as being, at least in part, still in period 1. Even though many investment decisions that will shape the deregulated environment have long been made, it could well be years, or decades, before deregulation is accomplished. As a result, the ex ante perspective is appropriate as to investment decisions in the interim.

Transition relief, particularly from the ex post mind set that typically accompanies its consideration, is often viewed as a mechanism that may be used to buy off opposition, increasing the political feasibility of change. Section 2 offers some ex ante perspectives on this consideration as well. The section closes by commenting briefly on the available empirical estimates of the transfers that might result from electric utility deregulation. Concluding remarks appear in section 3.

1. EX ANTE ANALYSIS OF TRANSITION RELIEF

1.1. Model and Assumptions

In the first period, investors make investment and insurance decisions. Government transition relief, modeled as compensation (windfall taxation, in the event of gains), is provided in period two, but is fully anticipated.² All costs and returns are denoted as period two receipts, to minimize notation and to avoid the need to make discounting explicit.

For simplicity, the model assumes that there are only two states -- one corresponding to maintenance of status quo policies and the other to regulatory reform.³ Reform is assumed to be implemented if and only if information and circumstances make it optimal (see note 24), from which it follows that the unmitigated private costs and benefits do not diverge from social benefits.⁴

² It is demonstrated in Kaplow (1987) that other forms of mitigation -- such as grandfathering, delayed implementation, and phase-ins -- are essentially the same as compensation in terms of ex ante effects, although typically less efficient in terms of ex post effects. As a result, they will not be examined here at length. See subsection 2.2.4.

³ It should be clear from this formulation that uncertainty, rather than change, is central to raising the transition issue presented here. For if it were already known in advance that a change would be implemented (with certainty) at a specified future date, such a change would be fully reflected in ex ante decisions, and its implementation would raise no argument for relief.

⁴ To the extent that either state represents a second best policy, a more complex model would be necessary to take into account the underlying first best for each state of the world, even if that first best is unattainable. Of

The following notation will be used:

U: investor's utility function; $U' > 0$, $U'' \leq 0$

K: amount of investment

$B_i(K)$: net benefits from investment in state i ; the two states are denoted by "r" for reform and "o" for the "null" state, or status quo; $B_i'' < 0$, $i=o,r$; $\nexists K$ such that $B_i'(K)=0$, $i=o,r$;⁵ $\Delta B \neq 0$,⁶ where $\Delta B = B_o - B_r$ -- " Δ " notation will have the analogous meaning for derivatives of B

a: private insurance premium

b: portion of loss $(B_o - B_i)$ covered by private insurance in state i

c: lump sum tax, or "compensation premium"⁷

d: portion of loss $(B_o - B_i)$ covered by government compensation in state i

p: probability of state r

The social optimization problem is:⁸

course, with such second best problems, the usual qualifications to the welfare analysis would apply.

⁵ Given the stated conditions on the second derivatives, if such a K existed, it would be the optimum and there would be no moral hazard.

⁶ If the optimal investment (K) given no insurance ($b=0$) is such that $\Delta B=0$, then the optimum entails no insurance. In this simple case, there is no risk to insure against in the first instance.

⁷ This compensation premium, also anticipated in period 1, is designed to assure that the government policy breaks even, and in such a manner that the investor, when behaving optimally, does not receive a net transfer (on an expected value basis) from the government. If reform entailed losses, it would be obvious that utility would increase with greater relief (and conversely for gains) if such a term was not added. The result is to permit an efficiency analysis of transition relief -- including effects on the allocation of risk -- separate from distributive effects that, if desired, could, in principle, be achieved independently.

Max U subject to:

d

$$(1) c = pd\Delta B$$

(2) private maximization (optimal insurance decision):

Max $U|_d$ subject to:
b

$$(i) a = pb\Delta B$$

(ii) investor maximization, which consists of:

$$\text{Max}_K U|_{d,b}$$

Note that the utility function is simply:

$$\begin{aligned} (3) \quad U &= (1-p)U_o + pU_r \\ &= (1-p)U[B_o - a - c] + pU[B_r - a - c + (b+d)\Delta B] \end{aligned}$$

As stated in its general form, this model assumes that insurance consists of coverage amounting to a percentage of the losses or gains, the percentage being set by contract. Implicit is that both states and investment levels are not observable and cannot be inferred from other information.⁹ These

⁸ Some readers may note a similarity with the model in Pauly (1974), in which the government here could be viewed as a second insurance company in his model. This model differs in the assumptions made concerning both the government and the insurance company. The problem he addresses concerning multiple insurance companies is implicitly assumed away here, justified by the ability of insurance companies to make coverage contingent on what one has arranged with other insurance companies, as is common, for example, with health and liability insurance.

⁹ The latter is unrealistic, in that when there are only two states and the insurance company observes a loss (or gain) equal to zero in one state and a nonzero effect in the other, the state can be inferred. The purpose of this assumption is to capture the effects of a more complex world -- with many states and sources of gains and losses -- in the simplicity of the two-state model. Assuming that the insurance company can make no inferences from the observation is a worst case assumption, in that it most inhibits insurance from addressing risk and incentives simultaneously. Of course, insurance of this form (percentage of loss covered) is similar to that often observed in many contexts, such as medical insurance. (The fact that modest deductibles are also present is a minor detail that does not significantly affect the analysis.)

assumptions will be relaxed in subsection 1.3.

