Exit,
Voice,
and
Loyalty

Responses to Decline
in Firms, Organizations,
and States

Albert O. Hirschman

Harvard University Press
Cambridge, Massachusetts
and London, England
A Special Difficulty in Combining Exit and Voice

The groundwork has now been laid for telling the reader about the empirical observation that was mentioned in the Preface as the origin of this essay. In a recent book, I tried to explain why the Nigerian railways had performed so poorly in the face of competition from trucks, even for such long-haul, bulky cargo as peanuts (which are grown in Northern Nigeria, some eight hundred miles from the ports of Lagos and Port d’Harcourt). Specific economic, socio-political, and organizational reasons could be found for the exceptional ability of the trucks to get the better of the railroads in the Nigerian environment; but having done so I still had to account for the prolonged incapacity of the railroad administration to correct some of its more glaring inefficiencies, in spite of active competition, and proposed the following explanation:

The presence of a ready alternative to rail transport makes it less, rather than more, likely that the weaknesses of the railways will be fought rather than indulged. With truck and bus transportation available, a deterioration in rail service is not nearly so serious a matter as if the railways held a monopoly for long-distance transport—it can be lived with for a long time without arousing strong public pressures for the basic and politically difficult or even explosive reforms in administration and management that would have been required. This may be the reason public enterprise, not only in Nigeria but in many other countries, has strangely been at its weakest in sectors such as transportation and education where it is subjected to competition: instead of stimulating improved or top performance, the presence of a ready and satisfactory substitute for the services public enterprise offers merely deprives it of a precious feedback mechanism that operates at its best when the customers are securely locked in. For the management of public enterprise, always fairly confident that it will not be let down by the national treasury, may be less sensitive to the loss of revenue due to the switch of customers to a competing mode than to the protests of an aroused public that has a vital stake in the service, has no alternative, and will therefore “raise hell.”

In Nigeria, then, I had encountered a situation where the combination of exit and voice was particularly noxious for any recovery: exit did not have its usual attention-focusing effect because the loss of revenue was not a matter of the utmost gravity for management, while voice did not work as long as the most aroused and therefore the potentially most vocal customers were the first ones to abandon the railroads for the trucks. It is particularly this last phenomenon that must be looked at more closely, for if it has any generality, then the chances that voice will ever act in conjunction with exit would be poor and voice would be an effective recuperation mechanism only in conditions of full monopoly “when the customers are securely locked in.”

As a preliminary to generalizing about this sort of situation, another example, closer to home, may be helpful. If public and private schools somewhere in the United States are substituted in the story for the railroads and lorries of Nigeria, a rather similar result follows. Suppose at some point, for whatever reason, the public schools deteriorate. Thereupon, increasing numbers of quality-education-conscious parents will send their children to private schools. This “exit” may occasion some impulse toward an improvement of the public schools; but here again this im-

2. Private schools being costly and income distribution unequal, the public schools will of course be deserted primarily by the wealthier parents. Nevertheless, the willingness to make a financial sacrifice for the sake of improving the children’s education differs widely within a given income class, especially at intermediate levels of income. In its pure form, the phenomenon here described is best visualized for a school district with many middle-class parents for whom
Exit, Voice, and Loyalty

pulse is far less significant than the loss to the public schools of those member-customers who would be most motivated and determined to put up a fight against the deterioration if they did not have the alternative of the private schools.

In the preceding examples, insensitivity to exit is exhibited by public agencies that can draw on a variety of financial resources outside and independent of sales revenue. But situations in which exit is the predominant reaction to decline while voice might be more efficacious in arresting it can also be observed in the sphere of private business enterprise. The relation between corporate management and the stockholders is a case in point. When the management of a corporation deteriorates, the first reaction of the best-informed stockholders is to look around for the stock of better-managed companies. In thus orienting themselves toward exit, rather than toward voice, investors are said to follow the Wall Street rule that “if you do not like the management you should sell your stock.” According to a well-known manual this rule “results in perpetuating bad management and bad policies.” Naturally it is not so much the Wall Street rule that is at fault as the ready availability of alternative investment opportunities in the stock market which makes any resort to voice rather than to exit unthinkable for any but the most committed stockholder.  

