

NBER Working Paper Series

TOWARD A MORE GENERAL THEORY OF REGULATION

Sam Peltzman*

Working Paper No. 133

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR
AND SOCIAL INSTITUTIONS

National Bureau of Economic Research, Inc.
204 Junipero Serra Boulevard, Stanford, CA 94305

April 1976

Preliminary; not for quotation.

NBER working papers are distributed informally and in limited number for comments only. They should not be quoted without written permission of the author.

This report has not undergone the review accorded official NBER publications; in particular, it has not yet been submitted for approval by the Board of Directors.

This study has been supported by a grant from the National Science Foundation to the National Bureau of Economic Research for research in law and economics.

*Graduate School of Business, University of Chicago, and NBER.

TOWARD A MORE GENERAL THEORY OF REGULATION*

by

Sam Peltzman
Graduate School of Business, University of Chicago
and National Bureau of Economic Research

George Stigler's work on the theory of regulation is one of those rare contributions--rare for the rest of us, though not for him--which force a fundamental change in the way important problems are analyzed. Stigler's influence will be clear in this paper. There is perhaps no more telling evidence of this influence than that its basic motivation was my dissatisfaction with some of Stigler's conclusions. (It was a dissatisfaction that Stigler shared, since I can report that we simultaneously reached one of the conclusions elaborated here--that regulatory agencies will not exclusively serve a single economic interest.) My intellectual debt to Stigler is so great that this paper emerges as an extension and generalization of his pioneering work.

What Stigler accomplished in his Theory of Economic Regulation was to crystallize a revisionism in the economic analysis of regulation that he had helped launch in his and Claire Friedland's work on electric utilities.¹ The revisionism had its genesis in a growing disenchantment

*This study has been supported by a grant from the National Science Foundation to the National Bureau of Economic Research for research in law and economics. The paper is not an official Bureau publication since it has not yet undergone the full critical review accorded Bureau publications, including review by the Bureau's Board of Directors. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

¹G. J. Stigler, *The Theory of Economic Regulation*, 2 *Bell J. of Econ. and Mgt. Science* 3 (1971) and G. J. Stigler and C. Friedland, *What Can Regulators Regulate? The Case of Electricity*, 5 *J. of Law and Econ.* 1 (1962).

with the usefulness of the traditional role of regulation in economic analysis as a deus ex machina which eliminated one or another unfortunate allocative consequence of market failure. The creeping recognition that regulation seemed seldom to actually work this way, and that it may have even engendered more resource misallocation than it cured, forced attention to the influence which the regulatory powers of the state could have on the distribution of wealth as well as on allocative efficiency. Since the political process does not usually provide the dichotomous treatment of resource allocation and wealth distribution so beloved by welfare economists, it was an easy step to seek explanation for the failure of the traditional analysis to predict the allocative effects of regulation in the dominance of political pressure for redistribution on the regulatory process. This focus on regulation as a powerful engine for redistribution shows clearly in such works as Jordan's Producer Protection and Posner's Taxation by Regulation.² The common role of regulation in this literature is as a fulcrum upon which contending interests seek to exercise leverage in their pursuit of wealth. A common, though not universal,³ conclusion has become that, as between the two main countending interests in regulatory processes, the producer interest tends to prevail over the consumer interest.

In one sense, Stigler's work provides a theoretical foundation for this "producer protection" view. However, its scope is much more general. It is ultimately a theory of the optimum size of effective political

²W. A. Jordan, Producer Protection, Prior Market Structure and the Effects of Government Regulation, 15 J. Law and Econ. 151 (1972) and R. A. Posner, Taxation by Regulation, 2 Bell J. of Econ. and Mgt. Science 22 (1971).

³Posner, op. cit., is an important exception.

coalitions set within the framework of a general model of the political process. Stigler seems to have realized that the earlier "consumer protection" model comes perilously close to treating regulation as a free good. In that model the existence of market failure is sufficient to generate a demand for regulation, though there is no mention of the mechanism that makes that demand effective. Then, in a crude reversal of Say's Law, the demand is supplied costlessly by the political process. Since the good, regulation, is not in fact free and demand for it is not automatically synthesized, Stigler sees the task of a positive economics of regulation as specifying the arguments underlying the supply and demand for regulation.

The way he does this abstracts almost completely from pure allocation questions. The essential commodity being transacted in the political market is a transfer of wealth, with constituents on the demand side and their political representatives on the supply side. Viewed in this way, the market here, as elsewhere, will distribute more of the good to those whose effective demand is highest. For Stigler, the question of which group will have the highest effective demand translates very quickly into a question of numbers. In this view, "producer protection" represents the dominance of a small group with a large per capita stake over the large group (consumers) with more diffused interests. The central question for the theory then becomes to explain this regularity of small group dominance in the regulatory process (and indeed the political process generally). The way the question is posed already foreshadows one of the results of the theory. For in Stigler's model, unlike most market models, there are many bidders, but only one is successful. There is essentially a political auction in which the high bidder receives the

right to tax the wealth of everyone else, and the theory seeks to discover why the successful bidder is a numerically compact group. The answer lies essentially in the relationship of group size to the costs of using the political process.

To summarize the argument briefly, the size of the dominant group is limited in the first instance by the absence of something like ordinary-market-dollar voting in politics. Voting is infrequent and concerned with a package of issues. In the case of a particular issue, the voter must spend resources to inform himself about its implications for his wealth and which politician is likely to stand on which side of the issue. That information cost will have to offset prospective gains, and a voter with a small per capita stake will not, therefore, incur it. In consequence the numerically large, diffuse interest group is unlikely to be an effective bidder, and a policy inimical to the interest of a numerical majority will not be automatically rejected. A second major limit on effective group size arises from costs of organization. It is not enough for the successful group to recognize its interests; it must organize to translate this interest into support for the politician who will implement it. This means not only mobilizing its own vote, but contributing resources to the support of the appropriate political party or policy: to finance campaigns, to persuade other voters to support or at least not oppose the policy or candidate, perhaps occasionally to bribe those in office. While there may be some economies of scale in this organization of support and neutralization of opposition, these must be limited. The larger the group that seeks the transfer, the narrower the base of the opposition and the greater are the per capita stakes that determine the strength of opposition, so lobbying and campaigning costs will rise faster

than group size. The cost of overcoming "free riders" will also rise faster than group size. This diseconomy of scale in providing resources then acts as another limit to the size of the group that will ultimately dominate the political process.

In sum, Stigler is asserting a law of diminishing returns to group size in politics: Beyond some point it becomes counterproductive to dilute the per capita transfer. Since the total transfer is endogenous, there is a corollary that diminishing returns apply to the transfer as well, due both to the opposition provoked by the transfer and to the demand this opposition exerts on resources to quiet it.

Stigler does not himself formalize this model, and my first task will be to do just this. My simplified formal version of his model produces a result to which Stigler gave only passing recognition, namely that the costs of using the political process limit not only the size of the dominant group but also their gains. This is at one level, a detail, which is the way Stigler treated it, but a detail with some important implications--for entry into regulation, and for the price-output structure that emerges from regulation. The main task of the paper is to derive these implications from a generalization of Stigler's model.

A Stiglerian Model of Regulation

I begin with the presumption that what is basically at stake in regulatory processes is a transfer of wealth. The transfer, as Stigler points out, will rarely be in cash, but rather in the form of a regulated price, entry restriction, etc. I shall ignore that detail here, and the resulting model applies to any political wealth redistribution. A particularization to price and entry regulation comes later. I treat the relevant

political process as if control of the relevant taxing power rests on direct voting, though this too is meant only for simplification. Though appointment of a regulatory body may lie effectively with a legislature, a committee thereof, or an executive, the electorate's receptivity to these intermediaries ought to be affected by the performance of their appointees. With Stigler, I assume that beneficiaries pay with both votes and dollars. However, again as a simplification, I assume that the productivity of the dollars to a politician lies in mitigation of opposition. A more general model might make "dollars" (broadly defined to include, for example, employment of former regulators) a source of direct as well as indirect utility to the regulator. In this model, though, direct political support-- "votes"--is the object sought directly by the regulator. More particularly he seeks to maximize net votes or a majority in his favor. There is no presumption that the marginal utility of a majority vanishes at one. Greater majorities are assumed to imply greater security of tenure, more logrolling possibilities, greater deference from legislative budget committees, etc. The crucial decision that the regulator (or would-be regulator) must make in this model is the numerical size of the group he (promises) favors, and thus implicitly the size of the group he taxes. At this stage, I retain Stigler's presumption that the agency confers benefits on a single victorious group, and the essential purpose of the model is to elaborate the limits on this group's size.

To put this formally, the regulator wants to maximize a majority M , generated by

$$(1) \quad M = n \cdot f - (N - n) \cdot h ,$$

where

n = number of potential voters in the beneficiary group

f = (net) probability that a beneficiary will grant support

N = total number of potential voters

h = (net) probability that he who is taxed (every non- n) opposes.

Note that, because both gainers and losers face transaction and information costs, f and h are not either zero or unity, but depend on the amount of the group member's gain or loss. There are similar costs facing the regulator, so he cannot exclude nonsupporting beneficiaries. At this stage, I assume that gains and losses are equal per capita within groups. This nondiscrimination assumption serves both to simplify the problem and to force Stigler's result of a single politically dominant economic interest, but the assumption is subsequently dropped. I also assume that ignorance does not lead to perverse or biased voting. If a beneficiary, for example, does not know enough to vote for his benefactor, his voting decision is not biased for or against the benefactor. Either he does not vote, or he decides how to vote by tossing a fair coin. In either case, the f in (1) will be zero, and M will be the (same) difference between votes for and votes against. With nonparticipation by the ignorant f (or h) is simply the probability that a beneficiary (or loser) votes, while with random voting by the ignorant f is the difference between the probability of a favorable and unfavorable vote by the beneficiary.