1.2. Risk Neutrality

The model is easiest to analyze when the investor is risk-neutral. In that case, the following result is straightforward.

Proposition I: If $U''=0$, the optimal insurance policy has $b=0$ (independent of d), $\text{sign } \partial U/\partial b = -\text{sign } b$, the optimal government policy has $d=0$, and $\text{sign } \partial U/\partial d = -\text{sign } d$.¹⁰

Proof: Initially, consider the first order conditions for optimization over K :

$$(4) \quad \frac{\partial U}{\partial K} = (1-p)U'_O B'_O + pU'_r \cdot [B'_r + (b+d)\Delta B'] = 0$$

When the investor is risk-neutral, U' is equal across states, so we have:

$$(5) \quad (1-p)B'_O + p[B'_r + (b+d)\Delta B'] = 0$$

To determine $\partial U/\partial b$ from (2), one can substitute for a , using constraint (i), and keep in mind that K is a function of b , whereas (2) has d taken as given.

$$(6) \quad \frac{\partial U}{\partial b} = (1-p)U'_O \cdot [B'_O K_b - pb\Delta B' K_b - p\Delta B] \\ + pU'_r \cdot [B'_r K_b - pb\Delta B' K_b - p\Delta B + (b+d)\Delta B' K_b + \Delta B] = 0$$

The sum of the first U'_O term and the first and fourth U'_r terms equals zero, based on $\partial U/\partial K$ (4). Divide all remaining terms by p and group as follows:

¹⁰ The derivative with respect to d is not really a partial derivative, but the notation is used to avoid the confusing alternative "dd."

$$(7) \quad -b \left[(1-p)U'_O \Delta B' K_b + pU'_R \Delta B' K_b \right] - \Delta B \left[(1-p)U'_O - (1-p)U'_R \right] = 0$$

Define

$$(8) \quad \bar{U}' = (1-p)U'_O + pU'_R,$$

so we now have:

$$(9) \quad (1-p)\Delta B(U'_R - U'_O) - b\Delta B'K_b\bar{U}' = 0$$

In the risk-neutral case, the first term equals zero, and the sign of $\partial U/\partial b$ will be that of $-b\Delta B'K_b$. $\Delta B'K_b$ is the moral hazard term. Consider the case in which $\Delta B'$ is positive. This indicates that the marginal return is greater in the state involving no reform. In that instance, one would expect an increase in the level of insurance coverage (which applies in the state involving reform) to cause an increase in the level of investment, by making the marginal return in that state, as faced by the investor with insurance coverage, closer to the return in the state involving no reform. Thus, the term should be positive.¹¹ To demonstrate this formally, we can determine K_b from (5) by differentiating that first order condition with respect to b :

$$(10) \quad (1-p)B''_O K_b + p[B''_R K_b + (b+d)\Delta B''K_b + \Delta B'] = 0$$

Rearranging yields:

$$(11) \quad K_b[(1-p)B''_O + p(B''_R + (b+d)\Delta B'')] = -p\Delta B'$$

The bracketed term on the left side simply equals U_{KK}/U' , which, at the optimum, will be negative.¹² Therefore, K_b and $\Delta B'$ have the same sign, and

¹¹ The same argument indicates that, when $\Delta B'$ is negative, the expected sign for K_b reverses as well.

¹² If $U_{KK}=0$, either $\Delta B'=0$ or K_b is infinite. In the latter case, the result for sign K_b holds by the same analysis. $\Delta B'=0$ implies, from (5), that $B'_i=0$ (for $i=0,r$), which is ruled out by assumption. See note 5. In addition, for

the first two results of the proposition follow.

Using the fact that $b=0$, one can now examine the first order condition for d in a parallel manner to that for b . Such manipulation directly yields that the sign of $\partial U/\partial d$ will be the same as that of $-d\Delta B'K_d$. $\Delta B'K_d$ is a moral hazard term for d -- government compensation -- like that for private insurance. Paralleling the above derivation again, it can be shown that, in this case, $K_d = K_b$, which completes the proof. *Q.E.D.*

Remarks: First, a few words on the intuition of the result, which should be rather transparent in this simple case. From (4) and (5), it can readily be seen that the effect of compensation and insurance is to drive a wedge between the private and social investment maximization problems, since, by assumption, the B_i functions reflect the social value of investment. For private insurance, this sort of result is the familiar moral hazard effect. [Arrow (1963), Shavell (1979).] When the investor is risk-neutral, there is no mitigating gain due to risk spreading, so optimal coverage -- both private and government -- is zero. Government relief, here modeled as compensation, acts like an externality in the sense that costs or benefits from adjusting the level of investment are no longer fully borne by the investor. In a moment, it will be demonstrated that this fundamental result -- the inefficiency of any government relief -- carries over completely to more complex cases in which the investor is risk averse.

Second, one might inquire into the relevance of the risk neutrality result. With many government reforms, electric utility deregulation included, one would expect the incidence of most transition gains and losses to fall on investors in widely held corporations, most of whom hold diversified

$d \in [0,1]$, it can be shown (see Proposition IV) that $(b+d)$ is in the same interval, which is sufficient for U_{KK} to be strictly negative.

portfolios. Therefore, as a rough approximation, modeling investors as risk-neutral might be quite appropriate, and the analysis could stop here.¹³

Third, it is worth noting briefly that all of the analysis thus far (and to follow) is equally applicable to gains and to losses due to transition policy. This point should be emphasized because in many transition contexts it is far more common to hear cries for transition relief for losses due to reform, but (not surprisingly) no such demands that gains be taxed away or otherwise mitigated. In a few contexts, such as the windfall profits tax accompanying oil price deregulation, windfalls are explicitly taken into account. Overall, regardless of which concern dominates in a particular transition battle, it is important to recognize that the incentive effects (and, as should be clear from the analysis to follow, risk effects as well) are fully symmetric with respect to whether gains or losses are involved.