While it is most clearly revealed in the private-public school case, one characteristic is crucial in all of the foregoing situations: those customers who care most about the quality of the product and who, therefore, are those who would be the most active, reliable, and creative agents of voice are for that very reason also those who are apparently likely to exit first in case of deterioration.

One interest of this observation is that it could define a whole class of economic structures where a tight monopoly would be preferable, within the framework of the “slack” or “fallible” economy, to competition. But before jumping to this conclusion, we must take a closer look at the observation by translating it into the ordinary language of economic analysis.

In terms of that language, the situations just described have more than a faint odor of paradox. We all know that when the price of a commodity goes up, it is the marginal customer, the one with the smallest consumer surplus, the one, that is, who cares least, who drops out first. How is it then that with a decline in quality the opposite seems quite plausible: Is it possible that the consumers who drop out first as price increases are not the same as those who exit first when quality declines? 4 If this question were to be answered in the affirmative, it would be easier to understand why combining exit and voice is so troublesome in some situations.

The basic reason for our paradox lies in the still insufficiently explored role of quality (as contrasted with price) in economic life. Traditional demand analysis is overwhelmingly in terms of price and quantity, categories which have the immense advantage of being recorded, measurable, and finely divisible. Quality changes have usually been dealt with by economists and statisticians

3. The passages in quotes are from B. Graham and D. L. Dodd, *Security Analysis*, 3d ed. (New York: McGraw-Hill, 1951), p. 616. The argument is spelled out in some detail in ch. 50, “Stockholder-Management Controversies.” In the fourth edition of this work (1962), the authors return only briefly to this argument, and seem to be aware that the institutional odds are heavily stacked against any substantial success of their exhortations: “In quixotic fashion perhaps,” they say wistfully, “we wanted to combat the traditional but harmful notion that if a stockholder doesn’t like the way his company is run he should sell his shares, no matter how low their price may be” (p. 674).

4. Appendix C refers to this possibility as the “reversal phenomenon.” The discussion in the following pages should be read in conjunction with Appendixes C and D by those who find diagrams clearer than language.
Exit, Voice, and Loyalty

through the concept of the *equivalent* price or quantity change. An article of poor quality can often be considered to be simply less in quantity than the same article of standard quality; this is the case, for example, of the automobile tire which lasts on the average only half as long (in terms of mileage) as a high quality tire. Alternatively, poor quality can often be translated into higher costs and prices; for example, increased pilferage in the rendering of railroad freight service will result in higher insurance premiums. In the latter case, a large part of the quality deterioration can be described by the statement: “now everybody really pays more for the same railroad service than before.” To the extent that this statement is correct, there would be no reason to expect the effect of quality deterioration on demand (that is, for who gets out first) to be any different from the effect of a uniform rise in price. In other words, if a quality decline can be fully expressed as an equivalent rise in price that is *uniform for all buyers* of the article, the effects on customer exit of the quality decline and of the equivalent rise in price would be identical.

The crucial point can now be made. For any one individual, a quality change can be translated into equivalent price change. But this equivalence is *frequently different for different customers of the article because appreciation of quality differs widely among them*. This is so to some extent even in the just mentioned case of automobile tires and of increased pilferage of freight sent by rail. Appreciation of the longer life of quality tires will depend on the time discount of each individual buyer. In the case of rail freight, the increase in the insurance premium fully offsets only the increase in average direct monetary costs which is occasioned to the shipper by the worsening in service. For some shippers this may be all they care about, but there will surely be others for whom the lessened reliability of rail service represents costs (in inconvenience,

A Special Difficulty in Combining Exit and Voice

reputation of their own reliability, and so forth) that cannot be fully made good through an insurance scheme. That appreciation of quality—of wine, cheese, or of education for one’s children—differs widely among different groups of people is surely no great discovery. It implies, however, that a given deterioration in quality will inflict very different losses (that is, different equivalent price increases) on different customers; someone who had a very high consumer surplus before deterioration precisely because he is a connoisseur and would be willing to pay, say, twice the actual price of the article at its original quality, may drop out as a customer as soon as quality deteriorates, provided a nondeteriorated competing product is available, be it at much higher price.