The probability of support may now be specified as

$$(2) \quad f = f(g) \quad ,$$

where, g is the per capita net benefit, and is

$$(3) \quad g = \frac{T - K - C(n)}{n},$$

with

T = total dollar amount transferred to the beneficiary group

K = dollars spent by beneficiaries in campaign funds, lobbying, etc., to mitigate opposition

$C(n)$ = cost of organizing both direct support of beneficiaries and efforts to mitigate opposition. This organization cost increases with n , but we place no restrictions on the shape of the marginal cost curve.

It is assumed that (2) holds for any subset of the electorate, in the sense that any coalition of size n faces the same costs of organization and has members with the same responsiveness to benefits. Thus, the number of votes in support depends on n in two offsetting ways: a larger n provides a broader base for support, but dilutes the net gain per member and so the probability of a member's support.

As a further simplification I assume that the regulator chooses K as well as T . The process could be modeled with the benefitted group itself determining the appropriate K , but in doing so it would be motivated by the same forces affecting a regulator who would ask K as a price for conferring the benefit. Thus, I treat it as a detail whether the beneficiaries "bid" a K and "ask" a T , or whether the regulator asks a K and bids a T .

The transfer is assumed generated by a tax at the rate t on the wealth (B) of each member outside the benefitted group, so

$$(5) \quad T = t \cdot B(N - n), \quad \text{or} \quad t = \frac{T}{B(N - n)}.$$

For application to problems of regulation, B can be thought of as a typical consumer's surplus and t a regulated price if producers are beneficiaries, or B might be a producer's surplus and t the difference between the surplus maximizing price and the regulated price where consumers are beneficiaries. At this level of generality, though, I simply treat B as a negative function of t .⁴ Opposition is assumed generated by the tax rate and mitigated by voter education expenditures per capita (z), so

$$(6) \quad h = h(t, z)$$

$$(7) \quad z = K/(N - n) .$$

In keeping with Stigler's model, I assume that, in the relevant range, benefits are subject to decreasing returns so that

$$(8) \quad f'_g > 0, \quad f'_{gg} < 0$$

(unless specified otherwise subscripts will denote partial, or where appropriate, total derivatives from here on). A complementary assumption is made for z :

$$(9) \quad h_z < 0, \quad h_{zz} > 0$$

(opposition is measured in positive units), and there are assumed to be increasing political costs to taxation:

⁴This treatment is less innocent than it appears. It implicitly rules out a "pure" transfer--i.e., one with no allocative effects. There may be forms of wealth whose supply is totally inelastic with respect to taxes, but, as a general matter, these cannot be presumed to suffice the demands of the political process--or even yield costless taxes, once tax administration and evasion costs are allowed for. The general proposition that every tax affects the wealth base being taxed has important implications for the evaluation of the whole range of government redistributive policies. See Gary Becker's comments on this article.

$$(9a) \quad h_t > 0, \quad h_{tt} > 0.$$

In this characterization of the political process, then, office holders or candidates to replace them must pick the size (n) of the group they will benefit, the amount (K) they will ask that group to spend for mitigating opposition and the amount (T) they will transfer to the beneficiary group. The necessary conditions for these choices to yield the maximum majority, the presumed goal for the office seeker, are

$$(10) \quad M_n = 0 = -(g + m)f_g + f - h_t \left(\frac{tB}{B + tB_t} \right) - h_z \cdot z + h$$

$$(11) \quad M_T = 0 = f_g - h_t \left(\frac{1}{B + tB_t} \right)$$

$$(12) \quad M_K = 0 = -f_g - h_z,$$

where

$m = C_n$, the marginal cost of group organization.

Combining (10)-(12) and making use of the definitions yields the following solution for n

$$(13) \quad \frac{n}{N} = \left[1 - \frac{f_g \cdot g}{f + h - (a - m) \cdot f_g} \right]$$

where

$a =$ average cost of organization (C/n).

Thus, if there are constant or increasing organization costs ($a - m \leq 0$), the bracketed term is less than one. We are unprepared yet to say much more than that political wealth maximization distributes benefits to a fraction of the electorate, and since we have ruled out net gains to regulation, no other result would make sense.

Before some forces affecting the size of this ratio are elaborated, it is worth dwelling on (11) for a moment. This condition--essentially that the marginal political return from a transfer must equal the marginal political cost of the associated tax--has an important subsidiary implication. Since both f_g and h_t are positive, an interior maximum can occur only if the term $(B + tB_t)$ is also positive. This term is the marginal product of t in raising revenue from a member of the losing group. That it must be positive implies that these losers must be taxed less than the interests of the winners would dictate (a revenue maximizing tax--i.e., $B + tB_t = 0$).

This result is portrayed in Figure 1. The function $R(t)$ is (h_t/f_g) . With diminishing returns in g and increasing costs in t , R_t is positive and increasing in the relevant range. The marginal revenue from t , $(B + tB_t)$, is decreasing in t , and the revenue maximizing tax is t_m where this marginal revenue is zero. However, with $R(t)$ positive at any $t > 0$, t_m cannot be a political equilibrium. The equilibrium, from (11), must occur at something like $t_a < t_m$

Thus we have an important first principle of regulation: even if a single economic interest gets all the benefits of regulation, these must be less than a perfect broker for the group would obtain. The best organized cartel will yield less to the membership if the government organizes it than if it were (could be) organized privately. This principle is independent of organization or campaigning costs, but rests on the heed the

political process must pay to marginal opposition. (Condition (11) holds even if K and C are assumed zero.) It suggests that what the "capture" literature treats as an ad hoc detail--that "the political process automatically admits powerful outsiders to the industry's councils"⁵--is in fact integral to regulatory processes. The principle also suggests that failure of regulation to maximize cartel profits need not, as Posner has suggested, arise as an efficient substitute for other forms of taxation.⁶ Even if more efficient substitutes exist and are used, a rational regulator will still tax cartel profits to secure his own position.

This logic may be pushed a step further. It will pay the rational regulator to exploit differences within the group that, taken as a whole, either wins or loses. The ability to do this may be constrained by "due process" considerations, but not typically to the point that a uniform tax must be levied or gain transferred to each member of a group. Therefore, the regulator's choice problem is not limited to selecting the appropriate size of an interest group to benefit or tax; it includes selection of an appropriate structure of benefits and costs. Once we drop the simplification of uniform taxes (prices), the identification of regulation with any single economic interest can no longer be maintained as a general proposition.

To see this, consider the following restricted problem: The regulator has decided on the total wealth that must be transferred to one

⁵Stigler, supra n. 1 at 7.

⁶Posner, supra, n. 2.

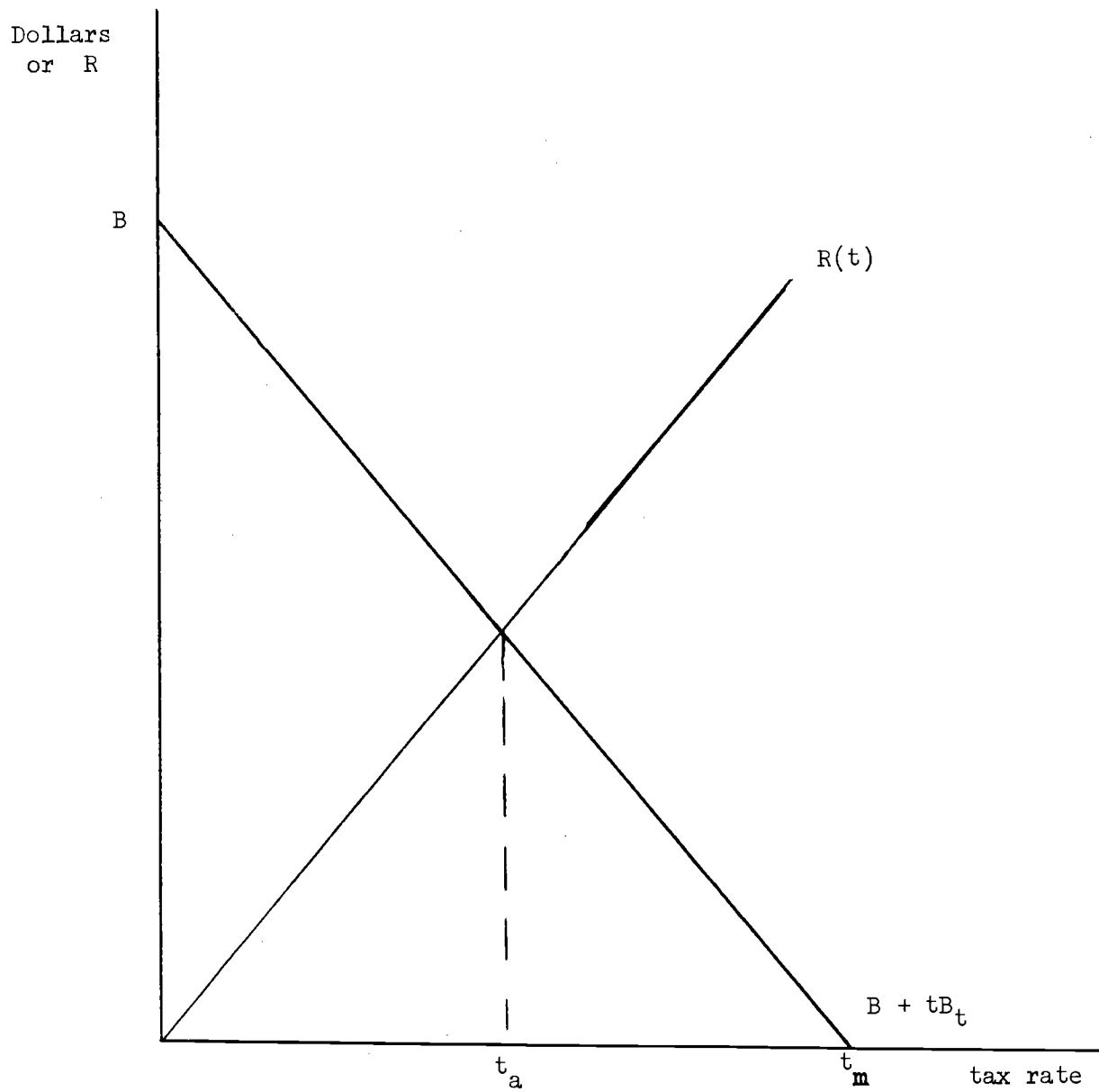


Figure 1

economic interest (say producers) from another, so that both T and n are data. However, he desires minimization of opposition (0) from consumers by exploiting differences among them in per capita wealth or the responsiveness of wealth to taxes (i.e., differences in the height and elasticity of their demands) or in their voting sensitivity to taxes. Assume that the $(N - n)$ consumers can be separated into 2 groups of size P_1 and P_2 respectively so that the last term in (1) may be written

$$(14) \quad 0 = P_1 h_1 + P_2 h_2 \quad . \quad (\text{Subscripts denote groups here.})$$