1.3. First Best Insurance

For the remainder of the analysis, the assumption of risk neutrality will be dropped. There are still, however, two instances in which the result of Proposition I concerning government relief is almost immediately applicable. When the insurance company can observe the states directly, a first best can be achieved through the familiar use of contingent commodities. Essentially, the insurance contract involves a state-contingent lump-sum transfer. The transfer is lump sum in that the amount is independent of the gains or losses, and thus independent of the level of investment as well. Clearly such a

¹³ If the government reform imposed systematic risk, which often will not be the case, it should be noted that, in principle, the government is in no better position to absorb such risk than are financial markets. See the discussion in Kaplow (1987) of the applicability of the Arrow-Lind results in this context.

scheme has no distorting effect on incentives, while it can fully equalize incomes across states.

Second, when the insurance company can observe the level of investment (ex ante or ex post), premiums can effectively be made a function of investment, fully counteracting the moral hazard effect. Once again, a first best is achieved, in this instance with full coverage ($b=1$, rather than the $b=0$ in the first case) but with a premium $a=pb\Delta B(K)$ -- the notation differing from that in constraint (2)(i) by indicating that the level of K is used in setting a . In both cases, the analysis of government relief is much the same.

Proposition II: *If states or investment levels are observed by the insurance company (or can be inferred from other information), the first best is achieved when $d=0$. Moreover, $d=0$ is the unique optimum, and $\text{sign } \partial U / \partial d = -\text{sign } d$.*

Proof: See Kaplow (1987).

Remarks: The intuition should be straightforward. With such information, the investor is able to join forces with the insurance company to obtain the results achievable in a risk-neutral world. Government compensation creates the same wedge as in the risk-neutral case. This time, the investor's reaction involves both insurance and investment decisions. In the risk-neutral case, no insurance was obtained regardless of government transition policy. In these cases, private insurance will fully, and without moral hazard, cover all of the loss *not* covered by transition relief. Thus the result will be that no risk is borne regardless of government relief, and the incentive distortion will be determined by the level of d , precisely as in the risk-neutral case.

Again, briefly consider the applicability of first best insurance. Of course, the widely held firm described in subsection 1.2 will not need insurance in the first place. For others, one or the other option may, in principle, often be available. For reforms that are limited in number and readily specified in advance (and distinguished ex post), a state-contingent contract should be feasible. For other reforms, it may be possible to approximate such results by use of a proxy variable.¹⁴ Investment levels might also be observable, at least in part. One complication is that investments such as those involving electric generating capacity can take many forms, and thus a model with many types of investment might be appropriate. This would not change the results unless, as might be the case, the insurance company can only observe, e.g., total investment in certain categories, and not how it is allocated within them. It may still be that total capacity, perhaps in combination with some other variables, would be an adequate proxy, but in general one cannot assume that this always will be the case.

1.4. Two-Stage Moral Hazard

When the special cases of subsections 1.2 and 1.3 are not fully applicable, one has what might be viewed as a two-stage moral hazard problem. The simple moral hazard problem is that involving the insurance company and investor when there is no government relief. The second stage is created by the prospect of government compensation. That prospect not only affects the investment decisions, but also, as noted briefly in the simple case in subsection 1.3, affects the insurance decision itself. Thus, a complete analysis of the problem in the general case must take into account the

¹⁴ For example, with reforms affecting the prospects of a defense contractor, a myriad of possible policy changes might be indexed by the size of the defense budget, or that portion of the budget devoted to a particular sector, such as fighter aircraft.

maximizing response in the insurance decision, and how it, in combination with government compensation, will affect incentives. Since it will no longer be the case that risk spreading will be identical (or unimportant) for different levels of government relief, it is necessary to assess the trade-off between the greater risk spreading and the cost in terms of incentives. Despite the greater complexity of the problem, however, the same, simple conclusions largely emerge intact. A brief summary of the main results follows.

Proposition III: $d=0$ is the unique global optimum.

Proof: See Kaplow (1987).

That $d=0$ is a global optimum follows from a simple lemma that can be described briefly. Consider the combination $(d, b|_d)$, referring to a given level of government relief (d) and the level of insurance coverage that is privately optimal conditional on that level of compensation. Compare the utility resulting from $(\tilde{d}=0, \tilde{b}=d+b|_d)$, evaluated at the level of investment that would have been chosen at d . Since the aggregate coverage is the same, the total of premiums ($a+c$) can be shown to be the same, and the investment level is the same,¹⁵ utility must also be the same. Thus, any utility achievable at $d \neq 0$ can be achieved at $d=0$. Proving uniqueness is a bit more complicated. The basic strategy is proof by contradiction. By positing a global optimum at $d \neq 0$, it follows that the transformed " \sim " scheme just described must be a global optimum. Some manipulation of the first order conditions allows one to rule out this possibility.