Here, then, is the rationale for our observation: in the case of “connoisseur goods”—and, as the example of education indicates, this category is by no means limited to quality wines—the consumers who drop out when quality declines are not necessarily the marginal consumers who would drop out if price increased, but may be intramarginal consumers with considerable consumer surplus; or, put more simply, the consumer who is rather insensitive to price increases is often likely to be highly sensitive to quality declines.

At the same time, consumers with a high consumer surplus are, for that very reason, those who have most to lose through a deterioration of the product’s quality. Therefore, they are the ones who are most likely to make a fuss in case of deterioration until such time as they do exit. “You can actively flee, then, and you can actively stay put.” This phrase of Erik Erikson 5 applies with full force to the choice that is typically made by the quality-conscious consumer or the member who cares deeply about the policies pursued by the organization to which he be-

longs. To make that kind of consumer and member "actively stay put" for a while should be a matter of considerable concern for many firms and organizations, and particularly for those, of course, that respond more readily to voice than to exit.

Before the varieties of consumer behavior in the case of connoisseur goods are further explored, a brief homage to the hoary concept of consumer surplus is in order, for it appears to have the useful property of measuring the potential for the exercise of influence on the part of different consumers. This potential is the counterpart of the concept's traditional content. Consumer surplus measures the gain to the consumer of being able to buy a product at its market price: the larger the gain the more likely is it that the consumer will be motivated to "do something" to have that gain safeguarded or restored. In this way it is possible to derive the chances for political action from a concept that has dwelt so far exclusively in the realm of economic theory.  

Evidently the nature of the available substitute has something to do with the question whether or not connoisseur goods will be rapidly forsaken, in case of deterioration, by the more quality-conscious customers. In the discussion of the exit and voice options in Chapter 3, it was assumed that the only available competing or substitute good was initially of inferior quality, but carried the same price tag. Usually, of course, many other combinations of price and quality exist: in particular, consumers may often have had some hesitation between the good they actually bought, a better-quality substitute with a higher price, and a poorer-quality substitute with a lower price. Suppose now that only the former type of substitute exists

6. For a similar transformation of a time-honored economic concept, the gain from trade, into a political category, namely the influence a trading partner may acquire in the gain-receiving country, see my National Power and the Structure of Foreign Trade (Berkeley: University of California Press, 1945, rev. ed. 1969), ch. 2.

and that the quality of the connoisseur good normally bought by a group of consumers deteriorates. In this case it is immediately plausible that the consumers who valued the deteriorating good most will be the first ones to decide that it is worth their while to go over to the higher-quality, higher-price substitute. If only a lower-price, lower-quality good is available, on the other hand, these highly quality-conscious consumers, even though they suffer greatly as a result of quality deterioration, will stick with it longer than their less quality-conscious colleagues. These and similar propositions can be easily proved by indifference curve analysis.  

Hence the rapid exit of the highly quality-conscious customers—a situation which paralyzes voice by depriving it of its principal agents—is tied to the availability of better-quality substitutes at higher prices. Such a situation has, for example, been observed in the field of housing. When general conditions in a neighborhood deteriorate, those who value most highly neighborhood qualities such as safety, cleanliness, good schools, and so forth will be the first to move out; they will search for housing in somewhat more expensive neighborhoods or in the suburbs and will be lost to the citizens' groups and community action programs that would attempt to stem and reverse the tide of deterioration. Reverting to the public-private school case, it now appears that the "lower-priced" public schools have several strikes against them in their competition with private schools: first, if and when there is a deterioration in the quality of public school education these schools will lose the children of those highly quality-conscious parents who might otherwise have fought deterioration; second, if, thereafter, quality comes to decline in the private schools, then this type of parents will keep their children there for much longer than was the case

7. See Appendix D, which also discusses in more technical terms a number of other points made in this section.
Exit, Voice, and Loyalty

when the public schools deteriorated. Hence, when public and private schools coexist, with the quality of education in the latter being higher, deterioration will be more strenuously fought “from within” in the case of the private than in that of the public schools. And because exit is not a particularly powerful recuperation mechanism in the case of public schools—it is far more so in that of private schools which have to make ends meet—the failure of one of our two mechanisms is here compounded by the inefficiency of the other.