To simplify still further, treat z as fixed and equal for both groups. Minimization of (14) then involves forming the Lagrangian

$$(15) \quad L = P_1 h_1 + P_2 h_2 + \lambda(T - t_1 B_1 P_1 - t_2 B_2 P_2) \quad ,$$

where the term in parentheses is the constraint that the sum of sub group taxes is fixed, and setting the first partials with respect to t_1 , t_2 and λ equal to zero. The resulting expression for the opposition minimizing t_1 is (primes denote derivatives)

$$(16) \quad t_1 = \frac{B_2 + \frac{TB_2'}{P_2 B_2} - \frac{h_2'}{h_1'} B_1}{B_1' \left(\frac{h_2'}{h_1'} \right) + B_2' \left(\frac{P_1 B_1}{P_2 B_2} \right)} \quad .$$

The denominator is negative, but only the last two terms in the numerator are negative. This means that a negative t_1 cannot be ruled out. Thus if one group of consumers has sufficiently large per capita demand (B_2), sufficiently low demand elasticity (B_2') and tax responsiveness (h_2') relative to the other group, the latter may become part of the winning

group (get a subsidized price). On a similar argument, some producers may be taxed even if most are benefitted. The regulator's constituency thus cannot in general be limited to one economic interest.

The structure of (16) shows that t_1 is affected not only by some obvious characteristics of that group (its wealth and voting response to t_1) but also by characteristics of the other group. I shall return to this subsequently, for (16) hints at some important implications for the structure of prices emerging from regulation--e.g., that this will be the resultant of forces pushing both for and against profit-maximizing price discrimination.

I want now to return to (10)-(13) and discuss some forces affecting the size of the winning group. The Stigler model leads, after all, to more than the near truism that n/N is less than one; it more nearly asserts that the ratio is close to zero. So let us examine the effect of three variables whose importance the Stigler model asserts--support, opposition and organization costs.

In general, if x represents a variable affecting choice of n (and T and K), we want to determine the vector of total derivatives: $[dn/dx, dT/dx, dK/x]$. This can be found by solving

$$(17) \quad [M_{ij}][di/dx] = -[M_{ix}] \quad ,$$

where

$[M_{ij}]$ = matrix of cross partial derivatives, $i, j = n, T, K$

$[M_{ix}]$ = vector of the cross-partial of M_i w.r.t. x .

I now treat three simple cases:

1. A parametric shift in the support function, f , (which leaves f' unaffected). From (10)-(12) we obtain

$$(18) \quad \begin{bmatrix} M_{nf} \\ M_{Tf} \\ M_{Kf} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

and from (17) and the second order condition for a maximum M (that $[M_{ij}]$ be negative definite), we obtain the following sign conditions

$$(19) \quad \text{sign } dn/df = \text{sign } C_{nn},$$

where

$$C_{ij} = \text{cofactor of } M_{ij}.$$

Since $C_{nn} > 0$ by a second-order condition for a maximum, $dn/df > 0$ -- i.e., an increase in the probability of support for a given g increases the size of the winning group. Or, as Stigler might wish to put it, the difficulty of translating the transfer into votes leads the regulator to concentrate benefits. For the other variables we have

$$(20) \quad \text{sign } dT/df = \text{sign } C_{nT},$$

and

$$(21) \quad \text{sign } dK/df = \text{sign } C_{nK},$$

which are uncertain and negative respectively. The underlying reasons may be seen by writing out the co-factors

$$(22) \quad C_{nT} = [M_{TK} \cdot M_{nK} - M_{Tn} \cdot M_{KK}]$$

$$(23) \quad C_{nK} = [M_{Tn} \cdot M_{TK} - M_{nK} \cdot M_{TT}] .$$

$M_{TK} > 0$, because an increase in K reduces opposition and makes an increase in T more attractive. $M_{nK} < 0$, because an increase in K also dilutes the net gain, and makes concentration of the transfer on a smaller group more attractive. M_{TT} , M_{KK} are both negative, because of diminishing returns. This leaves M_{Tn} , whose sign is ambiguous: an increase in n dilutes the gain to the winners, which would induce an increase in T . But the increase in n also concentrates the opposition, and this pushes for a reduction in T . The only restriction that can be imposed (from the second-order conditions) is $(M_{Tn} + M_{Kn}) < 0$, which is enough to imply $C_{nK} < 0$ and $dK/df < 0$, but is insufficient to predict the sign of C_{nT} . If buying off a more concentrated opposition is sufficiently important to render $M_{Tn} \leq 0$, then $dT/df < 0$.

2. A parametric shift in the opposition function, h . This yields precisely the same result as a shift in support (the vector of the relevant cross-partials is the same as the right-hand side of (19)), and this symmetry between the effects of support and opposition is perhaps one of the chief insights of Stigler's model. If a more effective political support technology (a rise in f) induces a more numerous winning group, a more effective opposition technology must lead the regulator to permit a larger group to escape taxation as well. Some losers will then be made winners when there is a rise in opposition. This is better stated in the reverse. The difficulty of translating a tax into political opposition (a low h) induces the regulator to tax the many and thus to concentrate his favors on a few. Hence the filtering of information through the noise of a political

process that forces consideration of many programs simultaneously acts unambiguously, as Stigler intuited, to restrict the size of the winning group. This filtering must be done by both winners and losers, and this makes it simultaneously unattractive to spread the benefits and attractive to spread the losses over large numbers.

3. A parametric shift in the cost of organizing a group for political support. Stigler argues that the cost of organizing support (e.g., the cost of overcoming the "free rider" problem) also restricts n . However, on closer inspection, this is not obvious. Consider a rise in the $C(n)$ of (3) which, for simplicity, leaves marginal cost unchanged. Then, focusing only on dn/dC , we obtain

$$(24) \quad \text{sign } dn/dC = \text{sign}(M_{nC}C_{nn} + M_{TC}C_{Tn} + M_{KC}C_{nK}) .$$

This will be ambiguous for reasons apart from ambiguity about C_{Tn} . Stigler's argument focuses essentially on M_{nC} , which is indeed negative and induces a smaller n . However, because of diminishing returns to per capita gains, a rise in C will lead to an offsetting decrease in K ($M_{KC} < 0$). On balance, this fall in K requires a rise in n ($C_{nK} < 0$). That is, if K is reduced, restoring optimum effectiveness of lobbying and education efforts requires concentration of these efforts on a smaller group of losers. To obtain Stigler's result, one must conjecture that this sort of secondary effect is outweighed by the initial impulse to concentrate gains to offset the effect of increased organization costs.

It is well to summarize the results of this formalization of Stigler's model:

1. With a few ambiguities, the thrust of imperfect information about both the gains and losses of regulatory decisions and of costs of

organizing for political favors is to restrict the size of the winning group.

2. But this winning group will not obtain even a gross gain through political action as great as is within the power of the political process to grant it.

3. Moreover, even if groups organize according to an economic interest (producers v. consumers), political entrepreneurship will produce a coalition which admits members of the losing group into the charmed circle.

I now apply these principles specifically to price-entry regulation, and derive implications for the price-profits outcome and the demand for new regulation.

The Politics of Price-Entry Regulation

A generalization of the Stiglerian model of political transfers just discussed would be to write the politician's objective function as:

$$(25) \quad M = M(W_1, W_2) \quad ,$$

where W_i = wealth of group i , and where $M_i > 0$, but where we assume no intergroup dependencies, so that $M_{12} = 0$. This is then maximized subject to a constraint on total wealth (V):

$$(26) \quad V = W_1 + W_2 = V(W_1, W_2) \quad ,$$

where $V_i > 0$, but where $V_{12} < 0$. That is, the total wealth to be distributed is limited: market failures aside, one group's wealth can be increased only by decreasing the other's. Let us now suppose that the two groups vying to achieve benefits or mitigate losses from the political

process are consumers and producers, and that the process is constrained to provide these gains and costs through the setting of a maximum or minimum price together with control of entry. In this case, we can specialize the majority generating function (25) as

$$(27) \quad M = M(p, \pi) \quad ,$$

where

p = price of the good

π = wealth of producers, $M_p < 0$ and $M_\pi > 0$.

The implicit assumption here is that the powers of the state are sufficient to, on the one hand, enforce competition, so that any $\pi > 0$ translates into political support, and on the other, to ban sale of the good or price it out of existence, so that any consumer surplus provides some votes or stills some opposition. A somewhat more elegant, though not necessarily more insightful formulation would define (27) with respect to an anarchistic reference point. I retain the Stiglerian assumption that the political returns to higher π or lower p are diminishing ($M_{pp} < 0$, $M_{\pi\pi} < 0$).⁷

⁷ $M_{pp} < 0$ is not, of course, strictly implied by diminishing returns, and we shall see later that so strict a condition is unnecessary. If we have the simple function $M = M(S)$, where S = consumer's surplus, rather than p , and $S = \int_P^{P'} Q(p) dp$, where $Q(p)$ is the demand curve, and $Q(P') = 0$, then diminishing political returns requires $M_{SS} < 0$. However, this is related to M_{pp} by

$$M_{pp} = M_S \cdot S_{pp} + S_p^2 \cdot M_{SS}$$

where

$$S_p = -Q < 0 \quad \text{and} \quad S_{pp} = -Q_p > 0 \quad .$$

Thus M_{pp} may be positive even if $M_{SS} < 0$, but $M_{pp} < 0$ is sufficient for $M_{SS} < 0$.