Proposition IV: $d \in [0,1) \Rightarrow b \in (0, 1-d)$ and $d=1 \Rightarrow b=0$.

¹⁵ The same level of investment is, in fact, optimal for both configurations, which, in turn, implies that $a+c = \tilde{a}+\tilde{c}$.

Proof: See Kaplow (1987).

The intuition is straightforward. At $d=1$, there is no further risk to spread, and $b \neq 0$ imposes two costs: it creates risk not otherwise present and it creates *private* incentive costs. Note that, although $b < 0$ would diminish total incentive costs, since all are borne by the government, there is no private benefit, and in fact there will be a private cost, since, for purposes of private insurance, incentive costs are by reference to a baseline that takes d as given.

For $d \in (0,1)$, $b > 0$ follows from the fact that, when $b=0$, a marginal increase in b creates risk spreading benefits and imposes no *private* incentive costs. $b < 1-d$ follows because, at $b=1-d$, a marginal reduction in b entails no risk bearing costs and reduces private incentive costs.

Simple results concerning the signs of various derivatives are a bit more complicated without further restrictions because of a number of complex and subtle income effects that arise. In the simplest cases, assuming a nonincreasing absolute risk aversion utility function, it can be shown, as one would expect, that total coverage ($b+d$) is greater when $d > 0$ than when $d=0$. The intuition parallels that indicating $d \in (0,1) \Rightarrow b > 0$: if one considers the optimal $b^* = b|_{d=0}$, one would expect that a strictly positive d would entail the desirability of greater total coverage. The reason is that the risk-spreading benefits of increasing total coverage beyond b^* are the same as in the $d=0$ case, but the incentive costs, at the margin, are less when $d > 0$.

Remark: It should come as no surprise that adding risk aversion and taking into account that optimal private insurance is often insufficient to achieve a first best does not affect the overall conclusion. The simple reason is that there is no externality in the private insurance decision, so

government action cannot improve upon it. [See Shavell (1982).] Additional government coverage could have been mimicked by greater private coverage (with $d=0$), but was rejected by the investor, presumably because the greater risk spreading was not worth the incentive cost. The government stipulating that it will provide coverage, regardless of whether the investor would select it, hardly solves the problem. This is the lesson of Proposition III. Moreover, even if the government provides less coverage than the investor would have selected in its absence, there is still some incentive distortion, as described in the discussion of Proposition IV.

1.5. Market and Government Risks Compared

It is quite familiar for economists to argue that government relief against losses induced by market forces -- interpreted broadly to include changes in technology and consumer taste as well as natural disasters and acts of foreign governments -- should not be mitigated by government; otherwise, incentives will be disturbed. Similarly, opposition to windfall taxation -- e.g., of extraordinary gains in the case of a successful new product -- is based on the fact that such "windfalls" are merely the tail of a probability distribution what was part of the ex ante inducement for investment. Chopping off either (or both) tail(s) of the distribution of expected returns generally will distort incentives. This section has offered a formal model equally applicable to market and government risks. The intuition is that, just as a firm might overinvest if losses caused, e.g., by an unexpected decrease in demand, are mitigated, the same effect follows if similar losses caused, e.g., by increased competition due to deregulation, are mitigated by the government, whether through direct relief or adjusting the pattern of reform, such as by delaying implementation to reduce such transition effects.

2. AN EX POST PERSPECTIVE

2.1. The Time Consistency Problem

Section 1 emphasized the efficiency properties of transition relief from an ex ante perspective. It is often the case, however, that the transition relief to be provided in connection with any particular reform is not considered until after the fact: even if it is considered before implementation, it is considered at the last moment, when many of the relevant incentive effects are in the past.¹⁶

Of course, the lack of a stated transition policy is not in itself a problem; so long as there exists a well established approach or tradition concerning transition relief that is generally followed at such moments, the effect will be much the same as if a transition policy were announced (conditional upon enactment of the reform) in advance. One could argue that, at least in some areas, such patterns exist, although they may not be the product of careful deliberation. Yet there is frequently considerable

¹⁶ One modest, although interesting exception arises in connection with many tax reforms, and some others. If there exists, for example, an up-front subsidy that is about to be repealed because the subsidization effect is no longer deemed desirable, it is often understood that the repeal should be made effective as of the *announcement* date of the reform -- e.g., when it is first seriously proposed in committee -- rather than as of the *enactment* date, because otherwise there will be, in the interim, a flood of precisely the activity that the repeal is meant to discourage (at least relative to the subsidized status). Failing to have such retroactivity in the effective date would, in effect, provide full transition relief to those investing just prior to the reform. A more elaborate discussion of this and related phenomena appears in Kaplow (1986) and is formally analyzed in Kaplow (1987).

uncertainty surrounding the debate over appropriate transition relief. Moreover, the character of the uncertainty is not limited to doubts concerning the application of established, refined judgments to the unusual facts of particular transitions, but rather in significant part reflects the lack of a coherent, accepted framework in which to evaluate such issues.

Section 1 provides such a framework, but it is not clear how it should be applied, ex post, in a given context in the absence of an established transition policy. On the one hand, purely static analysis favors full relief, directly contrary to the dictate of Proposition III. The reason is that ex ante incentive effects, for better or worse, are in the past -- they are sunk. At the time of the reform itself, the only effects of relief are distributional. In the context of the model, they can be viewed as ex post risk mitigation. Offering relief of losses (or windfall taxation, for gains) will tend, on average, to shift dollars to higher marginal utility investors, producing an overall gain in welfare. One must be careful, of course, not to overshoot -- i.e., to the extent some have already made risk mitigation arrangements, full government relief combined with partial to complete private transfers would result in more than full compensation which, from a risk perspective, is as problematic as partial relief.

