A paper prepared for the

Economic Council of Canada

P.O. Box 527
Ottawa, Ontario
K1P 5V6

Un document préparé pour le

Conseil économique du Canada

C.P. 527
Ottawa (Ontario)
K1P 5V6

[redacted]

DISCUSSION PAPER NO. 282

Issues in Subsidization with Attention to the Subsidization of Passenger Rail Service

by Jerome T. Bentley *

* Senior economist,
Mathtech - The Technical Research and Consulting Division of Mathematica, Inc.
Princeton, NJ

The findings of this Discussion Paper are the personal responsibility of the authors and, as such, have not been endorsed by members of the Economic Council of Canada.

Discussion Papers are working documents made available by the Council, in limited number and in the language of preparation, to interested individuals for the benefit of their professional comments.

ISSN-0225-8013

May 1985

CONTENTS

LE FINANCEMENT PAR SUBVENTIONS, NOTAMMENT
DU TRANSPORT FERROVIAIRE DE VOYAGEURS

Résumé

La présente étude porte sur le financement, au moyen de subventions, des services assurés par des entreprises publiques. L'auteur accorde une attention spéciale aux subventions qui favorisent l'efficacité des services et contribuent à en accroître la qualité; il étudie les divers aspects économiques du facteur qualité. Il fait aussi une évaluation des divers moyens possibles pour en améliorer le niveau. Il examine ensuite les pratiques passées et actuelles dans le domaine du transport des voyageurs. Enfin, il analyse le cas particulier des services voyageurs qu'offrent les chemins de fer canadiens.

Les conditions nécessaires à une répartition efficace des ressources sont plus rigoureuses lorsqu'on tient compte à la fois de la qualité et du niveau de la production. Des marchés qui, autrement, seraient peut-être concurrentiels évoquent plutôt alors une situation de concurrence monopolistique. En outre, il est nécessaire de se renseigner sur les préférences des consommateurs sous-marginaux pour s'assurer qu'ils obtiennent des services d'une qualité optimale.

Parmi les divers moyens envisagés pour favoriser la qualité des services, le plus apprécié est la soumission en libre concurrence. Toutefois, l'existence d'importants coûts irrécupérables peut éliminer cette méthode comme élément d'une politique. Des incitations financières directes pour accroître la performance peuvent aussi contribuer efficacement à améliorer la qualité des services, si l'on dispose de mesures de la performance à la fois objectives et pratiques.

L'auteur analyse les pratiques passées et actuelles concernant trois aspects de la qualité des services voyageurs, soit le respect et la qualité des horaires, ainsi que le confort des voyageurs. Des incitations financières directes sont souvent utilisées pour encourager la régularité des services ferroviaires. Aux Etats-Unis, la société Amtrak a tenté, au moyen de subventions, d'améliorer le confort des voyageurs et la précision des horaires. Malheureusement, l'expérience a échoué, peut-être à cause de la difficulté d'établir et d'appliquer des normes pour le confort des voyageurs, et à cause aussi du conflit normal entre le respect des horaires et la réduction de la vitesse des trains.

Au Canada, la société VIA Rail éprouve de graves contraintes institutionnelles qui l'empêchent de bien contrôler la qualité de son service voyageurs. Etant donné la facturation ultérieure au prix coûtant majoré, que pratiquent les chemins de fer, VIA Rail a peu de contrôle sur une grande partie de ses coûts et elle n'a pas accès aux renseignements qui pourraient lui permettre de prendre des décisions éclairées en matière de politiques. De plus, à cause d'incitations insuffisantes, elle n'est pas portée à contrôler ses propres coûts directs. Bien sûr, des subventions fondées sur la propre performance de VIA Rail pourraient améliorer cette situation, mais il faudrait lui fixer au préalable un ensemble d'objectifs précis pouvant servir de critères pour évaluer ses réalisations. Enfin, beaucoup de problèmes de qualité des services tiennent à l'équipement. Des investissements s'imposeraient peut-être pour que l'on puisse améliorer de façon notable le transport ferroviaire de voyageurs au Canada.

ISSUES IN SUBSIDIZATION WITH ATTENTION
TO THE SUBSIDIZATION OF PASSENGER RAIL SERVICE

Abstract

The subsidization of services provided by government enterprises is examined in this study. Special attention is focused on subsidies that encourage efficiency and promote service quality. The economics of quality is studied and several alternative approaches to promoting service quality are evaluated. Next, past and current practices in the passenger transport sector are reviewed. Finally, the specific case of rail passenger service in Canada is studied.

The conditions necessary to achieve allocative efficiency are more stringent when both quality and output levels are considered. Markets that might otherwise be competitive tend to look like monopolistic competition. Information on the preferences of infra-marginal consumers is needed to assure that the optimal level of service quality is produced.

Of the approaches considered for promoting service quality, competitive tendering is preferred. However, the presence of large sunk costs may render this approach infeasible as a policy option. Direct performance incentive payments can also be an effective means of promoting service quality if objective measures of performance are available and enforceable.

Past and current practices directed to three aspects of passenger service quality -- schedule adherence, passenger comfort, and schedule quality -- are reviewed. Direct incentive payments are often used to promote on-time performance. In the United States, Amtrak has experimented with direct payments for passenger comfort and schedule quality. These experiments failed, presumably because of the difficulty of establishing and enforcing standards for passenger comfort, and the natural conflict between on-time performance and reduced running times.

In Canada, VIA Rail faces serious institutional constraints on its ability to control the quality of rail passenger service. Because of cost-plus *ex post* billing by railways, VIA Rail has little control over a major portion of its costs, and moreover, VIA Rail does not have access to information required to make intelligent policy decisions. In addition, VIA Rail does not have appropriate incentives to control its own direct costs. Subsidies to VIA Rail based on its own performance may improve this situation, but VIA Rail must first be given a clear set of

objectives against which its performance can be assessed. Finally, many service quality problems are equipment related. Capital investments may be required before significant improvements in Canadian rail passenger service can be achieved.

1. INTRODUCTION

Much of the analysis conducted in the economic literature considers price and output as variables endogenous to the firm's decision-making process. A third, and perhaps equally important variable, quality, has received considerably less attention. Indeed, to the extent that quality differences are recognized, economists typically consider output as having been produced and exchanged in different markets.

Similarly, the literature on the economics of regulation have focused primarily on how market failures can be addressed so that an efficient level of output can be achieved. The lack of attention given here to quality, however, is possibly more serious. Anytime a firm is constrained in terms of either price or output, but is left free to vary quality, market imperfections are likely to occur.

This study explores alternatives for assuring that the appropriate level of quality is embodied in public services. Although much of the discussion is general, it is intentionally focused on VIA Rail as a case study. The general model is one in which VIA Rail is considered as a broker for rail passenger services. That is, VIA Rail purchases inputs from other firms, and then sells trips, embodying alternative levels of quality, to rail passengers. The central question here is: How can VIA Rail, as a contractor, assure that the services it purchases embody appropriate levels of quality and that such services are produced at minimum costs?

It should be noted, however, that the model is sufficiently general to allow for other views of VIA Rail, or public enterprises in general. For

example, if one considers VIA Rail as a producer instead of a broker, the optimal incentive payments made by VIA Rail for service quality can be interpreted as optimal subsidies.

1.1 Summary of Findings and Conclusions

The major findings and conclusions of the topics discussed in each of the remaining four sections of this report are summarized below. These topics include the economics of quality, approaches to promoting service quality, practices in the transportation sector, and rail passenger service in Canada.

1.1.1 The Economics of Service Quality —

The conditions necessary to achieve the allocative efficiency associated with competitive markets when service quality is considered are more stringent than those necessary when only the level of output is considered. Specifically, competition along both quality and output spectrums is required.

Limited quality options may cause suboptimal resource allocation. As a policy matter, this means that both the level and variation in available quality are at issue. In addition, quality considerations may cause markets that would otherwise be competitive to yield monopolistically competitive results, even if only a single level of quality is actually produced.

Perhaps the most surprising result is that monopolists, as a general rule, do not necessarily under-produce service quality. In fact, they will either under- or over-produce quality depending on how marginal and infra-marginal buyers value service quality. The important point here is that information about infra-marginal buyers is needed to establish the optimal (in terms of allocative efficiency) level of service quality.

Subsidizing service quality may be appropriate under the following circumstances:

- Service quality generates non-user benefits which are not appropriable by the enterprise.
- Equity or distributional considerations are included among the objectives of the enterprise.
- The enterprise is a natural monopoly.

It should be noted that the natural monopoly case is applicable here only if the marginal cost of service quality declines. It should also be noted that lump sum subsidies, especially those paid on an ex post basis, should be avoided. Lump sum subsidies do not provide incentives for service quality improvements.

1.1.2 Approaches to Promoting Service Quality —

Of the approaches considered for promoting service quality, competitive tendering is preferred. Because of the play of competitive market forces, given levels of service quality tend to be produced at minimum cost. Contractual service quality provisions can be enforced by fiat

through the threat of contract termination -- the approach employed most often in the private sector.