I will also assume no intergroup political effects (such as envy or vindictiveness), so $M_{\pi p} = 0$. The relevant constraint here is given by cost and demand conditions, summarized by the profit function

$$(28) \quad \pi = f(p, c) \quad ,$$

where $c = c(Q)$ = production costs as a function of quantity (Q), and where over the range we shall be interested in, $f_p \geq 0$ and $f_{pp} < 0$, and, of course, $f_c < 0$. The formal problem for a successful regulator then is to maximize (I assume sufficient competition for the regulator's office) the Lagrangian

$$(29) \quad L = M(p, \pi) + \lambda(\pi - f(p, c)) \quad ,$$

with respect to p , π and λ , which yields

$$(30) \quad -\frac{M_p}{f_p} = M_{\pi} = -\lambda \quad .$$

This says that the marginal political product of a dollar of profits (M_{π}) must equal the marginal political product of a price cut ($-M_p$) that also costs a dollar of profits (f_p is the dollar profit loss per dollar price reduction). This result requires $f_p > 0$ (since $-M_p$, $M_{\pi} > 0$); which is merely a concrete application of the result in (11). That is, political equilibrium will not result in the monopoly or cartel-profit maximizing price ($f_p = 0$). The solution is shown graphically in Figure 2, where (27) is represented as a series of iso-majority curves ($M_i M_i$) obeying the assumed signs for first and second derivatives. Political equilibrium occurs at tangency (A) between the profit hill and an iso-

majority curve. On this formulation, pure "producer protection" can be rational only in the absence of any marginal consumer opposition to higher prices ($M_i M_i$ are all horizontal) and pure "consumer protection" requires no marginal support for higher profits.

This analysis says nothing about whether A in Figure 2 is anything more than trivially different from either top or bottom of the profit-hill. To make the analysis meaningful, we must either derive the appropriate political power function (the shape of the $M_i M_i$) or focus on the effects of changes in the underlying economic constraints. In the remainder of the paper I take the latter tack. That is, I set aside the question of who gets what share of the spoils to focus on the implications of the result that the spoils will in fact be shared. For example, note one implication of (30) for entry in regulation. Either naturally monopolistic or naturally competitive industries are more politically attractive to regulate than an oligopolistic hybrid. The inducement to regulate is the change in the level of $M_i M_i$ occasioned thereby. For an oligopoly with a price already intermediate between the competitive and monopoly price, the political gain from moving to A will be smaller in general than if the pre-regulation price is either at the top or bottom of the profit hill. This may help explain such phenomena as the concurrence of regulation of ostensible "natural monopolies" like railroads, utilities and telephones with that of seemingly competitive industries like trucking, airlines, taxicabs, barbers and agriculture. It may also rationalize the twin focus of antitrust on reducing concentration and protecting small businessmen, and the delay until comparatively recent times in applying the Sherman Act to less than the most concentrated industries. However, the model does not explain the dilatoriness of the government in regulating a gamut of unconcentrated retail and manufacturing markets.

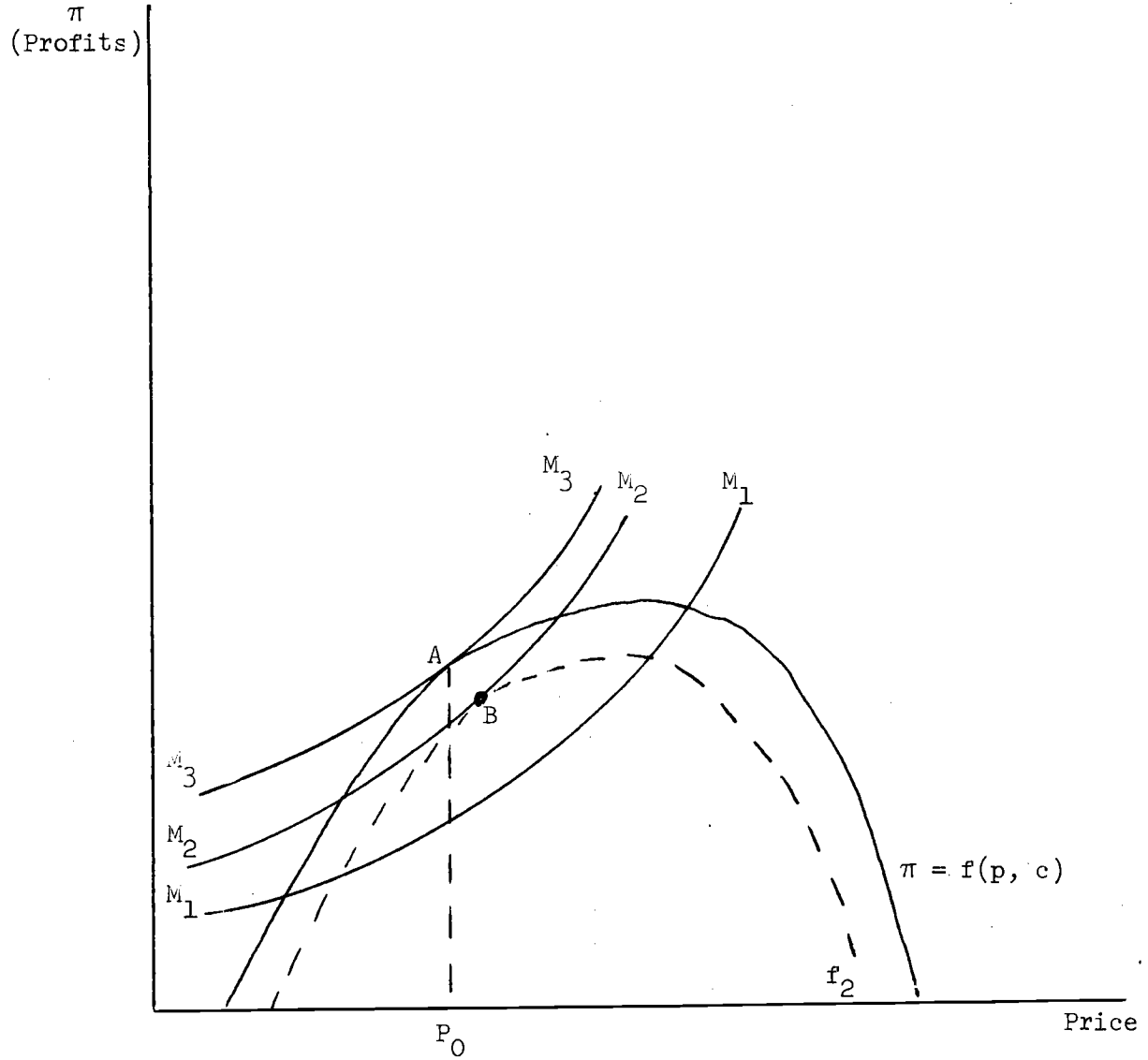


Figure 2

There is also implicit here a connection between regulation and productivity and growth. Reduction in costs or growth in demand will increase the total surplus (the height of the profit-hill in figure 2) over which a regulator might have control, and, pari passu, the political payoff for its redistribution.⁸ I have seen this point made before only in connection with welfare programs,⁹ and it deserves a systematic test. However, the association of new regulation with industries where demand and/or productivity is growing rapidly is frequent enough to be suggestive (electricity and telephones in the early 20th century, trucking and airlines in the 1930's and 1940's, natural gas in the 1950's, automobiles and drugs in the 1960's).

Some interesting implications for the pattern of regulatory choice can be derived from a more formal treatment of the interaction between productivity and growth and rational political choice. Consider a market already subject to regulation and in a political equilibrium such as A in Figure 2. Then consider the effects on this equilibrium of a parametric shift, dx , in either the cost or demand function. To obtain the effect of the shift on the p, π configuration generated by regulation, we must

⁸This is easiest to see for a constant cost competitive industry where demand increases. In that case, the no-regulation majority is unaffected by the increased demand (p and π are the same) but the gain to regulating the industry and moving to a majority maximizing (p, π) is increased. I demonstrate below that a similar result obtains for more complicated cases.

⁹See W. Allen Wallis' discussion in James Tobin and W. Allen Wallis, *Welfare Programs: An Economic Appraisal* at 54 (1968).

solve

$$(31) \quad [L_{ij}] \begin{bmatrix} dp/dx \\ d\pi/dx \\ d\lambda/dx \end{bmatrix} = -[L_{ix}] ,$$

where i, j denotes p, π or λ . In the case of a (marginal) cost shift, we obtain

$$(32) \quad \frac{dp}{dx} = \frac{-\lambda f_{px} + f_x \cdot f_p \cdot M_{\pi\pi}}{-(M_{pp} - \lambda f_{pp}) - f_p^2 M_{\pi\pi}}$$

The denominator is positive by a necessary condition for a maximum, so the sign of (32) depends on that of the numerator, which is positive.¹⁰ This is hardly surprising, since a rise in marginal cost leads to the same result without regulation. However, the insight provided by (32) is that the price increase has distinct "political" and "economic" components. The first term in the numerator $(-\lambda f_{px})$ is essentially a "substitution effect" akin to that facing an unregulated firm. A rise in marginal cost makes a higher price profitable. The second term is a "political wealth" effect: the surplus to be disposed of has shrunk, and this forces the regulator to reduce his purchases of political support. However, the usual marginal conditions familiar from consumer theory are applicable here. The regulator will, in general, not force the entire adjustment

¹⁰ $\lambda < 0$, from (30); $f_x = -c_x < 0$; $f_p > 0$, since profits are below a maximum; $M_{\pi\pi} < 0$ by assumption; and

$$f_{px} = -Q_p^c Q_x > 0 .$$

onto one group. In particular, consumers will be called on to buffer some of the producer losses. To see this more clearly, abstract from the substitution effect by assuming a change in fixed cost only, so $f_{px} = 0$. Then the profit hill in Figure 2 shifts down by a constant to f_2 , leaving the profit-maximizing price unchanged, but increasing the political-equilibrium price and buffering the fall in profits that would otherwise occur. Of course, as is the case in consumer choice, one cannot rule out "inferiority" of price decreases or profit increases.¹¹ But the "normal" purely political component of the response to cost changes involves consumers shielding producers from some of the effects of cost increases and producers sharing some of their gains from cost reductions.