A purely static, ex post perspective is recognized as naive in many contexts closely resembling this one. [See Kydland and Prescott (1977).] Even if there is no government reputation concerning transition policy to disrupt, such ex post decisionmaking is hardly conducive to establishing a credible long-run policy with contours directly contrary to current actions. Thus, one might argue that, despite the static, ex post argument for relief, it is best to forgo the opportunity, explicitly announcing that relief is not to be provided this time, and will not be provided in similar future

circumstances. To defend such an approach, one would have to rule out as infeasible the dominant strategy of providing relief this time, but announcing that it will never be done in the future.¹⁷ In addition, one would have to be sufficiently confident of the value of the contribution of the single action to the credibility of the long-run policy, and the discount rate (taking into account the relative importance of the particular transition at hand and the likelihood of substantial reforms in the future) could not be too high.

Obviously, in light of the complexity of these questions and the political and institutional contingency of some of the issues, no simple conclusions can be stated. Moreover, in practice, one would often wish to distinguish transitions by context. For example, there may exist traditions, for better or worse, in some areas -- such as tax reform or changes in common law rules by courts -- and different traditions (or no traditions) in other areas -- such as deregulation or changes in trade policy. One could imagine that these context-specific traditions could evolve somewhat independently given the different institutions involved in different policy realms as well as the general failure among many participants to understand the similarity of transition questions arising in these different areas.

2.2. Electric Utility Deregulation -- Ex Ante or Ex Post?

2.2.1. The Timing of Deregulation. -- In the particular case of electric utility deregulation, it is important to emphasize that, despite the long-term nature of many investments, there are still significant respects in which the ex ante perspective may be directly appropriate. It seems plausible that deregulation may be years or more than a decade away, at best. Noting that a substantial amount of capacity is currently being planned or is in early

¹⁷ Compare discussions of the desirability of a one-time capital levy.

stages of construction, the potential significance of investment decisions before deregulation is enormous.

Much of the efficiency to be realized as a result of deregulation will arise due to changes in the use of generating capacity, and, in turn, changes in its construction (size, location, type). [Joskow and Schmalensee (1983).] In fact, even those emphasizing the potential efficiency gains arising from the adjustment of generation capacity note that it would probably take years to decades before the necessary realignments were substantially complete. Given a positive discount rate, this significantly affects the potential net gains that could arise due to deregulation. It follows, therefore, that anything that could induce some of the shifts -- particularly those involving investment decisions with long-run impacts -- to be made in advance of deregulation would offer important potential gains.

The model in section 1 demonstrates that the very prospect of reform -- to the extent of its probability -- will have precisely such effects, *but only to the extent that the gains and losses will not be mitigated by transition relief*. If the general expectation is that there will be little, if any, such relief, the status quo must be viewed positively from this perspective. If, however, many expect that there is a serious prospect of relief in some form, it might be that significant gains could be achieved by a credible ex ante announcement (i.e., one in the present, well in advance of deregulation) that there will be no transition relief.¹⁸ How such an announcement could be made

¹⁸ This strategy must be sharply distinguished from delayed implementation and phase-ins. The strategy noted in text and that used in section 1 involve announcement -- as early in time as possible -- that future reforms, at the time they are deemed desirable, will be implemented without transition relief. Delayed implementation and phase-ins involve putting off action already deemed desirable to allow time for adjustment. First, ex ante effects, properly understood by reference to the time of *desired* implementation, are the same under equivalent amounts of direct compensation and delay. Second, any

is a difficult institutional question, particularly since the very fact that deregulation is well in the future suggests that the relevant actors are not yet willing to make commitments along these lines. It may be that, to the extent academic commentary is thought to influence policy, a growing sense that transition relief is undesirable in such contexts would affect the likelihood investors would attach to the prospect.

2.2.2. *The Ex Ante Decisionmakers.* -- There are three relevant groups of actors who are in a position to make important ex ante investment decisions. First, there are users of electricity. For example, some users of power have special access to cheaper sources (e.g., from hydroelectric cites -- Bohn et al. 1982, Plummer 1983) or might face lower rates than others (e.g., given declining block pricing structures -- Bohn et al. 1982);¹⁹ others might be paying higher rates for parallel reasons. If there were going to be transition relief based on the change in expected power costs, as measured by usage patterns existing at the time of deregulation, there would be inefficient ex ante incentives. Those who stand to lose from deregulation, expecting compensation or other relief for their losses, have inadequate incentives to begin economizing on power consumption in advance -- whether it takes the form of locational decisions or investments that change operations at existing sites. Those who might gain but for the prospect of windfall taxation or other mitigation will have inadequate incentives to locate efficiently or expand in anticipation of deregulation.²⁰

further delay imposes additional costs in terms of incentive effects after the time of desired implementation. See note 2.

¹⁹ Rate structures could, in principle, be changed without deregulation of generation, or maintained after it is implemented. The latter seems unlikely because the distribution utilities would be paying the same for each unit of power, regardless of the user, and would only need to recover the costs of the distribution network itself. Thus, it seems less likely that one would observe rates that made such significant distinctions among users -- either by class or by quantity of use.