The relevance of the foregoing observation is greatest in certain important discontinuous choices and decisions, such as between two kinds of educational institutions or two modes of transportation. If one assumes a complete and continuous array of varieties, from cheap and poor-quality to expensive and high-quality, then deterioration of any but the top and bottom variety will rapidly lead to a combination of exits: the quality-conscious consumers move to the higher-price, higher-quality products and the price-conscious ones go over to the lower-price, lower-quality varieties; the former will still tend to get out first when it is quality that declines rather than price that rises, but the latter will not be far behind.

The proposition that voice is likely to play a more important role in opposing deterioration of high-quality products than of lower-quality products can nevertheless be maintained for the case of a good with many varieties, if these varieties can be assumed not to be spread with equal density over the whole quality range. If only because of economies of scale, it is plausible that density is lower in the upper ranges of quality than in the lower and middle ranges. If this is so then deterioration of a product in the upper quality ranges has to be fairly substantial before the quality-conscious will exit and switch to the next better variety. Hence the scope for, and resort to, the voice option will be greatest in these ranges; it will be comparatively slight in the medium- and low-quality ranges.

This finding permits two inferences. First, it can be related to the discussion of education which suggested that the role of voice in fending off deterioration is particularly important for a number of essential services largely defining what has come to be called the “quality of life.” Hence, a disconcerting, though far from unrealistic, conclusion emerges: since, in the case of these services, resistance to deterioration requires voice and since voice will be forthcoming more readily at the upper than at the lower quality ranges, the cleavage between the quality of life at the top and at the middle or lower levels will tend to become more marked. This would be particularly the case in societies with upward social mobility. In societies which inhibit passage from one social stratum to another, resort to the voice option is automatically strengthened: everyone has a strong motivation to defend the quality of life at his own station. That cleavages between the upper and lower classes tend to widen and to become more rigid in upwardly mobile societies has become increasingly obvious;
but it has not been an easy observation to make in a culture in which it had long been taken for granted that equality of opportunity combined with upward social mobility would assure both efficiency and social justice.⁹

A rather different inference results if the assumption of a progressive thinning out of varieties at the upper end of the quality scale is brought into contact with the plausible notion that a combination of exit and voice is needed for best results. If this notion is accepted, then the recuperation mechanism may rely too much on exit at the lower end of the quality scale, *but suffer from a deficiency of exit at the upper end*. An illustration of the latter proposition will be found toward the end of the book.

A Theory of Loyalty

As was pointed out in earlier chapters, the presence of the exit option can sharply reduce the probability that the voice option will be taken up widely and effectively. Exit was shown to drive out voice, in other words, and it began to look as though voice is likely to play an important role in organizations only on condition that exit is virtually ruled out. In a large number of organizations one of the two mechanisms is in fact wholly dominant: on the one hand, there is competitive business enterprise where performance maintenance relies heavily on exit and very little on voice; on the other, exit is ordinarily unthinkable, though not always wholly impossible, from such primordial human groupings as family, tribe, church, and state. The principal way for the individual member to register his dissatisfaction with the way things are going in these organizations is normally to make his voice heard in some fashion.¹

As an aside, it is worth noting that, with exit either impossible or unthinkable, provision is generally made in these organizations for expelling or excommunicating the individual member in certain circumstances. Expulsion can be interpreted as an instrument—one of many—which "management" uses in these organizations to restrict resort to voice by members; a higher authority can then in turn restrict the powers of management by prohibiting expulsion, as is for example done to protect consumers when a public service is supplied in conditions of monopoly. But when exit is a wide-open option and voice is largely nonexistent, as in the relations between a firm and its customers in competitive markets, expulsion of a member or customer is a pointless affair and does not need to be specifically prohibited. One way of catching that somewhat rare bird, an organization where exit and voice both hold important roles, may be to look for groupings from which members can both exit and be expelled. Political parties and voluntary associations in general are excellent examples.

The Activation of Voice as a Function of Loyalty

A more solid understanding of the conditions favoring coexistence of exit and voice is gained by introducing the concept of loyalty. Clearly the presence of loyalty makes exit less likely, but does it, by the same token, give more scope to voice?