¹¹Such inferiority is in fact essentially ruled out here by the absence of intergroup dependencies. This plays the same role here as utility independence does in ruling out inferior goods in consumer choice theory. The closest analogy to the conventional consumer choice problem would be where the regulator always sets a marginal price equal to (a constant) marginal cost and then merely allocates the resulting surplus among producers and consumers by fashioning a suitable two-part or declining-marginal price scheme. In this case, the surplus is the regulator's "income" which can be used to purchase the "goods" producer or consumer support at a price of \$1. If the utility (votes) of the two goods is independent, declining marginal utility will assure that both are normal.

This analogue helps illuminate the attraction of regulation to markets with growing productivity and demand. The increased surplus, which is the regulator's income, generates a larger utility (vote) gain from moving from either corner (monopoly or competition) to the vote maximum, again so long as there are diminishing political returns to both producer and consumer wealth.

The case of a shift in demand is more complex, because the demand function enters indirectly into the M function: M_p depends on the relationship between price and consumer surplus, which depends on the height of demand. Formally, a change in demand, dy , yields

$$(33) \quad \frac{dp}{dy} = \frac{-\lambda f_{py} + M_{py} + f_y \cdot f_p M_{pp}}{-(M_{pp} - \lambda f_{pp}) - f_y^2 M_{pp}}$$

Again, the first term of the numerator is a profit-maximizing "substitution" effect which is positive,¹² and the last term a political wealth effect which is, in this case negative ($f_y > 0$). The middle term represents the effect of the demand shift on political "tastes"--i.e., on the slope of the $M_i M_i$ in Figure 2, but this effect is ambiguous.¹³ For example, if a rise in consumer income raises the payoff to price reductions, $M_{py} < 0$, and the political-wealth effect is reinforced. Ignoring this taste change, the results are symmetric with those of a cost change. Consider a rise in demand such that $f_{py} = 0$.¹⁴ The political wealth effect

¹² Ignoring, as usual, any offsetting changes in the slope of demand.

¹³ In particular

$$M_p = M_S \cdot S_p = -QM_S$$

where S again denotes the underlying consumer surplus. So

$$M_{py} = -QM_{SS}S_y - M_S Q_y$$

Since M_S , Q_y and $S_y > 0$, while $M_{SS} < 0$, the sign of M_{py} is ambiguous.

¹⁴ This requires an appropriate change in the slope of the demand curve, since

$$f_{py} = (P - C_Q) \cdot Q_{py} + Q_y$$

So some $Q_{py} < 0$ is required for $f_{py} = 0$.

will nevertheless induce a price reduction because the diminishing political returns to both profit increases and price decreases make a combination of the two the best strategy for political "spending" of more wealth.

What emerges from this discussion is more a working hypothesis than an a priori conclusion about the nature of price and profit adjustment under regulation. If the political wealth effect is empirically important, it will be manifested in attenuation of price changes when demand changes, and their amplification when costs change; vice versa for profit changes. In the case of the latter, the wealth-effect components of the counterparts to (32) and (33) may be written

$$(34) \quad \frac{d\pi}{dx} = \frac{f_x}{1 + f_p^2 \left(\frac{M_{\pi\pi}}{M_{pp} - \lambda f_{pp}} \right)}$$

$$(35) \quad \frac{d\pi}{dy} = \frac{f_y}{1 + f_p^2 \left(\frac{M_{\pi\pi}}{M_{pp} - \lambda f_{pp}} \right)}$$

These are both smaller absolutely than what would obtain under pure producer protection (which yields simply f_x or f_y). We can then summarize the interaction between cost and demand changes and regulatory utility maximization as follows: Define variables π' and p' as the difference between regulated and profit maximizing profits and prices respectively. The purely political effects of changes in underlying economic conditions are then for dp'/dx and $d\pi'/dx > 0$; dp'/dy and $d\pi'/dy < 0$. Among the empirical implications of these forces would be:

1. Regulation will tend to be more heavily weighted toward "producer protection" in depressions and toward "consumer protection" in

expansions. Thus, for example, it is not useful to view events like the Robinson-Patman Act and the NRA as "inconsistent" with the intent of anti-trust legislation; this intent is endogenous. Similar arguments apply to the structure of taxes (the corporate-personal tax mix should offset changes in the share of GNP earned by capital), tariffs (more free trade when demand grows or costs fall), etc.

2. Government intervention and regulation are both normal goods. Though this generalization has exceptions, the difference between the no-regulation iso-majority curve and the regulatory equilibrium (i.e., the incentive to regulate) grows with the level of demand. As a further generalization, the income elasticity of producer protection ought to be less than that of consumer protection. This follows from the negative wealth effect of demand growth on equilibrium price, which makes for an increased consumer share of the total surplus as demand (income) increases.

3. The tendency of regulation to change prices infrequently, sometimes called "regulatory lag" ought to be stronger when demand changes than when costs change. This follows from the opposing wealth and substitution effects in the case of a shift in demand (but not in the case of a cost change). Here failure to change a price can be interpreted to mean that the opposing effects offset one another.

4. Some reexamination of studies, such as Stigler and Friedland's, which show regulation to be ineffective is called for. In the first place the result ought to be sensitive to the dynamics of supply and demand. In a growing, technologically progressive industry, producer protection ought to yield to consumer protection over time, even if, on average, there is no effect. (Stigler and Friedland's data do show some secular trend toward

lower prices).¹⁵ Secondly, deviations about the zero mean effect should be systematic: high cost, low demand markets will have prices elevated by regulation and low cost high demand markets will have prices reduced. Finally, as a generalization of (2), entry of regulation is not exogenous. It should occur first in the low cost high demand markets. This last point indicates some of the complexity engendered by the interaction of the static and dynamic aspects of the model: whether entry of regulation into any market raises or lowers prices depends on whether the market was initially competitive or monopolistic. Once that initial adjustment has been made, subsequent cost and demand changes will govern any redistribution from the initial position.

5. If regulation is evaluated against a zero-profit (fair rate of return) benchmark, we might be tempted to conclude that positive profits imply a "captured" regulator and thereby expect a positive correlation between prices and profitability. In fact the observed correlation ought to be negative. Whatever its source--increased demand or lower costs--an increase in the profit hill of Figure 2 generates a political incentive to move toward a combination involving higher profits and lower prices. Thus, quite apart from any private profit maximizing incentives toward this configuration, the most profitable regulated firms ought to have the lowest prices. More precisely, the gap between the profit maximizing and regulated price will be positively correlated with the gap between the former and the "fair-rate-of-return" price.

¹⁵ Stigler and Friedland, *What Can Regulators Regulate?*, p. 7. Their estimate is that regulation had no effect on electricity rates in 1912, and lowered prices by about 10 percent in 1937.

6. The model also yields predictions on the bias of regulation. Briefly, elastic demand and economies of scale create a bias favorable to consumers. The reason is that these sorts of demand and cost conditions enhance the consumer surplus gained while mitigating the producer surplus lost due to a price reduction. To see this formally, first introduce a parameter, w , into the slope of the demand curve at equilibrium, so that a positive dw implies a less elastic demand. By appropriate reformulation of the right-hand side of (31), we obtain the vector

$$(36) \quad - \begin{bmatrix} L_{pw} \\ L_{\pi w} \\ L_{\lambda w} \end{bmatrix} = - \begin{bmatrix} M_{pw} - \lambda f_{pw} \\ M_{\pi w} \\ -f_w \end{bmatrix} = - \begin{bmatrix} M_{pw} - \lambda f_{pw} \\ 0 \\ 0 \end{bmatrix} ,$$

where we set $f_w = 0$ (i.e., assume that the less elastic demand passes through the initial price quantity combination). Both M_{pw} and $-\lambda f_{pw}$ are positive: A less elastic demand reduces the consumer surplus and vote productivity of a price reduction, while it enhances the profitability and vote productivity of a price increase.¹⁶ The signs of the relevant total derivatives, then become

¹⁶ Again starting from $M_p = -QM_S$, we get $M_{pw} = -QM_{SS} \cdot S_w$, and $S_w > 0$, so with $M_{SS} < 0$, $M_{pw} > 0$. Since

$$f_p = (P - C_Q)(Q_p + w) + Q$$

$$f_{pw} = (P - C_Q) > 0 ,$$

and with $\lambda < 0$, $-\lambda f_{pw} > 0$.

$$(37) \quad \text{sign } dP/dw = \text{sign } L_{pw} > 0$$

$$(38) \quad \text{sign } d\pi/dw = \text{sign } f'_p \cdot L_{pw} > 0 .$$

That is, a less elastic demand induces the regulator to "relocate" toward the northeast on any iso-majority curve in Figure 2.

For the scale-economies case, introduce a parameter, v , into marginal cost and assume that a negative dv leaves profits at the old equilibrium unchanged. That is, if there is a lower marginal cost in the neighborhood of equilibrium, it is sufficiently higher at lower outputs to leave total costs unchanged. This sort of characterization of increased scale economies implies the vector

$$(39) \quad - \begin{bmatrix} L_{pv} \\ L_{\pi v} \\ L_{\lambda v} \end{bmatrix} = - \begin{bmatrix} M_{pv} - \lambda F_{py} \\ M_{\pi v} \\ -f'_v \end{bmatrix} = - \begin{bmatrix} \lambda Q_p \\ 0 \\ 0 \end{bmatrix} .$$

The term, λQ_p , is positive. The diseconomies of smaller outputs when $dv < 0$ make a price increase less profitable (and so a price decrease more attractive politically). This renders the derivatives, dp/dv and $d\pi/dv$, both positive, so more scale economies induce a move to the southwest on any iso-majority curve.