Second are those supplying factors to utilities. Analogous reasoning applies to their incentives concerning, e.g., future exploration and production decisions.

Finally, and most obvious, are the utilities themselves. (Golub (1982) discusses plants that would increase and decrease in value with deregulation.) Initially, one might simply apply the general reasoning of the model, as it clearly relates to capacity expansion, fuel choice, and locational decisions. The complication with this group concerns whether utility management can itself be considered the decisionmaker, or whether some modification in the analysis is necessary to take into account the role of regulatory commissions. In this regard, it has been noted (Golub 1982) that utility regulation is far more intrusive than many other forms of regulation (e.g., occupational safety, airline or taxi entry restrictions) because it is directly targeted at controlling the rate of return. This is relevant because of its effects on the incentives of utility management, through such regulatory decisions as whether (and when) a new plant will be allowed in the rate base and what return will be permitted. In addition, there are other institutions involved in regulating such decisions as the siting of a new plant that may need to be convinced of the need for the investment.

Regardless of how one chooses to model this process -- a complex issue this paper will not attempt to address -- it seems clear that the ex ante perspective will usually remain important to some degree. For example, a utility might be willing to begin planning and constructing a plant today, despite the anticipation that CWIP will not be included in the rate base or

²⁰ An important subgroup consists of residential users, although one might suspect that they would not be as careful a group of calculators on this score, at least not until deregulation becomes quite likely.

that a low return might be allowed tomorrow, because it anticipates that shortly after completion there is a significant prospect of deregulation, which will make the investment profitable. Of course, if from the ex post perspective at the time of deregulation, it appears that the utility is receiving a significant windfall gain as to its already completed plant, transition relief (taxation) might be provided. If anticipated, this ex post mitigation would be inefficient in terms of ex ante incentives. Similarly, a utility being urged to build new capacity by a commission that seems willing to allow a generous return might hesitate if deregulation is a serious prospect in cases where the plant would not be profitable in the competitive environment. However, if relief were to be provided in such instances, this efficient negative incentive would be distorted as well.

2.2.3. *The Meaning of Transition Relief.* -- Finally, a brief word is in order concerning the ambiguity of the meaning of transition relief in this context. The problem, related to the complex effects of rate of return regulation, is suggested by a simple example. Consider a utility that has already recovered one-third of the present value of the cost of investing in a plant through provision for CWIP. [This example is noted in Golub (1982).] If there is a gain or loss from deregulation, does it belong wholly to the utility, or is one-third "owned" by the ratepayers? The difficulty clearly concerns the ambiguity of "ownership"; in the case of a typical corporation, one can ask whether bondholders or preferred shareholders should be deemed "owners." The answer to this semantic question should be based on the purpose for which it is asked. Here, the focus has been on incentives. For a typical corporation, the common shareholders bear residual gains and losses, and thus they are the relevant "owners" for purposes of who controls decisionmaking. For a utility, the question of who is the relevant decisionmaker has been seen

to be more complicated. Thus, the ambiguity of ownership -- which one might instinctively deem relevant to determining who should bear gains and losses -- and therefore the ambiguity of transition relief -- i.e., is it "relief" to give one-third of the gains to the consumers who have paid for one-third of the plant -- should be resolved, if incentives are the issue, in terms of how the complex decisionmaking process is believed to operate.

2.2.4. *How to Give Transition Relief.* -- As indicated at the outset, this investigation has focused on compensation and windfall taxation as mechanisms for transition relief, not independently considering alternatives such as phase-ins or grandfathering, which are aspects of some gradual deregulation proposals [Plummer 1983, Lindsay and Pfeffer 1983], because such mechanisms have similar ex ante effects and less desirable ex post effects. [See note 2.] It is not even the case, however, that all formula for compensation and windfall taxation are equally efficient. Thus, if transition mitigation is contemplated, it is necessary to consider what form it should take.

Consider, for example, Golub's (1982) suggestion that "Relegating the theoretically ideal method of lump-sum transfers to the textbooks, any practical scheme [of relief] would alter the pricing structure of the newly freed prices, potentially creating new inefficiencies." Of course, in the limit, providing full transition relief through altering newly freed prices could be accomplished by taxes and subsidies that insured that such prices would be at precisely the levels they would have been but for deregulation! It is easy enough to see that if a windfall tax is based on future prices and output, one not only creates the ex ante inefficiencies that have been the focus of this investigation, but additional ex post inefficiencies as well.

We should not be so quick to "relegat[e] . . . lump sum transfers to the textbooks" because the practical grounds for doing so are weaker here than, for example, in the optimal tax context. To provide appropriate full (or partial) windfall extraction (or compensation), one needs to estimate the difference between value with and without regulation. Regardless of the scheme one chooses, short of maintaining regulation, the greatest difficulty will be in estimating what the hypothetical future path of regulation -- in terms of prices and induced investment decisions -- would have been. Pegging windfall taxes or compensation to future prices and output only helps address uncertainty with regard to the deregulated path. But, of course, immediate changes in asset prices (many priced on securities markets) caused by deregulation will themselves be based on expectations concerning the deregulated path. There is no obvious reason that relief could not be calculated in the same manner as investors evaluate the underlying securities. Modest errors are not particularly costly, and the best information at the time is sufficient (i.e., waiting for future realizations is unnecessary) because such information also provides the basis for market movements. The benefit, of course, in addition to avoiding the administrative cost of maintaining the tax/subsidy scheme into the future, is that future incentive distortions are avoided.