That the answer is in the positive can be made plausible by referring to the earlier discussion of voice. In Chapter 3 two principal determinants of the readiness to resort to voice when exit is possible were shown to be:

(1) the extent to which customer-members are willing to trade off the certainty of exit against the uncertainties of an improvement in the deteriorated product; and

(2) the estimate customer-members have of their ability to influence the organization.

Now the first factor is clearly related to that special attachment to an organization known as loyalty. Thus, even with a given estimate of one's influence, the likelihood of voice increases with the degree of loyalty. In addition, the two factors are far from independent. A member with a considerable attachment to a product or organization will often search for ways to make himself influential,

¹ There is no intention here to associate absence of exit with "primitiveness." Edmund Leach has noted that many so-called primitive tribes are far from being closed societies. In his classic study Political Systems of Highland Burma (1954) he traced in detail the way in which members of one social system (gumsha) will periodically move to another (gumlaa) and back again. Exit may be more effectively ruled out in a so-called advanced open society than among the tribes studied by Leach.
especially when the organization moves in what he believes is the wrong direction; conversely, a member who wields (or thinks he wields) considerable power in an organization and is therefore convinced that he can get it “back on the track” is likely to develop a strong affection for the organization in which he is powerful.2

As a rule, then, loyalty holds exit at bay and activates voice. It is true that, in the face of discontent with the way things are going in an organization, an individual member can remain loyal without being influential himself, but hardly without the expectation that someone will act or something will happen to improve matters. That paradigm of loyalty, “our country, right or wrong,” surely makes no sense whatever if it were expected that “our” country were to continue forever to do nothing but wrong. Implicit in that phrase is the expectation that “our” country can be moved again in the right direction after doing some wrong—after all, it was preceded in Decatur’s toast by “Our country! In her intercourse with foreign nations, may she always be in the right!” The possibility of influence is in fact cleverly intimated in the saying by the use of the possessive “our.” This intimation of some influence and the expectation that, over a period of time, the right turns will more than balance the wrong ones, profoundly distinguishes loyalty from faith. A glance at Kierkegaard’s celebrated interpretation of Abraham’s setting out to sacrifice Isaac makes one realize that, in comparison to that act of pure faith, the most loyalist behavior retains an enormous dose of reasoned calculation.

When is loyalty functional?

The importance of loyalty from our point of view is that it can neutralize within certain limits the tendency of the most quality-conscious customers or members to be the first to exit. As has been shown in Chapter 4, this tendency deprives the faltering firm or organization of those who could best help it fight its shortcomings and its difficulties. As a result of loyalty, these potentially most influential customers and members will stay on longer than they would ordinarily, in the hope or, rather, reasoned expectation that improvement or reform can be achieved “from within.” Thus loyalty, far from being irrational, can serve the socially useful purpose of preventing deterioration from becoming cumulative, as it so often does when there is no barrier to exit.

As just explained, the barrier to exit constituted by loyalty is of finite height—it can be compared to such barriers as protective tariffs. As infant industry tariffs have been justified by the need to give local industry a chance to become efficient, so a measure of loyalty to a firm or organization has the function of giving that firm or organization a chance to recuperate from a lapse in efficiency. Specific institutional barriers to exit can often be justified on the ground that they serve to stimulate voice in deteriorating, yet recuperable organizations which would be prematurely destroyed through free exit. This seems the most valid, though often not directly intended, reason for the complication of divorce procedures and for the expenditure of time, money, and nerves that they necessitate. Similarly the American labor law sets up a fairly complex and time-consuming procedure for one trade union to take

2. In terms of figure 3 of Appendix B, a person whose influence (that is, the likelihood that he will be able to achieve full quality recuperation) is correctly expressed by a point as high as V3 will be willing to trade off the certainty of the competing product against even a little hope of recuperation for the traditional product. Thus he will choose voice. He who has little influence and knows it, on the other hand, is not likely to take kindly to such a trade-off. If he is to opt for voice rather than exit, he will normally require the certain availability of the competing product to be matched by the near-certainty of recuperation for the traditional variety.
## The Elusive Optimal Mix of Exit and Voice

In the earlier chapters of this volume, much has been made of situations in which exit drives out voice and assumes a disproportionate share of the burden involved in guiding a firm or organization back to efficiency after the initial lapse. It was shown that, in certain situations, voice could function as a valuable mechanism of recuperation and deserves to be strengthened by appropriate institutions. With a commendable sense of balance I have turned my attention, in the last few pages, to situations in which it is the exit option's turn to be almost wholly shunned, much to the detriment, once again, of effective recovery. Having come full circle I shall stop very soon. But in view of the numerous baroque ornamentalations which have been added to the circle as it was being traced a somewhat schematic finale may be useful.