Pending a systematic test of the empirical relevance of these propositions, I point out potential pitfalls. The long history of "producer" regulation of agriculture (price supports, marketing restrictions, etc.) seems consistent with the model, given the conventional

wisdom about low supply and demand elasticities in this sector. However, the cartelization of airlines, trucking, railroads, and taxicabs where there are either constant or decreasing costs is obviously troublesome. A more general problem is how to distinguish the political incentives here from corresponding profit-maximizing incentives which push in the same direction, if we want to use the result to predict the behavior of established regulators rather than the entry pattern in regulation.¹⁷

7. Finally, I note an implication for the theory of finance. Regulation should reduce conventional measures of owner risk. By buffering the firm against demand and cost changes, the variability of profits (and stock prices) should be lower than otherwise. To the extent that the cost and demand changes are economy-wide, regulation should reduce systematic as well as diversifiable risk.

There is no obvious risk pattern among currently regulated firms: electric, gas and telephone utility stocks rank among the least risky while airline stocks are among the most risky. However, in one case of new regulation (of product quality), I found that both total and systematic risk of drug stocks decreased substantially after regulation.¹⁸ A crude test on railroad and utility stock prices shows the same pattern, though

¹⁷As an example of the kind of entry pattern that can be predicted, consider a competitive industry with inelastic demand and supply. The political equilibrium here is closer to the monopoly equilibrium than it is with elastic demand and supply. Hence such an industry is more likely to attract regulation than one with elastic demand and supply. Similarly a natural monopoly with elastic demand and supply makes an inviting target for regulation.

¹⁸See S. Peltzman, The Benefits and Costs of New Drug Regulation, in R. Landau, ed., *Regulating New Drugs*, 1973 at 205-206.

the effect is weak. I correlated annual (December to December) changes in the log of the Standard and Poor's or Cowles indexes of railroad and utility stock price indices¹⁹ with those of the industrial index (which I treat as a diversified portfolio of stocks of unregulated firms) for equal periods spanning the onset of regulation. I took 1887 as the first year of railroad regulation, and 1907 as the start of utility regulation. (New York began regulating that year.) The indexes of systematic risk (estimated as the regression coefficient on industrial stock price changes) were, with standard errors in parentheses,

	<u>Before Regulation</u>	<u>After Regulation</u>
Railroads (1871-86, 1887-1902)	.74 (.24)	.56 (.17)
Utilities (1871-1906, 1907-42)	.67 (.12)	.60 (.10)

The total risk of these stocks relative to industrials (the ratio of standard deviations of annual changes) were.

	<u>Before Regulation</u>	<u>After Regulation</u>
Rails	1.16	.85
Utilities	.97	.84

All of the differences go in the right direction, but none are significant. The main point of this exercise is simply to hint what further research might be useful.

¹⁹Standard and Poor's Corp., Security Price Index record (1971).

The Structure of Regulated Prices

I have argued that the rational regulator will not levy a uniform tax nor distribute benefits equally. Rather, he will seek a structure of costs and benefits that maximizes political returns. This search for political advantage will in turn lead the regulator to suppress some economic forces that might otherwise affect the price structure. For example, the cost of serving a group of customers or their elasticity of demand will have a different impact under regulation than it will in an unregulated market because of the absence of political constraints in the latter case. The substitution of political for economic criteria in the price formulation process has several interesting implications which I shall elaborate. It is at the heart of the pervasive tendency of regulation to engage in cross-subsidization--i.e., the dissipation of producer rents on sales to some customers by setting below cost prices to others. We shall see that this cross-subsidization follows a systematic pattern in which high cost customer groups are subsidized by low cost customers. Further, this pattern of price discrimination emerges from a process in which conventional profit maximizing price discrimination as well as other economic forces leading to price differences are attenuated.

A convenient starting point for this analysis is the problem first set out in equations (14)-(16), where the regulator seeks a tax structure to minimize opposition. Here I want to consider the effect on the resulting tax structure when a change occurs of the type that would ordinarily lead the gainers to seek a change in only one of the two tax rates. As an example, suppose per capita wealth rises for one group only. In the price regulation analogue to this problem, this would lead to a rise in one group's demand, and a profit maximizing monopolist might then raise

that group's price, but not the other group's price. Under regulation, however, no such specialization of a tax increase will be tolerated, because this would violate the basic principle that opposition from the two groups must be equated at the margin.

This point can be demonstrated formally with the same framework used previously. Specifically let there be a parameter shift, dx , in the wealth of group 1 only. Then trace the effects of this shift on t_1 and t_2 . These effects are obtained by solving

$$(40) \quad \begin{bmatrix} dt_1/dx \\ dt_2/dx \\ d\lambda/dx \end{bmatrix} = - [L_{1x}, L_{2x}, L_{\lambda x}] \begin{bmatrix} L_{11} & L_{12} & L_{1\lambda} \\ L_{21} & L_{22} & L_{2\lambda} \\ L_{\lambda 1} & L_{\lambda 2} & L_{\lambda\lambda} \end{bmatrix}^{-1}$$

where the subscripts 1, 2 on the r.h.s. refer to t_1 and t_2 . This has the following relevant solutions

$$(41) \quad \text{sign} \frac{dt_1}{dx} = \text{sign}[-L_{1x} \cdot L_{2\lambda}^2 - L_{\lambda x} \cdot L_{22} \cdot L_{\lambda 1}]$$

$$(42) \quad \text{sign} \frac{dt_2}{dx} = \text{sign}[L_{1x} \cdot L_{\lambda 1} \cdot L_{2\lambda} - L_{\lambda x} \cdot L_{11} \cdot L_{\lambda 2}]$$

The sign of (41) is ambiguous, since the first term in brackets is positive while the second is negative. The first term reflects the ability of the regulator to both maintain revenues and limit opposition by raising taxes on the now wealthier group 1 individuals, while the second term is a political wealth effect which induces lower tax rates. The more interesting result is that the sign of (42) is unambiguously negative. This occurs

first because of the incentive to substitute higher taxes on group 1, which creates the ambiguity in (41) and which in (42) requires an offsetting decrease in t_2 to maintain equilibrium. This incentive to a lower t_2 is reinforced by the political wealth effect. The analysis assumes no interdependencies between the two groups' political responsiveness or wealth (i.e., L_{12} is assumed to be zero). Thus what emerges here is that the regulator's striving for minimum opposition by equating opposition at the margin leads him to spread effects of economic forces which are local to all groups. This common element in the tax structure is provided by the wealth effect which leads the regulator to buy more of both relevant "goods" (less opposition from group 1 and from group 2).

This result can be applied to the regulation of prices by suitably generalizing the analysis of a single price summarized in Figure 2. That is, assume that there are two separable groups of buyers, so that the majority generating function (27) is

$$(43) \quad M = M(p_1, p_2, \pi) \quad ,$$

with $M_1, M_2 < 0$. The distinction between the two groups is economic rather than political, in that I assume only that there are cost and/or demand differences. Thus customers whom the regulator might wish to single out for benefits can be scattered among both groups, and p_1 and p_2 can be regarded as averages from another price structure conditioned by political forces. I suppress this structure here only to highlight the difference between a regulated and unregulated market's response to common economic forces. The cost/demand differences also give rise to the new profit function

$$(44) \quad \pi = f(p_1, p_2, c) \quad .$$

With no loss of generality, I assume that it costs nothing to produce the product for group 2, so $c =$ cost of production for group 1. Otherwise the properties of (44) and its simpler counterpart (28) are the same ($f_1, f_2 \geq 0, f_{11}, f_{22} < 0, f_c < 0$). Again, to make the problem non-trivial, I rule out cross-group effects, so

$$M_{12} = M_{1\pi} = M_{2\pi} = f_{12} = 0 \quad .$$

We may now proceed to trace out the implications for the structure of regulated prices if there is a change of the sort that would lead, in an unregulated market, to a change solely of one group's price. As an example, let group 1's demand increase, so that, with independent demands and costs, the profit maximizing or short-run competitive price would rise for that group alone. The general problem now facing the regulator is to choose the set (p_1, p_2, π, λ) which maximizes the Lagrangian

$$(45) \quad L = M(p_1, p_2, \pi) + \lambda(\pi - f(p_1, p_2, c)) \quad .$$

(Note that we are dropping the restriction in equations (14)-(16) and (40)-(42) of fixed "tax receipts"--here profits--transferred to winners.) The first-order conditions for a maximum here are similar to (30); specifically

$$(46) \quad -\frac{M_1}{f_1} = -\frac{M_2}{f_2} = M_\pi = -\lambda \quad .$$

So both p_1 and p_2 will be held below the profit-maximizing level

($f_1, f_2 > 0$). Now let there be a parameter shift, dy , in group 1's demand, and let us see what effect this has on p_1 and p_2 . Consequently, we solve

$$(47) \quad \begin{bmatrix} \frac{d_i}{d_y} \end{bmatrix} = -[L_{iy}][L_{ij}]^{-1} ,$$

where $i, j = p_1, p_2, \pi, \lambda$. The left hand side of (47) is a vector of total derivatives; the first term on the right is a row-vector of partial derivatives, and the second term is a matrix of partial derivatives. To present the results in a manageable fashion, I define the following variables, and indicate their signs:

$$A = [(\lambda F_{22} - M_{22}) - f_2^2 \cdot M_{\pi\pi}] > 0$$

(by second-order conditions for a maximum);

$$B = f_1 \cdot M_{\pi\pi} \cdot (M_{22} - \lambda f_{22}) > 0$$

(by second order conditions and $f_1 > 0$);

$$C = f_2 \cdot M_{\pi\pi} (M_{11} - \lambda f_{11}) > 0$$

(by second-order conditions and $f_2 > 0$). I then show the results for the signs of dP_1/dy and dP_2/dy by components.

$$(48) \quad \text{sign} \frac{dP_1}{dy} = \text{sign:}$$

$$\begin{aligned} M_{ly} \cdot A? & \quad (\text{"taste" shift}) \\ -\lambda f_{ly} \cdot A > 0 & \quad (\text{"substitution"}) \\ -f_y \cdot B < 0 & \quad (\text{"political wealth"}) \end{aligned}$$

$$(49) \quad \text{sign} \frac{dP_2}{dy} = \text{sign:}$$

$$\begin{aligned} -M_{ly} \cdot f_2? & \quad (\text{taste shift}) \\ +\lambda f_{ly} \cdot f_2 < 0 & \quad (\text{substitution}) \\ -f_y \cdot C < 0 & \quad (\text{political wealth}) \end{aligned}$$

The results in (48) are similar to those in (33), where we analyzed the effects of a shift in demand on a single price. There is a change in consumer surplus with ambiguous effects on the responsiveness of group 1 to price reductions (i.e., its "tastes" for price reductions). There is a substitution effect, showing that it is "cheaper" for the regulator to collect transfers in the form of higher prices to the higher demand group. Finally there is a political wealth effect, showing that the regulator will use the expanded opportunity locus to shield group 1 from the full substitution effect.