Competitive tendering also affords the purchasing agent significant leverage in contract negotiations. In addition, this approach is the most effective means of dealing with service quality problems when performance standards are difficult to measure quantitatively. However, the presence of large sunk costs may render competitive tendering infeasible as a policy option.

Direct performance incentive payments are also an effective means for promoting service quality. In this case, however, objective measures of performance must be available and enforceable. Direct performance incentive payments may also be effective when employed together with competitive tendering.

1.1.3 Review of Service Quality Practices --

Direct incentive payments are often employed to encourage on-time performance. Both VIA Rail and Amtrak include such incentives in their contracts with railways. Mass transit authorities also sometimes incorporate incentives in their contracts (usually delay penalties) with private bus lines. The prevalence of these arrangements is at least partially attributable to the fact that objective measures of schedule adherence are available and enforceable.

Amtrak once experimented with direct performance incentives for promoting passenger comfort. Specific provisions for bonuses and penalties associated with different aspects of passenger comfort were included in

several contracts negotiated with U.S. railways in 1974. Current Amtrak contracts, however, contain no such provisions. Difficulties in monitoring and enforcing contracts and establishing objective measures of passenger comfort are often cited as reasons for the failure of this experiment. In addition, Amtrak has now assumed direct responsibility for most cleaning and maintenance activities which affect passenger comfort.

Mass transit authorities, in their dealings with private bus lines, often assure passenger comfort through competitive tendering. VIA Rail's contracts with railways have no direct incentives for promoting with passenger comfort.

Amtrak has also experimented with direct performance incentives for improving schedule quality. Specifically, railways could have received a one-time bonus for reducing running times. This experiment failed because schedule improvement bonuses were small relative to schedule adherence incentives.

1.1.4 Rail Passenger Service Quality in Canada —

Public policy directed to service quality must be established within a framework which considers the overall goals and objectives of Canadian rail passenger service. VIA Rail should be given a set of clear objectives so that it is possible to establish meaningful standards upon which VIA Rail performance can be evaluated. A clear set of objectives must also be established to determine the appropriate degree and nature of VIA Rail subsidization.

At present, VIA Rail faces serious institutional constraints on its ability to control the passenger services for which it is responsible. Specifically:

- Because of cost-plus ex post billings by railways, VIA Rail has little control over major portions of its costs. VIA Rail should be permitted to negotiate fixed-rate agreements with the railways.
- VIA Rail does not have access to information required to make appropriate policy decisions. VIA Rail should have direct access to all railway cost figures relevant to passenger services.

It is also important to note that, because of present subsidy arrangements, VIA Rail does not have appropriate incentives to control its own costs. Subsidies based on performance can be employed to encourage cost efficiency within VIA Rail itself. First, however, VIA Rail must be given a clear set of objectives against which its performance can be assessed.

Many service quality problems are created by equipment failures. VIA Rail does not make final decisions on funding for equipment purchases. In addition, VIA Rail has little control over equipment maintenance services. In any event, substantial investments in new or upgraded equipment may be required to significantly improve the quality of Canadian rail passenger service. Such investments should be undertaken only after a clear set of objectives for rail passenger service are established.

Several specific comments on VIA Rail on-time performance incentive agreements are appropriate. The most important of these comments include the following:

- Railways are not penalized for delays caused by equipment-related problems, a major cause of poor on-time performance.
- Railways have little incentive to improve on-time performance above 90 percent, or to prevent service degradation below 75 percent.
- The effectiveness of the on-time performance incentives are significantly reduced by the fact that railways are reimbursed on a cost-plus ex post basis.

As was noted earlier, however, improvements in VIA Rail's rolling stock may be required to improve on-time performance.

Many problems related to passenger comfort are also equipment-related. Apart from this, no incentives to promote passenger comfort are currently in place. Direct incentives may be tried, but Amtrak's experience should be considered. VIA Rail should also consider, as a long-run goal, to promote opportunities to tender services related to passenger comfort competitively.

Schedule quality is currently negotiated with the railways. Because of the natural conflict between reliability and schedule quality (i.e., trip time), direct incentives here may be ineffective.

1.2 Organization of Report

Immediately below, in Section 2 of this report, a fairly general discussion of the economics of quality is provided. The relationship between the goals of the enterprise, subsidies, and service quality is also discussed in this section. In Section 3, four alternative methods -- direct incentive payments, rate of return regulation, franchise-type agree-

ments, and competitive tendering -- are assessed for their effectiveness in dealing with service quality problems. A number of past and existing practices for dealing with service quality in the transport sector are reviewed and evaluated in Section 4. In particular, practices aimed at three general aspects of passenger service quality -- reliability, passenger comfort, and schedule quality -- are reviewed.

Section 5 is devoted specifically to VIA Rail. First, objectives of the enterprise, evaluation of alternatives, and subsidies are discussed. Next, a number of important institutional constraints -- as they are related to service quality -- are described. Specific comments on existing incentives for Canadian rail passenger service quality are provided, and some recommendations are offered.

The discussions in Section 2 and 3 are conducted at a fairly general and conceptual level. Those readers who are more interested in applications to rail passenger service may want to skip these two sections.

2. THE ECONOMICS OF QUALITY

In order to understand how alternative incentive schemes to improve service quality affect resource allocation, it is first necessary to understand how quality is determined under different market structures. This is the subject of this section of the report. The discussion below is conducted at a fairly general level, but some applications that are potentially relevant to the VIA Rail case are noted.

Appropriate incentives for the provision of service quality cannot be divorced from the overall objectives of the government enterprise. In VIA Rail's case, this naturally raises issues related to the appropriate level and nature of subsidization. These general issues are discussed immediately below in Section 2.1.

Next, in Section 2.2, the conditions necessary to achieve allocative efficiency with respect to product or service quality in a competitive market are described. These conditions are more restrictive than those required for allocative efficiency when only quantity (output) is considered; as a result, market failures associated with quality aspects of goods and services are likely to be far more prevalent than those associated with quantity.

The monopoly case is discussed in Section 2.3. Perhaps the most significant finding there is that allocative efficiency is not assured even if the monopolist is required to provide service quality up to the point at which price equals marginal cost of quality. This finding is significant because it means that the policy-maker needs information on the willing-

ness-to-pay for quality of inframarginal buyers in order to achieve allocative efficiency.

Finally, a number of intermediate cases are discussed in Section 2.4. Many of these intermediate cases evolve to market structures that look like monopolistic competition. The relationships between price and quality when constraints are placed in markets is also described in Section 2.4.

2.1 Government Enterprise Objectives and Subsidization Issues

The primary interest of this study is to describe incentive structures and other mechanisms that will encourage the government enterprise to produce appropriate levels of service quality. The appropriate level of service quality, however, can only be defined within a context of the overall objectives of the enterprise.

This discussion naturally requires that the general issue of subsidies for government enterprises be addressed. Depending on the objectives of the government enterprise and the nature of service quality markets, subsidization may be appropriate. Subsidization as an issue is especially relevant in VIA Rail's case. VIA Rail is already heavily subsidized; other things being the same, improvements in service quality are likely to increase the cost of rail passenger service and may require even higher levels of subsidization.* Of course, at least some of the costs of

---

* This, of course, presumes that the service quality of rail passenger service is currently produced at minimum cost.

improved service quality may be financed through higher fares since, presumably, passengers will be willing to pay more for higher quality.

2.1.1 Objectives of the Government Enterprise —

It is convenient to structure the discussion of government enterprise objectives within the context of formal benefit-cost analysis. In this context, the government enterprise defines a set of objectives, evaluates the benefits and costs (both explicit and implicit) of alternative projects, and then selects the "best" alternative.

For example, VIA Rail might consider a set of projects defined by a continuum of passenger service quality levels. A project is efficient relative to a baseline level of service quality if the incremental benefits associated with the project exceed its incremental costs. The most efficient project is the one associated with the greatest positive difference between total benefits and total costs. At least one project must be economically feasible (i.e., total project benefits must exceed total project costs); otherwise, the enterprise or society as a whole is better-off if the service is not provided at all.

The term "government enterprise" is somewhat paradoxical. As an "enterprise," its objective is to maximize the net present value of the firm. The "government" descriptor suggests that the firm is entrusted with other objectives which are not consistent with behavior as a profit

maximizer.* The discussion that follows attributes these broader social objectives to the government enterprise.

There are two generic classes of objectives that can be defined. These are:**

- Allocative efficiency.
- Equity.

The term "allocative efficiency" is used here in the Pareto sense. Specifically, allocative efficiency is achieved when resources are allocated (e.g., in producing service quality) such that no member of society can be made better-off without making some other member of society worse-off. If the objective of the government enterprise can be stated strictly in terms of allocative efficiency, service quality levels should be based on willingness-to-pay.

The objectives of the enterprise may not be limited to strict allocative efficiency. In particular, "equity" in distribution may also be an objective. For example, some services may be provided at levels which exceed willingness-to-pay because they are judged to result in a distribution of benefits that is thought to be "fair."