2.3. Political Feasibility of Reform

A final argument sometimes used to justify transition relief in many contexts, particularly deregulation, is that compensation of losers is necessary to buy off political opposition. (Similarly, some may hold off support unless those receiving "windfalls" are duly taxed.) A closely related factor concerns whether an industry as a whole will come out in favor of or against deregulation. If the net effect of reform on the industry will be

positive, this problem might not be serious. (Indeed, one has the related problem of enacting undesirable reforms when the most influential groups stand to gain.) If most industry members lose, their opposition might be politically difficult to defeat without an appropriate payoff. The lucky intermediate case, which may have some applicability in the context of deregulating electric utility generation, would be where there were balanced forces on opposing sides -- sufficient to air the issues and present the facts (and reveal distortions of political opponents) but largely offsetting in the end, leaving political institutions to make a relatively undistorted choice. In general, such influence will often be determined by factors such as the concentration of gains and losses, depending upon one's model of the political process. [See Downs (1957), Mueller (1979), Olson (1965).]

It is particularly difficult to determine whether the use of transition relief for purposes of enhancing the political feasibility of desirable reforms is a net benefit in the long run, given the starting point of a complex, distorted, second best process. The analysis of this paper suggests a few remarks that are often overlooked. First, the direct implication of section 1 is that there are direct incentive costs to such a long-run strategy. Since such buy-offs are generally viewed exclusively as an ex post settling up of accounts, ex ante effects are by definition ignored. Second, as has been noted by some who emphasize the public choice perspective, a policy of buying out losers from deregulation creates perverse long-run (ex ante) incentives for the political process itself -- notably in terms of rent-seeking behavior. [See Kitch (1977), McKenzie (1986).] Simply put, a tradition of buying out the losers from deregulation -- or, equivalently, one might note other examples, such as the repeal of protective tariffs -- increases the expected payoff to such groups from getting such regulation --

or tariffs -- enacted in the first place. This second ex ante effect may alone be enough to cast serious doubt on the desirability of transition relief as a strategy to enhance political feasibility. Of course, it is easier to condemn some aspects of the political process than to change them.

2.4. Significance of Transition Gains and Losses

I have made no independent inquiry into the significance of transition gains and losses in the context of electric utility deregulation, although given the likely effects one can only anticipate that transfers would be large -- sometimes on the order of billions of dollars for particular utilities or regions. This subsection reports briefly and comments on the recent work on this question. [Golub 1982, Golub and Hyman 1983.]

These studies attempt to measure the net impact of deregulation, in terms of the total transfer from consumers to owners that would arise in what they term a "laissez faire" transition to deregulation -- one roughly corresponding to no transition relief. Their total estimate, based on 1980 data, is \$179 billion of implicit regulatory rents that would go to the utilities.²¹ The major components of this are the excess of replacement over book cost for utility assets and the degree to which utility common and preferred stock and bonds are below book value. It is interesting to note that while these investigations place great importance on these measurements and suggest at many points that various forms of relief are warranted, they give virtually no indication of why they believe relief might be appropriate. One can only

²¹ Note that this study includes all electric utility assets, not merely generation capacity, although since the latter are a majority of the total, the order of magnitude should be similar to that if the data had been decomposed.

surmise from a few of the passing remarks that considerations of political feasibility -- from a purely ex post perspective -- are at the core.

This large estimate may significantly overstate the effect of deregulation. First, the financial condition of electric utilities has changed overwhelmingly since the time of their estimates. (In fact, it had changed substantially even by the time their results appeared.) They found \$43 billion of rents from preferred stocks and long-term bonds having values below book, using 1980 discount rates of 13.91% and 13.46% respectively. Their calculations (Golub 1982, table 5.2) indicate that a three percentage point drop would wipe out over half of that estimate. In fact, the drop since then has been on the order of four to five points.²² Similarly, their estimates have utility stock at a value 78% of book. By 1983, however, utility stocks were at 100% of book (Bryson and Brownell 1983),²³ and they no doubt have risen well above that since then. Second, their large estimate of the gap between replacement cost and book value may well have fallen over time and, more importantly, is not a very reliable estimate of the gap between market and book value. In particular, replacement cost is an upper bound, and one might suspect that those plants most costly to replace relative to book value are often the oldest, most outmoded equipment that would not be rebuilt in a competitive environment (or a regulated one, for that matter). Finally, they have a rather small estimate for the expected benefit to utilities from including abandoned equipment in the rate base (only a \$1 billion offset).

²² As of February 5, 1987, PG&E bonds, listed by the New York Times as an exemplar of utility bonds, were at 8.95% (a drop of 4.51) and two series of Commonwealth Edison preferred were priced at a yield of 8.7% and 10.2% (a drop of 3.7 to 5.2).

²³ Bryson and Brownell note that the utilities' return on equity rose approximately two points from 1979-1983, while that for the S&P 1200 fell over five points.

Depending upon what one assumes regulatory commissions will do in the future, this figure might be far higher, although it is quite difficult to predict with any confidence.