The more interesting result is (49), since group 2 would be unaffected in an unregulated market. Apart from the ambiguous "taste" effect, there are two forces under regulation leading this group to benefit from the higher demand of group 1. First, there is the converse

of the substitution effect. If it is now more attractive to tax group 1, then for any given tax receipt, the price to group 2 will be lower. Second, there is the same wealth effect that assists group 1. The regulator distributes the gains made possible by the higher demand partly in the form of higher profits,²⁰ partly in the form of a lower price to group 1 and partly in the form of a lower price to group 2. All the margins in (46), not just one or two, require adjustment when one group's demand increases and thereby increases the wealth available to the regulator.

This result is illustrated in Figure 3, where I focus on the structure of prices. Each of the curves labeled M_i is a locus of price combinations consistent with a constant level of support or opposition from consumers. These are negatively sloped, indicating that the regulator can maintain the fixed support level by trading lower prices to one group for higher prices to another. The M index increases toward the origin, since lower prices are preferred by both groups. For simplicity, I assume diminishing political returns to price reduction, so the M are convex from above. The point A is the combination of profit maximizing prices, but the rational regulator wishes to set lower prices than these. The frontier DGC shows the p_1, p_2 combinations which yield the desired level of producer wealth. It is negatively sloped since f_1 and f_2 are both positive (or zero at D and C respectively), and concave from above, since both f_{11} and f_{22} are negative. The

²⁰The result for the wealth component of $d\pi/dy$ is a more complex analogue to (35) with the same properties.

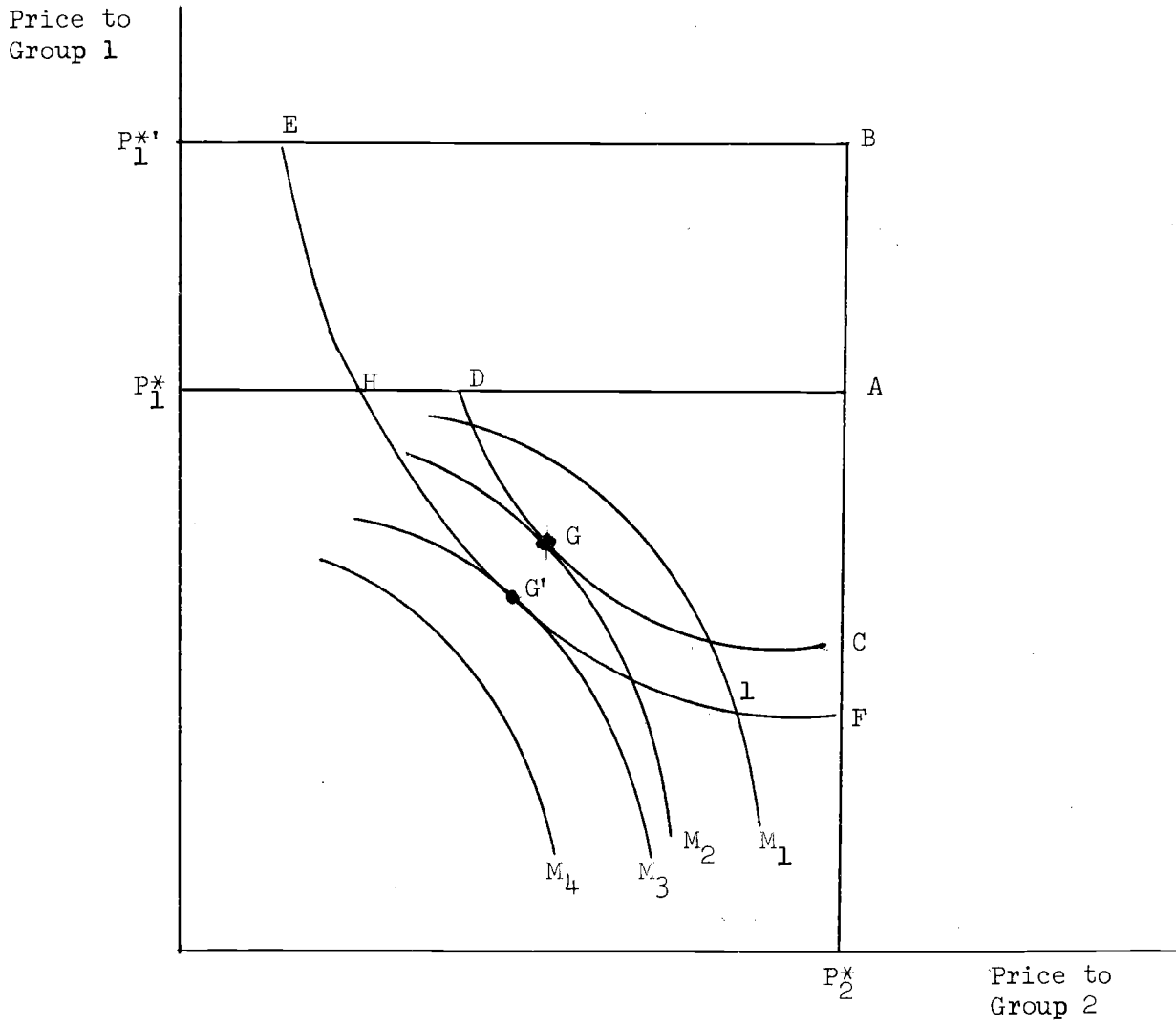


Figure 3

equilibrium at G is defined by the first two conditions in (46). Consider now the special case where the regulator desires to keep profits fixed and the group 1 demand increases. If p_1 at G exceeds marginal cost, the profit frontier will shift outward over a range of prices in the neighborhood of G . That is, with 1's higher demand, the same profit can be generated by a lower p_1 holding p_2 constant, or by a lower p_2 holding p_1 constant. (For simplicity, I have assumed that the p_1 at C also exceeds marginal cost, so that the frontier shifts out over the entire relevant range.) It is this shift to $EHG'F$ that produces a "wealth effect" toward a lower p_1, p_2 set, though there will also be a change in the slope of the frontier which will offset the incentive toward a lower p_1 .

The implication here is that, not only will the average level of prices under regulation be below what it would be in pure monopoly, but the structure of relative prices will depart from that in either pure monopoly or competition. The important contribution of politics is to suppress economically important distinctions and substitute for these a common element in all prices. On the demand side, this means that regulators will tax profits by attenuating profitable price discrimination. Discrimination is not eliminated, because there is a force--the substitution effect--unifying the interests of a discriminating monopoly and the regulator.²¹ It is countered by the wealth effect, so the empirical importance

²¹In the case of a pure change in 1's elasticity of demand--i.e., a change in the slope but not the height of demand--the relevant total derivatives of P_1 and P_2 are opposite, because only a substitution effect is at work.

of this effect will determine that of the unique political effect on the price structure. Equations (48) and (49) do shed this further light: the term, f_y , is proportional to the difference between price and marginal cost. So, the political element in pricing should be most prominent the more profitable the regulated firm.

Except that this last result does not hold, the case of a change in costs is similar to that of a change in demand. Specifically, a rise in the marginal cost of serving group 1 leads in addition to the conventional substitution effect raising p_1/p_2 , to a wealth effect raising both p_1 and p_2 .

This incentive to reward or tax all customers for the peculiar characteristics of some has interesting implications for the structure of regulated prices. Not only will profit-maximizing price discrimination be discouraged, but a peculiar form of price discrimination will replace it. This is usually referred to as "cross-subsidization" and, to the extent that this is not just another name for ordinary price discrimination, it connotes a structure in which an unprofitably low price for some is paid for from profits on sales to others. This sort of phenomenon seems difficult to reconcile with the producer protection view of regulation. Why, after all, would a surface transportation cartel wish to perpetuate unprofitable passenger train or short haul rail freight service? So far such questions have received no satisfactory answer, and the phenomenon tends to be viewed as "a process of ad hoc pacification" of vocal consumer groups.²² Our model suggests that the process is in fact systematic:

²²George Hilton, The Basic Behavior of Regulatory Commissions, 62 Am. Econ. Rev. 49 (May, 1972).

holding demand constant, the higher cost customers will receive the lower price-marginal cost ratios. Their peculiarly high costs will be spread among all customer groups by a rational regulator. Thus we need not appeal to ad hoc judgments about the political power of, say, train passengers or short haul freight users to explain the pattern of cross-subsidization. Instead, the model implies that we should observe either a higher level of costs (say for short hauls compared to long hauls) or more rapid increases in costs (for passengers compared to freight) for the subsidized group. More generally, the model sheds light on the tendency of regulation to produce rate "averaging" across dissimilar customer groups--e.g., charging similar electricity rates to rural and urban customers (which benefits the former) or similar auto insurance rates to rural and urban customers (which benefits the latter). The common element in these price structures is their suppression of cost differences.