---

* The implication is that the activities of the government enterprise would be undertaken in the private sector absent the presence of social goals inconsistent with pure profit-maximizing behavior.

** Public policies are often evaluated in terms of impacts on factors such as employment, international trade, and energy policy. In general, however, these sub-objectives can be classified as special cases of objectives related to allocative efficiency and welfare distribution.

In general, policy changes will cause a redistribution of benefits (and costs) across different members of society. If distributional considerations are included in the set of objectives for the government enterprise, then projects cannot be evaluated strictly by the willingness to pay criterion. In this case, the benefits (and costs) of distributional impacts must be evaluated in assessing service quality associated with alternative projects.

Allocative inefficiencies are created by failures stemming from problems caused by market structures and the existence of non-user benefits. Market structure problems, as they are related to service quality, are discussed in considerable detail in subsequent parts of this section.

Since non-user benefits do not accrue to consumers, they are not reflected in users' willingness to pay for services. The market failure is created by the inability of the producer to appropriate these benefits. Non-user benefits may be classified as follows:

- External economies.
- Option value.
- Existence value.

Each of these is discussed briefly below.

External economies occur when some member of society other than the consumer benefits from the consumption or production of the service. Rail passenger service, for example, may reduce congestion and pollution in some

communities, thus producing benefits for individuals other than rail passengers.*

Benefits from option value occur when some members of society value the opportunity to consume a service at some future date. An individual, for example, may receive value from the opportunity to use rail passenger service in the event of a future emergency. Society may also receive benefits collectively from the option to use rail service in the future because of events such as higher fuel prices.

Benefits from existence value arise when some members of society place value on the existence of a resource apart from its value in use (or potential use). The concept of existence value is often used in environmental economics to describe the value that some members of society place on the existence of rare species. It is reasonable to suspect that some individuals place existence value on rail passenger service.

The preceeding discussion has focused on non-user benefits. In general, it is possible that non-user costs may be associated with the consumption and production of a service. Just as it is appropriate to consider non-user benefits, non-user costs should be included in the evaluation of alternative projects.

2.1.2 Subsidization Issues —

As was noted earlier, subsidies may be appropriate depending on the objectives of the government enterprise and the nature of the market for

---

* Such community benefits are likely to be capitalized in property values.

services. If allocative efficiency is the goal of the enterprise, subsidies may be appropriate under the following circumstances:

- The natural monopoly case.
- The presence of external economies.
- The public good case.

It is well known that the natural monopoly -- characterized by declining long-run marginal costs -- cannot achieve allocative efficiency and, at the same time, obtain revenues sufficient to earn a normal return on investment. Indeed, subsidies to railroads have often been rationalized under this argument.

Several points relative to the VIA Rail case and the subsidization of natural monopolies are worth noting here. These are:

- Even if the long-run marginal costs of passenger service are declining, subsidies are justified only if there exists some project (e.g., some output and service quality level) for which total benefits exceed total costs.

- Lump sum payments made to subsidize the difference between project costs and revenues cannot be made ex post -- as is the current arrangement for VIA Rail, the CN, and the CP -- and still encourage production at minimum costs. In addition, lump sum subsidies cannot be employed to promote service quality.

- The enterprise might have decreasing long-run marginal costs in output, but increasing marginal costs in quality.

The possibility that marginal costs are decreasing in output, but increasing in service quality, raises an interesting policy issue for the

subsidization of government enterprises. The temptation here is to suggest that output, but not service quality, should be subsidized. But this clearly over-simplifies the problem since output cannot be defined absent some given quality level. As a practical matter, however, it is possible to consider marginal changes in service quality -- relative to some baseline level -- at a given level of output. The policy issue here is: Should improvements in service quality be subsidized given increasing marginal costs? Absent other considerations -- such as market failures distinct from the natural monopoly case, or distributional impacts -- it is difficult to rationalize direct subsidies for service quality improvements.

True external economies cannot be appropriated by the enterprise. As a result, the enterprise will tend to under-allocate resources to the production of output (or quality). Direct subsidization of the producer is one means of achieving allocative efficiency in this case. Lump sum subsidies are obviously ineffective in this situation since they provide no incentives for increasing either output or service quality.

It should also be noted that it may be possible that output changes generate external economies, but service quality changes do not. This will occur if the benefits of service quality accrue only to direct users. Of course, improvements in service quality may encourage consumption, and this in turn may generate external economies.

Similarly, the producer cannot appropriate benefits if the service can be characterized as a public good (or service). Again, direct subsidies can be employed to promote allocative efficiency in this case. Rail passenger service, however, cannot be classified as a public good <u>per se</u>, since individuals can be excluded from consumption. The public good

problem can arise, however, if members of society place existence value on rail passenger service. In this case, exclusion is not possible and subsidies may be appropriate to promote allocative efficiency.

Nonetheless, caution should be exercised in considering the appropriateness of subsidies to address the existence value problem. Direct subsidies are appropriate here only if existence value is affected by marginal changes in service quality or quantity.* If existence value is a function of the quantity of services available, then the direct subsidy should be tied to the level of output. Similarly, if existence value depends on service quality, then the direct subsidy should depend on the level of quality provided by the enterprise. In either case, the per-unit subsidy (per unit of output or quality) should equal marginal existence value.

Option value also raises some interesting issues related to the public good problem. Strictly speaking, option value does not, by itself, create a public good problem since it is possible to organize markets in which the producer can appropriate contingency benefits. Indeed, there are many examples of formally organized options markets.

A formal option market for rail passenger service does not currently exist, of course. Although it may theoretically be possible to organize such a market, transactions costs and other practical considerations make

---

* This is not to say that non-marginal existence value benefits should not be included in assessing projects within a benefit-cost analysis framework.

it difficult to envision an operational framework under which VIA Rail could appropriate option values.*

As was the case for existence value, caution should be exercised when considering the appropriateness of subsidies to address the option value problem. Specifically, option value must be affected by the marginal changes in output or improvements in service quality. Again, the direct subsidy should be tied either to output or service quality.

Subsidies may also be appropriate if distributional impacts are included in the objectives of the enterprise. As was noted earlier, consideration of distributional impacts require the decision-maker to assign a value (or benefit) to alternative welfare distribution regimes. If subsidies are employed because of objectives related to distributional impacts, lump sum subsidies should be avoided, especially those paid *ex post*. Lump sum subsidies do not provide direct incentives to improve service quality; *ex post* subsidies do not provide incentives to produce services at minimum cost.

2.2 The Competitive Case

Rosen (1974) provides an interpretation of the hedonic pricing model that is useful to describe how quality is determined in competitive markets. Perhaps more important, the model can be used to describe the conditions necessary for the existence of perfect competition in quality

---

* Under such an arrangement, VIA Rail would have to sell option passes to potential future users that would permit them to purchase tickets at some future date.

markets. Because these conditions are so restrictive, it is likely that quality markets that might otherwise be competitive resemble monopolistic competition in the real world.

Rosen begins by describing a market for a product that has a quoted price which implicitly reveals a function that relates a price and characteristics of the product. Specifically, the hedonic price equation is written:

$$P(z) = P(z_1, z_2, \ldots, z_n) \tag{2.1}$$

where $z_1, z_2, \ldots, z_n$ are the "quality" characteristics of the good or service. The hedonic price equation describes the minimum price at which a good or service with a given bundle of quality characteristics will be available to consumers. Since producers can increase quality only by using more resources, P(z) is increasing in all its arguments.

It can be shown that P(z) represents a trace of market equilibria for quality characteristics under the assumption that consumers maximize utility and producers maximize profits. First, write the consumer's utility function as:

$$U = U(x, z_1, z_2, \ldots, z_n) \tag{2.2}$$

where x is all other goods. If the price of x is set to one, the consumer's problem is to maximize utility subject to:

$$y = x + P(z) \tag{2.3}$$

where y is the consumer's budget. Note that since P(z) may be nonlinear, the budget constraint may also be nonlinear.

Next, define a bid function

$$\theta (z_1, z_2, \ldots, z_n; u, y) \tag{2.4}$$

which describes the amount the consumer is willing to pay for the various quality attributes at a given level of utility and income. Utility is maximized when $\theta (z^*; u^*, y) = P(z^*)$ and $\theta z_i (z^*; u^*, y) = P_i(z^*)$ or, in other words, when the bid function is tangent to the hedonic price function. This situation is depicted in Figure 2-1 for quality attribute z_i, given an optimal bundle of other characteristics.

The producer's problem is to select a bundle of attributes and an output level, Q, at which profits are maximized. That is,

$$\text{Maximize } \pi = Q \cdot P(z) - C(Q,z) \tag{2.5}$$

The next task is to relate the producer behavior to the implicit price equation. To do this define an offer function as:

$$\phi (z_i, \ldots, z_n; \pi) \tag{2.6}$$

The offer function is formed by eliminating Q from Equation 2.5 and solving for the optimal bid in terms of the quality characteristics of a product. In short, the offer function defines the set of unit prices that the firm is willing to accept for alternative designs of a product, given that profits are maximized with respect to optimum quantities.