Although their estimate of the net transfer can no longer be relied upon, their work does suggest that the magnitudes of transfers is likely to be quite large, which is the central point for considering the importance of the transition relief issue. Moreover, the authors properly emphasize that the concern is not solely with the net transfer between consumers and utilities, but rather with the magnitudes of individual transfers; their estimate was the net of presumably far greater figures that were largely offsetting. The improved financial condition of the utility industry does not suggest that there would not be significant individual gains and losses associated with deregulation, as it affects the profitability of particular plants and the prices in various regions.

3. CONCLUSION

The analysis of this paper has consisted of two primary points. First, a consistent, long-run analysis of transition relief in any context -- be it deregulation, tariffs, taxes, common law rules, or farm policy -- requires an ex ante perspective that examines effects on both incentives and risk bearing. The conclusion of both simple and more complex variants of the model is that the incentive effects determine the welfare properties, and thus, to the extent reforms being enacted are desirable,²⁴ transition relief should not be employed. Such relief distorts ex ante incentives by, in essence, making a portion of the costs and benefits of investment decisions external to the investor.

The second section examined many respects in which the ex ante perspective of the basic analysis might be seen as inappropriate -- particularly when considering a current proposal, such as the deregulation of electric utility generation. That section addressed the time consistency problem with

²⁴ This assumption has been used throughout. Note that, in the simple alternative case where a reform is assumed to be undesirable or a category of reforms being addressed are, on average, undesirable, the optimal transition policy is trivial: an infinite delay in implementation. This is not to say that the assumption here is unimportant. One obvious point of relevance in some contexts is that the nature of required transition relief might affect the behavior of political institutions determining which reforms to implement in the first place. In such instances, there would be additional efficiency effects from transition policy, and optimization would entail making an appropriate trade-off between the incentives of political actors and of private investors. For a brief discussion of these and related issues, see Kaplow (1986, 1987).

applying, ex post, a different transition policy than the one generally deemed optimal; the actual timing of electric utility deregulation and the resulting incentive effects; and the political feasibility of reform. Although all three issues require one to depart from the confines of the pure, ex ante model of section 1, in each case it was seen that the ex ante perspective offered important insights that go beyond or oppose conventional wisdom on transition relief. Moreover, most of those considerations demonstrate the importance of the government developing and maintaining a consistent, optimal transition policy, so that when future reforms are considered, many of these problems will have been favorably resolved.

REFERENCES

- Arrow, K.J., 1963, Uncertainty and the Welfare Economics of Medical Care, American Economic Review 53, 941-74.
- Bohn, R.E., R.D. Tabors, B.W. Golub, and F.C. Schweppe, 1982, Deregulating the Electric Utility Industry, MIT Energy Laboratory Technical Report No. MIT-EL 82-003.
- Bryson, J. and W. Brownell, 1983, Deregulation and the Efficiency of the Electric Power Industry, in: Plummer et al. (1983).
- Downs, A., 1957, An Economic Theory of Democracy (Harper & Row, New York).
- Golub, B.W., 1982, Deregulating the Electric Utility Industry: Financial Dislocation and Implicit Regulatory Rents, MIT Sloan School of Management Working Paper No. 1392-83.
- Golub, B.W. and L.S. Hyman, 1983, The Financial Difficulties and Consequences of Deregulation through Divestiture, Public Utilities Fortnightly, February 17, 19-25.
- Joskow, P.L. and R. Schmalensee, 1983, Markets for Power: An Analysis of Electric Utility Deregulation (MIT Press, Cambridge, Mass.).
- Kaplow, L., 1986, An Economic Analysis of Legal Transitions, Harvard Law Review 99, 509-617 (earlier version appeared as Harvard Law School Program in Law and Economics Discussion Paper No. 7).
- Kaplow, L., 1987, Optimal Transition Policy: Replacing Horizontal Equity with an Ex Ante Incentives Perspective (unpublished dissertation, Harvard University) (earlier version of relevant portions appeared as Harvard Law School Program in Law and Economics Discussion Paper No. 20).
- Kitch, E.W., 1977, Can We Buy Our Way Out of Harmful Regulation?, in: D.L. Martin and W.F. Schwartz, Deregulating American Industry: Legal and Economic Problems (Lexington Books, Lexington, Mass.).
- Kydland, F.E. and E.C. Prescott, 1977, Rules Rather than Discretion: The Inconsistency of Optimal Plans, Journal of Political Economy 85, 473-491.
- Lindsay, W. and J. Pfeffer, 1983, Potential Changes in Regulation of Bulk Power Sales by FERC, in: Plummer et al. (1983).
- McKenzie, R.B., 1986, Tax/Compensation Schemes: Misleading Advice in a Rent-Seeking Society, Public Choice 48, 189-194.

- Mueller, D.C., 1979, Public Choice (Cambridge University Press, Cambridge).
- Olson, M., 1965, The Logic of Collective Action (Harvard University Press, Cambridge).
- Pauly, M., 1974, Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection, Quarterly Journal of Economics 88, 44-62.
- Plummer, J., 1983, A Different Approach to Electricity Deregulation, in: J. Plummer, T. Ferrar, and W. Hughes, eds., Electric Power Strategic Issues (Public Utilities Reports, Arlington, Virginia).
- Shavell, S., 1979, Risk Sharing and Incentives in the Principal and Agent Relationship, Bell Journal of Economics 10, 55-73.
- Shavell, S., 1982, On Liability and Insurance, Bell Journal of Economics 13, 120-132.