I used this sort of model to rationalize differences in the price structure under government ownership and regulation. This required an assumption that purely political forces will be more prominent in the former regime.²³ It will take further empirical work to show whether the political impulse to uniform treatment of customers also affects regulated rates systematically. I can illustrate some of the promise and pitfalls by application to the airline rate structure. Keeler estimated price-marginal cost ratios for standard coach service in 29 regulated city-pair markets as of 1968.²⁴ He found that the most prominent cost difference in airline

²³S. Peltzman, Pricing in Public and Private Enterprises: Electric Utilities in the United States, 14 J. Law and Econ. (1971).

²⁴T. E. Keeler, Airline Regulation and Market Performance 3 Bell J. of Econ. and Mgt. Science 2 (1972).

service is distance related. Since major elements of cost are constant per flight, the per mile marginal cost falls continuously with a flight's distance. My model would imply that effective CAB regulation would convert this cost structure into a price structure whereby price/marginal cost rises continuously with distance--i.e., the fare-distance taper would be less severe than the cost-distance taper. One immediate problem is that profit maximizing discrimination would imply a similar price structure, since ground alternatives are more competitive over shorter distances. However, especially for standard coach service, where individual business travel tends to predominate over family and vacation travel (for which airlines offer discounts), the viability of ground alternatives is restricted. Gronau estimates that, for plausible values of time, airlines will essentially monopolize the relevant market for distances over 600 miles.²⁵ This implies that a profit-maximizing fare structure would have price/marginal cost ratios rising substantially more sharply with distance up to 600 miles than beyond. My model implies no such break, or at least a continual increase in this ratio in the over 600 mile segments.

To sort these forces out, I regressed the log of Keeler's estimate of price/marginal cost ($P - MC$) on two distance variables: the log of distance if the city pair is less than 600 miles apart and zero (i.e., one mile) otherwise (D_1) and log of distance if the distance exceeds 600 miles, zero otherwise (D_2). From Gronau's results, profit maximization implies that the coefficient of D_1 is positive, while that of D_2 is zero. Political support maximization implies that both coefficients are

²⁵R. Gronau, The Effect of Traveling Time on the Demand for Passenger Transportation, 78 J. Pol. Econ. 2 (1970).

positive, and, in the extreme, equal. The result is

$$P - MC = -.66 + .17D_1 + .17D_2 \\ (3.71)^1 \quad (4.49)^2$$

$$R^2 = .69 \quad \text{S.E.} \times 100 = 8.46$$

(t-ratios in parentheses). If the log of per-mile cost is regressed on D_1 and D_2 , the corresponding coefficients are both $-.26$. This association of a continuous increase in $P - MC$ with a continuous distance economy is strong support for the political support maximization model against simple profit maximization. The CAB essentially ignores the strength of ground competition for a particular flight, and simply spreads the same part (about 2/3) of any flight's distance related economy among all fares.

Now the pitfall: Keeler has recently updated his cost estimates to 1974.²⁶ There has been no important change in airline technology: per-mile costs still fall continuously with distance (the 1974 elasticity is $-.22$). There has been, though, a major change in the fare structure. For the 1974 data, the $P - MC$, distance relationship is

$$P - MC = .41 - .01 D_1 - .01 D_2 \\ (.23)^1 \quad (.68)^2$$

$$R^2 = .33 \quad \text{S.E.} \times 100 = 4.17.$$

The CAB has recently espoused the desirability of cost-based fares, and, more importantly, it has implemented them: the fare - and cost-distance gradients

²⁶ See Civil Aeronautics Board Practices and Procedures, U.S. Senate Committee on the Judiciary, 94th Cong., 1st Sess., 58 (1975).

are now essentially identical. To get there, the CAB has permitted fares on the longest flights in the sample to rise by under 30 per cent between 1968 and 1975, while those on the shortest have more than doubled. By 1974, much of the price-discrimination, at least on coach service, had vanished.²⁷

This implies that the CAB has been sacrificing producer and, in terms of my model, political wealth to the ghost of Pareto. I will not pretend that my model offers any insight into this recent behavior, however well it seems to explain matters up to 1968.²⁸ Perhaps, though, it does help explain recent

²⁷ The range of the P - MC variable was .47 in 1968 and .16 in 1974.

²⁸ The promise and pitfalls of the model are also illustrated by surface freight rates. The cost structure here is similar to air--a negative cost/mile-distance taper. This is most pronounced for rails, and they have experienced the most profound effects of the resulting political incentives: short-haul rates sometimes below marginal cost, regulatory inhibitions on elimination of such services and, recently, bankruptcies among short-haul specialists. This all appears consistent with the basic model, except that a simple extension should have firms and consumers treated similarly. That is, the firms in this industry happen to be crudely separable by an economic criterion--average length of freight haul. Maximization of political support from producers would then appear to require spreading some of the profit effects of high cost short-haul service to the long-haul specialists. Indeed the ICC has the power to do this by regulating divisions of joint rates. However, it has obviously not been sufficiently diligent in its use of the power to prevent striking differences in the prosperity of long- and short-haul specialists; differences which appear superficially greater than those that might be expected without regulation of rates and exit. This suggests two problems: (1) Why are the ICC's incentives to weld a coalition so much stronger in the case of consumers than producers? (2) What accounts for the difference between the ICC and CAB willingness to endanger the consumer coalition by permitting economic efficiency criteria to intrude in the rate structure?

There is finally a problem of appropriate units. A prime example of cost-based cross subsidization is first-class postage. The rate here ignores distance-related costs entirely, and so results in price/marginal cost declining with distance. The model can only hint at why weight happens to be the relevant unit for the Postal Service and distance for the ICC and CAB. One way by which a regulator can suppress cost differences is to ignore them entirely. However, in deciding which kinds of differences to ignore, he must also take account of the implications for profits. Hence my conjecture would have to be that weight-related costs are more important than distance-related costs in determining first-class postal service profits and vice versa for transportation. A further implication would then be that price/marginal cost in first-class postal service would be negatively related to marginal cost/pound, holding distance constant.

Congressional and Executive initiatives to reduce the CAB's regulatory powers.²⁹

The intra-group equilibrium aspects of the model reveal some implications for entry--both of regulators and regulated firms. First there is a clear incentive for regulators to limit entry (or seek the power to do so) quite apart from considerations of the producer interest. This stems directly from the fact that the politically appropriate price structure is invariably discriminatory (in the economic sense) when costs differ among customers. The proverbial "cream skimming" entrant must be prevented from serving the low cost customers and thereby preventing the regulator from spreading the low costs to others. On the other hand, we can expect the regulator to be more tolerant of entry which dampens the enthusiasm of producers for demand-based price discrimination. The regulator seeks to suppress the full effects of differences in the elasticity of demand, and his way can be eased by permitting entry into low-elasticity market segments. This last argument has more force in industries, like banking, where the primary regulatory control is over entry rather than price. In these cases, the regulator uses the entry control to produce indirectly the desired price structure. A testable implication would be that more entry is permitted in banking, say, the larger the gap between interest rates on small and large loans.

The obverse of the previous argument is that entry of regulation is more attractive the more disparate the price structure. This is independent of the pre-regulatory market structure. Competitively determined, cost-based price differentials create an opportunity for political gain

²⁹See U.S. Senate, op. cit. and U.S. Department of Transportation, Aviation Act of 1975 (1975).

through entry and/or price regulation designed to suppress the effects of cost differences, just as discriminating monopoly invites political suppression of the effects of demand elasticity differences.

In summary, the same forces that make regulators seek a broad-based coalition operate on the price structure. Opportunities for increasing producer wealth by price discrimination are not ignored, but they are never fully exploited. To do this would narrow the consumer base of the coalition. The uniquely political contribution to a price structure is to force a more uniform treatment of consumers than the unregulated market by weakening the link between prices and cost and demand conditions.

Concluding Remarks

This paper is concerned more with the design than the implementation of a research strategy. Much of the recent work in the theory of regulation has focused on political power relationships: which groups will have the muscle to extract gains from their regulatory process. I have largely begged this issue. In my general model, every identifiable group contains winners and losers, and even where all the winners are in one group they end up short-changed. This sort of result can hardly illuminate the nature of the underlying power relationships, but that shortcoming is purposeful.

In the way I have chosen to model the regulatory process, these power relationships play a role analogous to tastes in consumer choice theory. They shape the regulator's utility function. It has proved a highly rewarding research strategy for consumer choice theorists precisely to beg questions of taste formation and concentrate instead on the behavioral effects of changes in constraints in a regime of stable tastes. With some qualification, there is an analogous history in production theory.

I am suggesting here that the theory of politics has something to learn from this experience. Even if we can do no more than derive the most general properties of political power functions, there is much to learn about political behavior in a world where the constraints do change. And the specific contribution of economics to this venture will be enhanced if the constraints are those already familiar to economists. I have tried to show here how the most familiar sort of supply-demand apparatus can be converted into a constraint on regulatory behavior. Once this is accomplished the equally familiar analytics of supply-demand changes yield refutable implications about a wide range of regulatory behavior: when regulation will occur, how it will modify the unregulated price structure, even how it will change the division of the gains over time (with no change in relative political strengths).

Of course, no student of George Stigler can view the derivation of refutable implications as more than a first step. The usefulness of the model developed here awaits tests of these implications, of which the present paper is nearly devoid. The limited progress we have made in exploring political "tastes" is my main ground for optimism about the fruitfulness of a return to a more familiar theoretical mode.³⁰

³⁰ Some specification of power relationships is unavoidable. It is implicit, as Stigler has pointed out to me, in the choice of groups for which the model's regulator acts as broker. For example, why not posit a political redistribution between electricity producers and peanut vendors? Also, most of the results of the model are driven by "normality" of the political-wealth effect. Normality, in this context, is a specific assumption about power (inter) relationships.