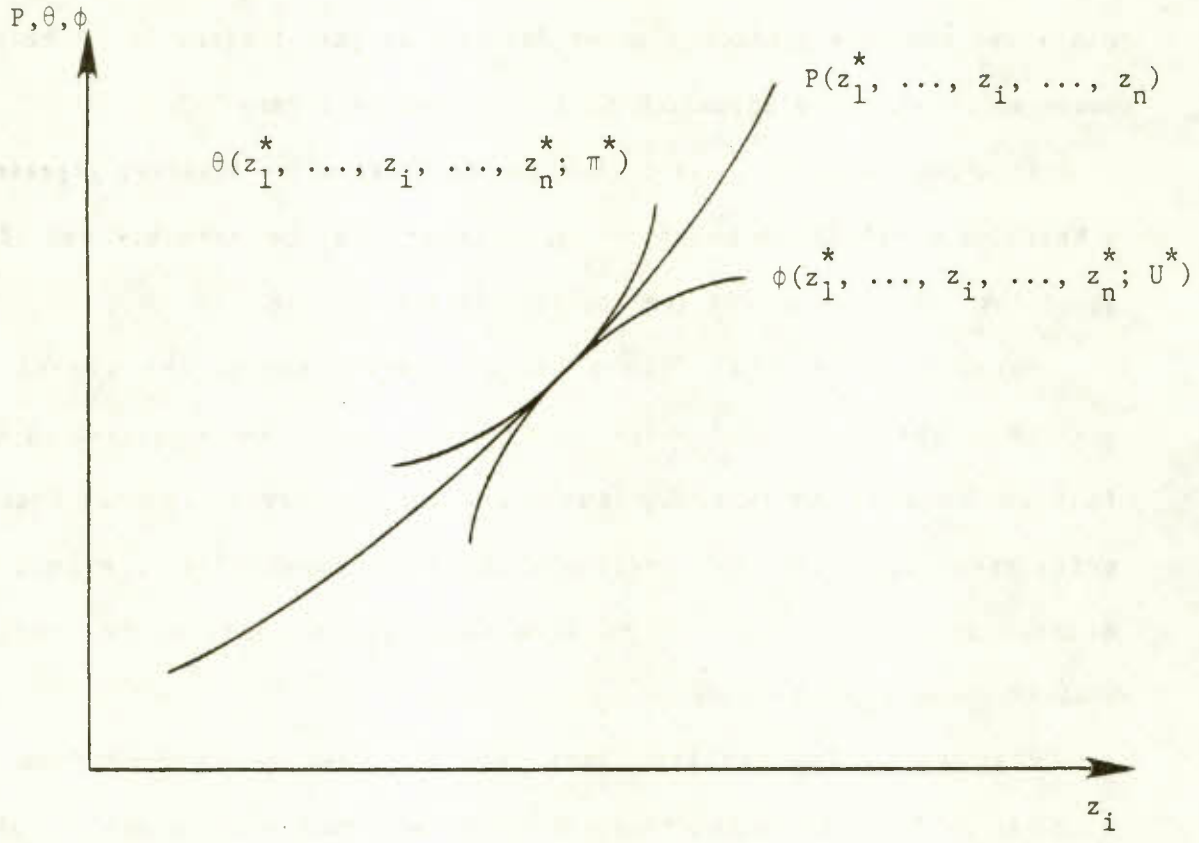


Figure 2-1. Bid and Offer Functions at Equilibrium

Of course, the producer is constrained by the set of market prices defined by the implicit or hedonic price equation. Thus, profits will be maximized when the producer's offer function is just tangent to the hedonic price equation. This situation is illustrated in Figure 2-1.

It should now be apparent that the implicit price equation represents a trace of equilibrium points -- i.e., tangencies between bid and offer functions — in the market for quality characteristics.

Note, however, that Figure 2-1 does not describe the market for quality in the traditional sense. The market can be characterized in this fashion, however, by deriving demand and supply curves from the hedonic price equation. Once such quality markets are identified, conventional methods to evaluate welfare changes under policy options focusing on quality changes can be made.

The demand for quality characteristics can be derived from the marginal implicit price equation for a consumer with a given set of tastes or preferences characterized by α. The marginal implicit price equation is formed as the first derivative of the hedonic equation with respect to some quality characteristic z_i, and can be identified for a given consumer through demand shifters described by α. More specifically, the demand for a consumer with a given set of tastes can be written:

$$D_{zi} = P_{zi}(z; \alpha) \tag{2.7}$$

Note that the marginal implicit prices and the quality market play the same role as ordinary prices do in conventional quantity markets. In brief,

they indicate the level of quality that will be demanded by the consumer at alternative implicit prices of quality imbedded in the product or service.

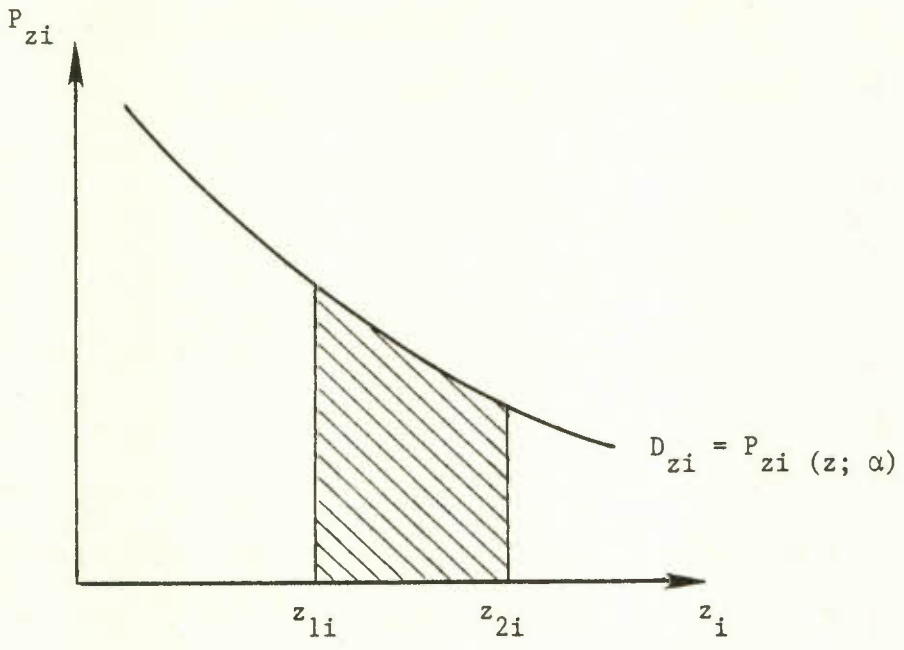
The demand curve for a single consumer for quality characteristics C_i is illustrated in Panel A of Figure 2-2. Changes in consumer welfare associated with an improvement in the quality of the product can be approximated by the shaded area in this figure.

Similarly, the quality supply curve for a given firm can also be formed from the hedonic price equation. Again, this is formed as the first derivative of the hedonic price equation with respect to some quality attribute, and is identified through a set of supply shifters characterized by β. Briefly, β is a vector of factors that cause differences in costs across producers in the market. The supply equation of a single firm for quality attribute z_i can be written:

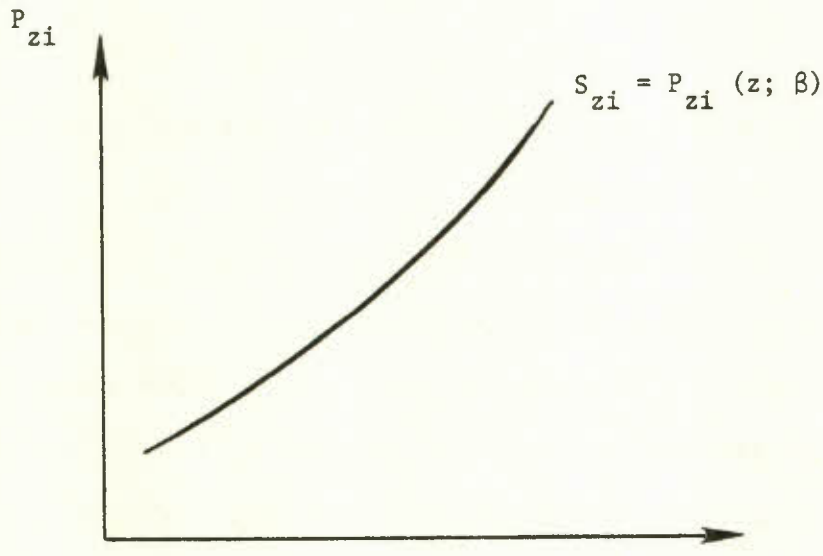
$$S_{zi} = P_{zi}(z; \beta) \tag{2.8}$$

The supply curve for a single producer with costs characterized by β is illustrated in Panel B in Figure 2-2.