First it is worth recapitulating how organizations array themselves in a table whose classification criterion is the absence or presence of our two reaction mechanisms. It hardly needs saying that this table is by way of a very rough summary. Qualifications to the pigeonholing here presented and borderline cases have been noted throughout the book.

This table brings out the basic contrast between organizations receiving information on member-customer discontent essentially through exit and hardly at all through voice such as business firms in a competitive market, and the more traditional human groupings exit from which is virtually unheard of while voice is available to the members in varying degrees. Organizations where both exit and voice play important roles are relatively few: the most important ones are voluntary associations of various types including, as a most important subcategory, competitive political parties. Also, certain types of clients of business firms will often attempt to influence the firms’ policies directly instead of choosing exit.

There are probably no organizations that are wholly immune to either exit or voice on the part of their members. The ones that have been listed in the cell corresponding to that category are those that, in their intended structure, make no explicit or implicit allowance for either mechanism. Exit is here considered as treason and voice as mutiny. Such organizations are likely to be less viable, in the long run, than the others; exit and voice being illegal and severely penalized, they will be engaged in only when deterioration has reached so advanced a stage that recovery is no longer either possible or desirable. Moreover, at this stage, voice and exit will be undertaken with such strength that their effect will be destructive rather than reformist.

On the other hand, there is no implication in the table that the organizations which are equipped with both feedback mechanisms are necessarily more advanced or viable than those which rely primarily on one alone. All

<table>
<thead>
<tr>
<th>Organizations whose members react strongly via</th>
<th>Exit</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary associations, competitive political parties, and some business enterprises, for example, those selling output to a few buyers</td>
<td>Yes</td>
<td>Family, tribe, nation, church, parties in non-totallitarian one-party systems</td>
</tr>
<tr>
<td>Competitive business enterprise in relation to customers</td>
<td>No</td>
<td>Parties in totalitarian one-party systems, terrorist groups, and criminal gangs</td>
</tr>
</tbody>
</table>
depends on the responsiveness of the organization to whatever mechanism or combination of mechanisms it is equipped with. If an organization equipped with exit is extremely sensitive to the loss of customers or members all is for the best; and similarly for an organization equipped with voice that takes complaints and protests from its members very seriously. But what if an organization is not particularly sensitive to the particular reaction it happens to provoke or does not possess the mechanism to which it would be sensitive? A large part of this book has been devoted to such cases of inadequate or wrong responses and the argument can be summarized by means of another table.

<table>
<thead>
<tr>
<th>Exit</th>
<th>Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitive business enterprise (for qualifications see Chapter 2)</td>
<td>Organizations where dissent is allowed, but is &quot;institutionalized&quot;</td>
</tr>
<tr>
<td>Public enterprise subject to competition from an alternative mode, lazy oligopolist, corporation-shareholder relations, inner cities, etc.</td>
<td>Democratically responsive organizations commanding considerable loyalty from members</td>
</tr>
</tbody>
</table>

The greatest interest centers here naturally on those “perverse” or pathological cases where an organization is in effect equipped with a reaction mechanism to which it is not responsive: those who are affected by quality decline do not vent their feelings in one way or another, but management happens to be inured or indifferent to their particular reaction and thus does not feel compelled to correct its course. The situations of this type which have been described at greatest length, particularly in Chapters 4 and 5, involved organizations whose decline gives rise to exit that does not bother management nearly as much as might be the case for voice. In Chapter 8, however, a situation of the symmetrically opposite type was encountered: an organization—the Executive Branch of the United States government—whose deteriorating performance under Lyndon Johnson led to numerous, but futile manifestations of voice when exit might have been more effective.