Market demand in a neighborhood of any given quality level will be determined by the distribution of consumer tastes and incomes, as consumers solve the constrained utility maximization problem. Similarly, market supply in the neighborhood of a given quality level will be determined by weighting individual firms' supply by the quality distribution. What results is a series of markets at alternative quality levels in which the price and output of a quantity of goods embodying different quality levels are determined.



Panel (a)



Panel (b)

Figure 2-2. Demand and Supply in the Quality Market

The critical question here is twofold: 1) under what conditions will the market for products embodying alternative levels of quality be competitive; and 2) under what conditions will markets achieve allocated efficiency with respect to quality? As the discussion below suggests, the answer to these questions is somewhat more complex than the case in which allocative efficiency with respect to only output is considered. The results have important implications for policy directed to assuring optimum quality levels in regulated markets.

There are two conditions under which quality markets will be competitive and allocative efficiency will be assured. These conditions are:

- An entire spectrum of quality is provided in the market and no single buyer or seller at any point along the quality spectrum can affect market price.
- All consumers have identical tastes and incomes and no one buyer or seller can affect price at the single quality level provided in the market.

It is easy to see that the standard competitive results will be obtained under the first condition. In this scenario, markets are characterized by a series of individual markets in which output and prices are determined competitively for each marginal change in quality. Consumers will be free to select any quality level they prefer, and all buyers in each submarket are marginal consumers. Competitive conditions will assure, in the long run, that the optimum level of output for each quality submarket will be obtained.

Similarly, if all consumers have identical tastes and incomes, only one distinct market in which product with a given set of quality attributes

will be produced. Since all consumers have identical tastes, every consumer will be a marginal buyer and competitive conditions will assure that both the optimum quantity and quality of the product or service will be produced.*

The basic problem is that neither of the two conditions described above are likely to be satisfied in real markets. First, and perhaps most obviously, consumer tastes and incomes differ. Second, the entire spectrum of quality is not likely to be provided in real markets. This is primarily due to the presence of fixed costs associated with the provision of a given quality level. In order to have an entire spectrum of quality provided in competitive markets, the level of consumer demand at each point in the quality spectrum must be sufficient to cover the fixed costs of a large number of producers for each quality level. This, of course, is rarely the case; as a result, only one or a few firms supply the market at a given quality level, and the alternative levels of quality provided on the market are usually limited. In either case, efficiency problems related to monopoly power or inframarginal buyers occur.

If the level of demand at each point in the quality spectrum is not sufficient to support many firms, markets that might otherwise be competitive take on a structure resembling monopolistic competition. Although a limited range of quality alternatives may be provided in the market, each quality alternative is provided by a single firm that differentiates its product along a quality spectrum. It is well known that monopolistically

---

* Allocative efficiency will also occur if consumers can be grouped by identical preferences and only a limited variety of quality levels are produced. As is explained later, however, this is not the case if several consumer groups exist and technology limits the quality choice to one.

competitive firms do not achieve allocative efficiency in output markets (i.e., they do not produce an output level at a point at which price and marginal costs are equal). It is also true that they will not tend to provide quality up to the point at which price and the marginal cost of quality are equated (given the assumption that additional resources are required to improve the quality embedded in the product). Perhaps the most disturbing outcome here, however, is that equating price and the marginal cost of quality does not necessary ensure that the optimum level of quality will be provided in the market. This problem is discussed in more detail in Section 2.3.

The fact that only a limited number of quality options are made available in the market also complicates the analysis. The problem is that in order to evaluate the welfare effects of changes in quality, one must consider tradeoffs in welfare gains and losses between marginal and infra-marginal buyers. This problem is illustrated in Figure 2-3. Suppose, for the sake of illustration, that we consider the demand for a trip between a given origin-destination pair. The quality aspects of the trip that matter to consumers are reliability (e.g., the probability of on-time arrival) and passenger comfort. Suppose further that there are only two options available in the market and that these are characterized by Alternative Mode A and Alternative Mode B.

The slopes of the line segments OA and OB are equal to the ratios of reliability and comfort that are embodied in the two alternative transportation modes. Alternative A embodies a relatively larger amount of reliability than Alternative B, which in turn, embodies a greater level of passenger comfort. The line segment EG is the efficiency frontier; that

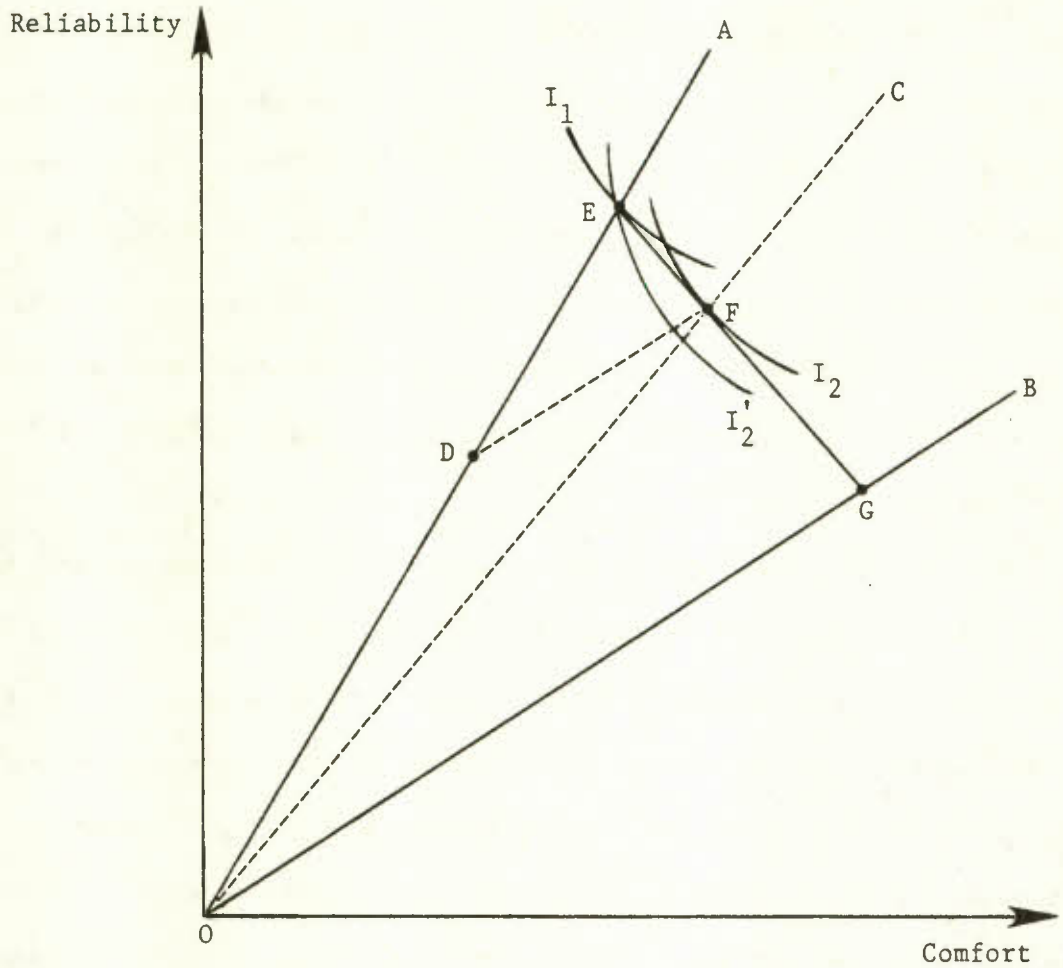


Figure 2-3. Restricted Quality Choice

is, it represents the amount of reliability and comfort that can be purchased with a fixed consumer budget.

Let indifference curve I_1 reflect the preferences of the first consumer for reliability and comfort (i.e., I_1 reflects the first consumers marginal rate of substitution between reliability and passenger comfort). The first consumer maximizes utility subject to the budget constraint by selecting Mode A at point E, at which the slope of the efficiency frontier and the indifference curve are equal.

The second consumer, however, with indifference curves labeled I_2, is an inframarginal buyer. This consumer would be an equilibrium at point F if Alternative C was available in the market. Alternative C embodies more comfort than Alternative A, and more reliability than Alternative B. However, if Alternative Mode C is not available in the market, the best that the second consumer can do is to select Alternative A at point E. Clearly, however, this consumer would be better off at point F. It is important to note that market experiments in the neighborhood of the existing price and quality level will not reveal the second consumer's true preferences for reliability and comfort. For example, marginal changes in the relative prices of the two alternative transportation modes -- which will cause the efficiency frontier to rotate -- will not reflect tangencies to the consumers indifference curve. This is not the case for the first consumer who is a marginal buyer. It is this problem that leads Spence (1975) to conclude that market experiments in the neighborhood of existing price and quality levels do not provide sufficient information to evaluate alternative policies directed to assuring optimum quality levels. He suggests that consumer surveys are required to measure welfare changes for infra-

marginal buyers. This issue is discussed in more detail later in this report.

It should be noted that the inframarginal buyer could reach point F in Figure 2-3 if he or she consumed OD amount of reliability and comfort embodied in Alternative A and then EF amount embodied in Alternative B. In short, it may be possible for the consumer to adjust behavior in markets in which limited quality options are provided by consuming a mix of products available in the market. This type of mitigating behavior, however, is only possible if all that matters to the consumer is the total amount of reliability and comfort that is consumed. In general, however, this will not be the case. That is to say, the consumer is not likely to be indifferent between two trips each embodying the same mix of reliability and comfort, and one trip a relatively high level of reliability and a second with a relatively high amount of comfort.*

The implications of limited quality options are directly relevant to the VIA Rail Case. In general, it is not feasible to provide a range of quality options for rail passenger service. For example, reliability for a given trip must be set a single level for all passengers making the trip. Indeed, given the interdependent nature of the rail passenger system, it is probably not even technically feasible to offer alternative levels of reliability over different trips. Similarly, aspects of passenger comfort such as temperature control and cleanliness must be fixed at a given level for large groups of passengers.

---

* See Hendler (1975).

It should be recognized that the passenger transport industry has long recognized differences in consumer tastes for the quality of service embodied in the trip. First class and coach accommodations represent one attempt by various segments of the industry to offer a range of alternative services. Airlines, in particular, offer a number of improved services for passengers who opt to fly first class. Taxi cabs tend to congregate at locations where large tips are anticipated, thus accommodating some passengers' preferences for reduced waiting times.

Because of fixed costs and technological constraints, however, most passenger services, including VIA Rail, can offer only limited variations in quality of service. This means that there are really two distinct but related policy decisions that must be made. That is, policy must decide the appropriate range of quality alternatives and the appropriate level of quality that should be embodied in each alternative.

2.3 The Monopoly Case

The discussion below provides a description of how the monopolist establishes a profit maximizing quality level. This profit maximizing quality level is then compared to the socially optimum quality level. The monopoly case is especially relevant for VIA Rail policy. Under existing institutional arrangements, VIA Rail is forced to negotiate with single suppliers for many of the services it provides to its passengers. Thus, the incentive structure for the monopolist to provide quality must be understood in order to establish appropriate incentives to provide socially optimum quality levels.

The major conclusions described below are as follows:

- The socially optimum quality level is achieved when the average willingness to pay for a marginal improvement in quality is just equal to the average marginal cost of the quality improvement.
- The unconstrained monopolist may produce a quality level either higher or lower than the socially optimum level.
- Even if the monopolist is forced to produce at a point at which price equals the marginal cost of quality, the socially optimum level of quality may not be produced.
- If price of output is set for the monopolist, but not quality, then the monopolist will always produce a quality level lower than the socially optimum level.

Each of these points is discussed below in more detail.

The model employed below to show monopoly incentives to provide quality follows Spence (1983). It is assumed that the monopolist has three decision variables, price (P), quantity (Q), and quality (z). For simplicity, it is also assumed that the product embodies only a single quality dimension.

The monopolist's profit function can be written:

$$\pi = Q \cdot P(Q, z) - C(Q, z) \qquad (2.9)$$

where P(Q, z) and C(Q, z) are the firm's inverse demand and cost functions, respectively.

For a given quantity, the monopolist will produce the profit maximizing level of quality when

$$\frac{\partial \pi}{\partial z} = QP_z - C_z = 0. \tag{2.10}$$

In words, the monopolist maximizes profits when the marginal revenue produce of quality is just equal to its marginal cost.

Next define consumer surplus S, as

$$S = \int_0^Q P(V, z)\, dv - Q \cdot P(Q, Z) \tag{2.11}$$

Note that consumer benefits are defined by P(V, z) and not P(Q, z). The atter, which represents the firm's inverse demand function, is determined by the marginal willingness-to-pay of marginal buyers. P(V, z), on the other hand, reflects the benefits received by inframarginal consumers as changes in quality occur.

Given Equations 2.10 and 2.11, total surplus, W, can be written:

$$W = S + \pi \tag{2.12}$$

when

$$\frac{\partial W}{\partial z} = \int_0^X P_z\, dv - C_z = 0. \tag{2.13}$$

Equation 2.13 states that the socially optimum level of quality is obtained when the total consumer benefits associated with a marginal change in quality are just equal to the marginal cost of improved quality, or equiva-

lently, when the average consumer valuation (at the margin) is equal to the average (per unit of output) marginal cost of the quality change.

To see how this result compares with the profit maximizing quality change, note that

$$\frac{\partial W}{\partial z} = \frac{\partial S}{\partial z} + \frac{\partial \pi}{\partial z} = \int_0^Q P_z \, dv - Q \cdot P_z + \frac{\partial \pi}{\partial z} \qquad (2.14)$$

When $\partial\pi/\partial z = 0$, the sign of $\partial w/\partial z$ will depend on the relative magnitudes of $\int P_z \, dv$ and $Q \cdot P_z$. That is, the monopolist will produce too little or too much quality as $1/Q \int P_z \, dv \gtrless P_z$. Note that the first expression is the average consumer valuation of the quality change (at the margin) while P_z represents the valuation of marginal consumers. Thus, the monopolist will select the socially optimum quality only if marginal consumers are representative of all consumers.

It is important to note that the quality problem exists independently of the monopolist's tendency to restrict output. In fact, even if the monopolist is required to price at the marginal cost of quality (and output), no assurance can be given that the optimum quality will be provided.* In short, the resource misallocation occurs because the monopolist's revenues and hence, profits, depend on the behavior of marginal consumers. Indeed, bids by marginal consumers can cause the monopolist to produce too much quality.

Another case relevant to this study is the situation in which the monopolist is constrained to charging a fixed price. In this case, the

---

* See Spence (1983). The proof is too lengthy to be reproduced here.

monopolist will always produce too little quality. To see this, rewrite the firm's revenue function in terms of the demand function, Q(p, z), and note that

$$\frac{\partial W}{\partial z} = \int_{p}^{\infty} Q_z (V, z)\, dv + \frac{\partial \pi}{\partial z} > \frac{\partial \pi}{\partial z} \qquad (2.15)$$

Accordingly, when $\partial\pi/\partial z = 0$ (i.e., maximum profits), $\partial w/\partial z > 0$ and society will benefit from the production of improved quality.

2.4 Intermediate Cases

When quality, along with price and output, are considered endogenous in markets, a number of interesting intermediate cases arise. Although these intermediate cases are presented below for the sake of completeness, they all have applications in different regulatory contexts, and some are directly relevant to the VIA Rail case.

In particular, the following intermediate cases are described:

- Fixed prices in markets with rival or competing firms.
- Price adjustments in markets with constrained quality levels.
- Quality and monopolistic competition.

Each of these cases is described below in more detail.

2.4.1 Fixed Prices with Rival Firms —

If only price and output are considered endogenous in the market, economic theory predicts that price setting (i.e., setting the price per unit of output) by a regulatory authority will create either a surplus or a shortage, depending on whether the regulated price is set above or below the market clearing price. Indeed, this theory accurately predicts the result observed in U.S. agricultural markets in which Federal price supports have regularly created substantial surpluses. However, another and equally valid view, is that the surpluses in these markets are the result of farmers' inability to compete along the quality spectrum. That is to say, the shortage or surplus in a fixed price market will result only if sellers are unable to vary the quality as well as the quantity of output.

Kahn (1970) notes that if price is prevented from falling to marginal costs in a short run, or average costs in a long run, then sellers in a competitive market will, to the extent that they are able, increase the quality embodied in the good or service until profits are eliminated. Conversely, if the regulatory authority sets prices too low in the market, sellers will respond by reducing the quality of the good or service. Thus, when quality is considered as a decision variable for the firm, it generates behavior which allows the market to clear at any fixed price. Given that all consumers can be treated as marginal consumers, fixing price at any level other than the competitive equilibrium price will cause the market to produce either greater or less quality than the socially optimum level.

It is also important to note that firms will adjust quality levels in the face of fixed prices in a manner that will tend to eliminate profits, even if entry into the industry is restricted. Prior to deregulation, the U.S. airline industry serves as a good example of how service inflation occurs in an industry with fixed prices, even in the absence of freedom of entry to the market. Fairly recently, the American Civil Aeronautics Board (CAB) regulated both prices and entry into the industry in an effort to protect the industry from "ruinous competition". However, because of excessive service inflation, the CAB found it necessary to specify maximum quality parameters in the industry such as maximum distances between seats, quality of free meals, beverages served on tourist flights, and required carriers of the airline industry to charge for in-flight motion pictures.*

However, the CAB paid far more attention to prohibiting price competition than quality of service competition, and there is little doubt that, as a result, the U.S. airline industry produced a greater than optimal level of service quality. As Gellman observes:

> While it is undeniably desirable that the level of service afforded the traveling public be raised continually, it is somewhat ludicrous to find virtually unrestricted service competition prevailing in this industry while prices are more or less rigidly controlled (in Hollander, 1968).