Several conclusions follow from the general observation that an organization may be arousing, through its decline, one kind of reaction from its members when its recovery would be more powerfully stimulated by another kind. One advantage of this formulation is that it points immediately to a variety of remedies or a combination of them. Take the case where an organization arouses primarily exit to which it is far more insensitive than it would be to voice. Corrective policies obviously include efforts to make the organization more responsive to exit, but also efforts to have the members of the organization switch from exit to voice. In this fashion, the range of possible remedial measures is broadened. For example, when railroads do not react vigorously to the loss of customers the typical proposal is to introduce stronger “financial disciplines” in the hope that the railroad managers will then react to the loss of revenue like private enterprise threatened by bankruptcy. It is now clear that as an alternative or complementary step it is worth looking into ways and means of strengthening voice on the part of the customers. This can be done directly, by reducing the cost and increasing the rewards of voice, as well as indirectly, by raising the cost of exit and even by reducing the opportunities for it.

Similarly, when an organization arouses but ignores voice while it would be responsive to exit, thought must be given both to making exit more easy and attractive by appropriately redesigned institutions and to making the organization more responsive to voice. The approach to the improvement of institutional designs that is advocated
here widens the spectrum of policy choices that are usually considered and it avoids the strong opposite biases in favor of either exit or voice which come almost naturally to the economist and political scientist, respectively.¹

But a word of caution is now needed about what our approach cannot yield: it does not come out with a firm prescription for some optimal mix of exit or voice, nor does it wish to accredit the notion that each institution requires its own mix that could be gradually approached by trial and error. At any one point of time, it is possible to say that there is a deficiency of one or the other of our two mechanisms; but it is very unlikely that one could specify a most efficient mix of the two that would be stable over time. The reason is simple: each recovery mechanism is itself subject to the forces of decay which have been invoked here all along. This is so not just to add a final touch of philosophical consistency, but for more mundane reasons as well. As has already been mentioned, the short-run interest of management in organizations is to increase its own freedom of movement; management will therefore strain to strip the members-customers of the weapons which they can wield, be they exit or voice, and to convert, as it were, what should be a feedback into a safety valve. Thus voice can become mere “blowing off steam” as it is being emasculated by the institutionalization and domestication of dissent which was described toward the end of the previous chapter. And exit can be similarly blunted. As was shown, organizations and firms that are ostensibly competing and are normally sensitive to exit can learn to play a cooperative, collusive game in the course of which they take in each other’s disgruntled customers or members. To the extent that the game is played successfully by competing organizations or firms, exit, compensated as it is by entry, ceases to be a serious threat to the deteriorating organizations.

¹ See the discussion of the Friedman proposal on education in ch. 1, above.

While management thus finds ways of whittling down the effectiveness of the reaction mode that, in any given set of circumstances, appears to be preferred by consumer-members, the latter are in a way easing management’s task by relying evermore on this mode and letting the other atrophy. As was mentioned at various times (particularly with regard to voice), the effectiveness of the less familiar mode becomes not only more uncertain, but tends to be increasingly underestimated. The reason is that effective use of the less familiar mechanism requires that its power be discovered or rediscovered whereas the preferred mechanism is routinely familiar; since quite properly we dare not believe in creative discoveries until they have happened, we will underestimate the effectiveness of voice when exit is dominant and vice versa. Once members have a slight preference for, say, voice over exit, a cumulative movement sets in which makes exit look ever less attractive and more inconceivable. As a result, voice will be increasingly relied on by members at a time when management is working hard to make itself less vulnerable to it.

For these reasons, conditions are seldom favorable for the emergence of any stable and optimally effective mix of exit and voice. Tendencies toward exclusive reliance on one mode and toward a decline in its effectiveness are likely to develop and only when the dominant mode plainly reveals its inadequacy will the other mode eventually be injected once again.

The invigorating results that can be achieved by the shock effect of such an injection have recently been demonstrated when the consumers’ voice was suddenly introduced, primarily through the courage and enterprise of Ralph Nader, into an area where exit had long been the dominant and almost exclusive mode. In the opposite case, when voice is the dominant reaction mode, exit can be similarly galvanizing. It may be asked why this should be so, why exit, an act of withdrawal, should suddenly prove...
to be influential when the returns from "working from within" via voice have been declining. Exit is not usually undertaken for the purpose of gaining more influence than one had as a member. That is nevertheless the way it often works out, especially when exit is a highly unusual event. Social psychologists have noticed that "the disappearance of the source of communication leads to a change of opinion in its favor." 2 Exit is unsettling to those who stay behind as there can be no "talking back" to those who have exited. By exiting one renders his arguments unanswerable. The remarkable influence wielded by martyrs throughout history can be understood in those terms, for the martyr's death is exit at its most irreversible and argument at its most irrefutable.