Some authors (see, for example, Trapani and Olson, 1982) correctly predicted that deregulation of the U.S. airline industry would bring both reduced prices, and reduced quality of services. What is perhaps more interesting is that deregulation also resulted in a wider variety of ser-

---

* See Kahn (1970), Vol. II, p. 10.

vice quality options being available to consumers. The success of "no frills" airline passenger service illustrates this point. Prior to deregulation, airline passengers could only choose between quality differences embedded in first class and coach accommodations.

There is no doubt that passengers preferring relatively low levels of service quality are better off, given the availability of no frills service. It is important to recognize that such an option would not be available with prices fixed at some level in the market. In short, no single firm has an incentive to offer reduced service quality if it is forced to charge the same price as all other producers in the market.

It is obvious that service inflation will occur in markets in which price is set too high and no barriers to entry exist. The taxi cab market is an interesting study here. If price is set in excess of the costs of existing suppliers, entry into the market will occur in the absence of barriers. In such a market, entry will automatically improve the quality of the service, as passenger waiting times will be reduced. Thus, if price is set too high in such a market, passengers will be forced to accept and pay for service quality that is higher than the optimum level. Conversely, a price set too low will impose a lower than optimum service standard level. Douglas (1972) notes that implicit markets in taxi cab service develop as cabs congregate at hotels and other junctions where high tips are anticipated as a reward for reduced waiting time. This is an attempt by the market to adjust for differences in consumer preferences for service quality.

2.4.2 Restricted Quality —

Another interesting case is market behavior when the quality of goods or services are constrained to a single quality level.* There are really two cases in which this situation can arise. Quality levels can be dictated by a regulatory authority, by a franchisor, or a single quality may be imposed on the market because of technological constraints. For example, the waiting time aspect of taxi cab service is determined by the size of the cab fleet and the number of passengers in a given market area. A single quality level is also imposed on rail passengers with respect to such quality variables as on-time performance and some aspects of passenger comfort.

The fact that technological constraints impose constant quality in the market creates a difficult dilemma for the policy maker or regulator, given that different consumers prefer different levels of service quality. Earlier in this report, it was noted that marginal and inframarginal buyers are likely to have different preferences for service quality. As a result, optimal resource allocation cannot be assured in this case even if quality is produced to the point at which price and marginal cost of service are equated.

The response of competitive markets when constant service quality is imposed is enlightening. Douglas (1972), for example, notes that in unregulated taxi cab markets, either one of two sets of competitive equilibrium prices will be observed if consumers are grouped into two by

---

* In most cases, the arguments offered below will also hold if the number of quality options is restricted to some finite number.

willingness to pay for reduced waiting time. Passengers with high values of time, for example, will seek to reduce waiting time by bidding up taxi fares. This, however, will cause consumers with low preferences for reduced waiting time to exit the market, thus causing taxis to exit the market. This, in turn, will reduce the quality of service as waiting times increase. This same phenomenon was recognized later by Rosen (1974) in his interpretation of the hedonic pricing model. Specifically, Rosen notes that a condition for stable equilibrium is that consumers not be clustered at different points reflecting varying preferences for the quality embedded in products.

2.4.3 Monopolistic Competition —

As was noted earlier in this section, the dimension of quality often generates market structures that resemble monopolistic competition when they might otherwise be competitive. There are two ways that this can happen. In the first case, the level of demand at any given point along the quality spectrum is not sufficiently large to support more than a single producer. In this case, several firms produce in the market, but each has its product differentiated by varying levels of embodied quality. The market is monopolistically competitive because, presumably, the output produced by any one firm competes with close substitutes.

The second case arises when quality of service itself is affected by the level of capacity in the industry. In this case, firms tend to look like monopolistic competitors even when constant quality is imposed on the market. DeVanney (1975), and DeVanney and Saving (1983) analyze taxi cab

markets in which the quality of service in the market is affected by industry capacity through its impact on waiting times. The implications of their analyses, however, are much broader than a narrow application to taxi cab markets.

For example, trip frequency and the availability of excess seats (i.e., the load factor) affect the quality of services embodied in passenger transport services. Ippolito (1981), and Trapani and Olson (1982) provide empirical evidence that these quality factors affect the demand for airline travel. The presumption here is that trip frequency and excess seating capacity also affect the perceived quality of received services offered to railroad passengers. In any event, the point is that the perceived quality of services offered in the market can be improved by expanding capacity in the industry.

DeVanney and Saving (1983) show that when service quality is affected by industry capacity, individual firms in the long run tend to look like monopolistic competitors in that the demand curve is tangent to the downward-sloping portion of the average cost curve (when output is measured along the quality dimension). This occurs even when constant quality is imposed in the market. Thus, it appears that the market is inefficient in that excess capacity exists, even the long run.

However, DeVanny and Saving also demonstrate that this long-run solution is socially optimum in that the cost of excess capacity will be exactly equal to the increased benefits of improved service-quality (i.e., reduced waiting times). This finding is consistent with an earlier discovery by DeVanny (1976) that the monopolist will provide the optimum level of capacity with respect to service quality if profits are zero. It is

important to stress, however, that in both cases, it is assumed that all consumers have identical preferences for service quality. As was noted earlier in this section, the presence of inframarginal buyers means that these markets may not automatically achieve the optimum level of service quality since, in maximizing profits, firms always respond to marginal instead of inframarginal consumers.

3. ALTERNATIVES FOR VIA RAIL AS A SERVICE BROKER: OPTIONS FOR REGULATING QUALITY

Several alternatives for VIA Rail to regulate the quality of services offered to rail passengers are described below. The implied model is one in which VIA Rail is viewed as a broker for passenger rail services. That is, VIA Rail purchases rail services or inputs to rail services from private sector firms, and then sells these to rail passengers.

It is recognized, of course, that characterizing VIA Rail simply as a passenger service broker somewhat simplifies current institutional arrangements. For example, VIA Rail purchases services from two rail carriers -- the Canadian Pacific and Canadian National Railroads -- such as track right-of-way, operating crews, etc. VIA Rail also can be viewed as a producer in that it provides equipment, on-board service personnel, ticket agents, and marketing and management services.*

Nonetheless, the discussion is sufficiently general to be applicable to a wide range of possible alternatives for VIA Rail participation in the production process itself. For example, one alternative scenario is be one in which VIA Rail acts purely as a service broker and represents rail passengers simply as a bargaining agent in purchasing all inputs required for the production of passenger services from private sector firms. An alternative scenario is one in which VIA Rail itself produces all passenger and related services. In any case, the question here is how to design a

---

* VIA Rail can be considered a producer under this scenario because services purchased from the CN and CP can be viewed as intermediate inputs to the production of final services.

structure which assures that the appropriate level of service quality is produced.

In particular, the following alternative arrangements for VIA Rail as a service broker are explored:

- Performance incentive agreements.
- Rate of return regulation.
- Quality by fiat -- i.e., service quality standards stipulated by contract.
- Competitive tendering.

Each of the alternatives described above are evaluated in terms of their ability to provide incentives for allocating resources efficiently in the production of service quality. In addition, the information needed to effectively implement the various alternatives is also described. At this point, the discussion is abstracted somewhat from difficulties in negotiating contracts and policing providers of services to assure that contractual arrangements are satisfied. The discussion also ignores difficulties in measuring specific aspects of specific service quality. All of these, and other detailed topics are handled in the discussion provided in Section 4 of this report.

In order to limit the discussion to a manageable level, the approaches described below are evaluated in terms of their potential effects on allocative efficiency under the assumption that no non-user benefits associated with service quality exist. That is, the evaluation weighs the willingness to pay for service quality by rail passengers against the costs

of providing service quality. It is recognized that achieving allocative efficiency may be only one of several possibly conflicting objectives of VIA Rail. For example, the subsidization of passenger rail service may be justified in terms of achieving a socially desirable distribution of income (e.g., subsidize passenger rail services because they are used by members of society in lower income classes). As was noted earlier in this report, subsidies may also be justified if it is believed that rail passenger service generates benefits to society at large which are not captured directly by rail passengers.

In any event, the discussion provided below is sufficiently general to encompass distributional objectives. For example, to the extent that these objectives are appropriate concerns for VIA Rail, the willingness to pay for service quality by rail passengers must be adjusted to account for other social goals. In short, willingness to pay by rail passengers is taken as the benchmark for the social benefits associated with service quality, and to the extent that other benefits are associated with quality, these other benefits must be added to passengers' willingness to pay.

3.1 Performance Incentive Agreements

The design of an incentive scheme to provide the socially optimum level of quality is described below. The scenario is couched in a setting in which VIA Rail negotiates a contract which provides inducements to the

firm supplying services to provide the appropriate level of service quality.*

More specifically, consider VIA Rail negotiating a contract under the following setting:

- VIA Rail negotiates with a single producer for a given service which has one quality dimension.
- VIA Rail and the firm have already negotiated the level output to be provided, and the price of that output at some baseline or minimum quality standard.
- VIA Rail sets the price for marginal improvements in service quality, and the firm then decides which level of quality to produce.
- VIA Rail's objective is to achieve the socially optimum level of service quality.