The critique of the optimal mix concept thus leads to a triple suggestion. In order to retain their ability to fight deterioration those organizations that rely primarily on one of the two reaction mechanisms need an occasional injection of the other. Other organizations may have to go through regular cycles in which exit and voice alternate as principal actors. Finally, an awareness of the inborn tendencies toward instability of any optimal mix may be helpful in improving the design of institutions that need both exit and voice to be maintained in good health.

It is even conceivable that this book might have a more direct influence. By bringing out the hidden potential of whatever reaction mode is currently neglected, it could encourage resort to either exit or voice, as the case may be. Such at least is the stuff writers' dreams are made of.

Appendix A.
A Simple Diagrammatic Representation of Voice and Exit *

A slight modification of the traditional demand diagram makes it possible to show the manner in which exit and voice are generated as a result of quality deterioration. In the following, the quality-elasticity of demand, that is, the exit-response of consumers to downward changes in quality, is assumed to be given, without regard to the possibility and prospective effectiveness of voice.

Figure 2a shows demand as a function of quality, measured along the vertical axis, rather than of price. Distance from the origin increases the poorer the quality; in this fashion the demand curve retains its usual downward slope. Normal quality is shown by point \( L_0 \), while post-lapse quality is shown by \( L_1 \). Figure 2b gives the conventional representation of unit price along the vertical axis. Both figures show quantity bought on the horizontal axis. When quality drops from \( L_0 \) to \( L_1 \) with unit price remaining unchanged, demand declines from \( Q_0 \) to \( Q_1 \) and total revenue losses are shown on figure 2b by the area of the rectangle \( Q_1 Q_0 P_0 T \). This is the exit or "E" rectangle. To what extent this revenue loss will curtail or perhaps wipe out the firm's profits depends of course on cost conditions which are not shown here.

Voice, on the other hand, depends on the number of non-exiting customers \( OQ_1 \) and on the degree of deterioration \( L_0 L_1 \) in figure 2a. Hence the amount of voice likely to be forthcoming is proportional to the area of the rectangle \( L_0 TP_1 L_1 \). This is the voice or "V" rectangle.

In general the \( E \) and \( V \) rectangles are not directly additive. But for any drop in quality, exit and voice somehow combine to exert influence on management; if the share of each in the total influence could be estimated, one could draw, for this drop, the vertical scale of the price diagram (figure 2b) so that the \( E \) and \( V \) rectangles would represent correctly their respective influence shares. Figure 2, for example, has been drawn so as to show that when quality declines from \( L_0 \) to \( L_1 \), exit will be twice as

* See text pages 22–23 and 33–36.
Figure 2. Voice and exit when demand is a function of quality effective as voice. If, within a certain range, the effectiveness of both voice and exit depend exclusively on, and vary directly with, their rectangle areas, then the decisive factor in determining the shares of voice and exit in the aggregate amount of pressure that is brought to bear on the firm is the quality-elasticity of demand. Under these conditions, higher elasticity means higher aggregate effectiveness of voice and exit only if it is assumed that the negative effect on recuperation due to the decline in the voice rectangle is more than compensated by the increase in the exit rectangle.

Voice can inflict direct costs on management as complaining customers occupy the time of the firm’s personnel and succeed in having defective merchandise “fixed up” or exchanged. To the extent that this is the case, voice has directly adverse monetary consequences which could then be represented in figure 2b. Suppose, for example, that half of the nonexisting customers complain and that the average complaint results in a cost that amounts to one-half of the article’s sales price; then voice would inflict a monetary loss equal to one-fourth of the $ORTQ_1$ rectangle. (Note that voice acts directly on profits while exit acts on profits via revenue.) It should be emphasized, however, that the effectiveness of voice does not at all depend on being convertible in this fashion into the measuring rod of money. Some comments closely related to this point appear on pages 72–74.