The lessons learned in the previous section on how the single supplier in the market, the monopolist, allocates resources for the production of quality are useful in developing an optimal incentive scheme. The problems are illustrated in Figure 3-1. First, if the monopolist is free to maximize profits with respect to quality, it will interpret the demand curve it faces in the market for quality as the schedule of prices and quality that reflects preferences of marginal buyers. This is shown as the downward-sloping curve labeled in Figure 3-1 as $P_z = D_z$. The corresponding marginal revenue curve facing the monopolist is labeled M_z in Figure 3-1. Given an

---

* It is recognized here that, in fact, VIA Rail faces considerable institutional constraints in its ability to negotiate with the CN and CP. These are discussed in detail in Section 5 of this report.

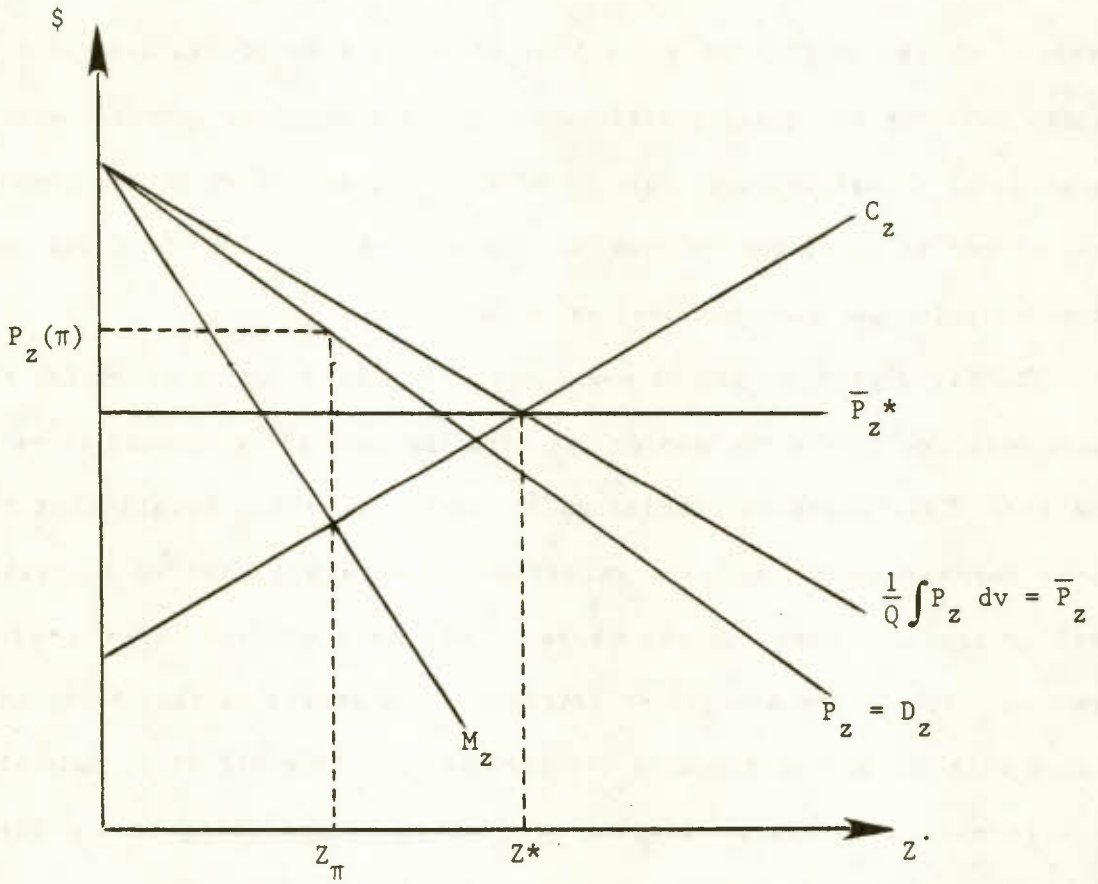


Figure 3-1. Optimal Service Quality Price Schedule

upward-sloping marginal cost of quality curve, C_z, the monopolist will maximize profits by producing at quality level Z_π.

Given the example illustrated in Figure 3-1, the monopolist will underallocate resources to the production of quality for two reasons: 1) it follows the demand curve associated only with marginal buyers in the market, which, in this case, is less than the true social demand curve (which reflects the quality preferences of inframarginal buyers); and, 2) it produces a quality level associated with a point at which the marginal revenue and marginal cost of quality are equated, instead of at the point at which price and marginal cost are equated.

The first problem can be overcome if VIA Rail, as an agent for rail passengers, confronts the monopolist with the true social demand curve for quality. This curve is labeled as $\bar{P}_z$ in Figure 3-1. Recall that this curve represents the average valuation of marginal changes in service quality for all buyers in the market, both marginal and inframarginal. However, even if the monopolist interprets its demand as that being associated with the social demand curve for quality, it would still underallocate resources in the production of quality by producing at a point at which $\bar{P}_z$ exceeds the marginal cost of quality.

The critical question then can be stated: How can the monopolist be induced to produce the socially optimum level of quality -- labeled Z^* in Figure 3-1? The answer here is relatively straightforward. VIA Rail must negotiate with the monopolist, a fixed price per unit of quality exactly equal to $\bar{P}_z^*$. Given that the monopolist faces such a price schedule, it will automatically produce the optimum level of quality, since this output

level will be consistent with profit maximization under a quality fixed price regime.*

It is also obvious that if VIA Rail sets the per unit price of quality too high, the monopolist will have every incentive to produce too much quality, and conversely. It is important, however, not to lose sight of the general case here; specifically, it is not always true that the monopolist, free to set any price it wishes for quality, will not necessarily provide service which embodies too little quality. The degree to which the monopolist will over- or underallocate resources to the production of quality will depend on the net impact of differences in preferences in between marginal and infra marginal buyers, and the tendency for the monopolist to produce at a point at which price exceeds marginal cost. In short, it is possible for the monopolist to produce too much quality if marginal buyers value improvements in quality more than inframarginal buyers.

Perhaps even more relevant to this case, however, is the fact that the monopolist faced with a fixed price per unit of output schedule -- and a fixed level of demand at that -- will <u>always</u> underallocate resources to the production of quality. Indeed, this is the relevant case here, since under the initial assumption, an initial contract for a fixed price and fixed output (i.e., the number of trips) has already been negotiated. In short, the firm is likely to produce even less quality than the unregulated monopolist, since it receives no additional revenue for improving quality.

It is also important to recognize that the optimum unit price schedule for quality does not assure that the monopolist will produce at a point at

---

* There are some cases described later in Section 4 for which nonlinear penalty schedules may be appropriate.

which profits equal zero, even if the contract negotiated for baseline or minimum quality services exactly covers baseline avoidable costs. Suppose, for example, that the optimum unit price schedule for quality leaves the monopolist in a position at which it earns profits on the production of quality above the minimum or baseline standard. If VIA Rail wishes to negotiate a total contract with the monopolist which eliminates all economic profit, then it must deduct a lump sum payment from the baseline agreement.

This lesson is simple, but important. Under current negotiating rules, VIA Rail is required to negotiate a baseline contract that must cover theoretically, at least, some measure of the railroad's avoidable costs. The above example makes it clear that VIA Rail cannot generally assure that the optimum level of quality is provided, and still leave the monopolist in a position at which no economic profit is earned.

This finding is important in another respect. VIA Rail (and Amtrak) negotiate into its contracts penalties for deterioration of on-time performance below a minimum or baseline standard. However, in Amtrak's case, these penalties can only be assessed against previously earned incentive credits; in VIA Rail's case, penalties can be assessed against credits earned by other trains, but total penalties in any period cannot exceed total earned performance bonuses. That is to say, the penalties for performance levels below the baseline standard cannot leave the monopolist in a position where it would receive total compensation less than the negotiated avoidable cost payment.

Such an arrangement has been criticized by several observers of Amtrak operations. But one important point seems often to be obscured. As a

practical matter, the restriction on performance penalties not only reduces incentives for railroads to provide service quality but it also restricts the ability of VIA Rail (and Amtrak) to negotiate contracts which assure appropriate levels of service quality, and at the same time eliminate economic profits for railroads.*

There is another comment that is appropriate at this point. Specifically, some observers of the industry have suggested that national passenger service corporations can watch to see if a negotiated contract which provides service quality incentives produces profits for the carrier, and then negotiate downward the incentives in the subsequent contract. Such adjustments, however, must be made through changes in lump-sum payments, and not through the incentive schedule itself. Given a lower unit incentive price for service quality, the monopolist will automatically respond by producing less quantity if costs are increasing in quality. In short, the per unit price schedule for service quality should be adjusted upward or downward only if VIA Rail guesses wrong on either the appropriate price of quality, or the marginal cost of quality, but not to reduce profits earned by railroads. Indeed, such a move will be counterproductive.

In the example described immediately above, VIA Rail played the role of a negotiator, (i.e., a broker) buying passenger services from a rail carrier. The analysis, however, can easily be extended to a counterfactual case in which VIA Rail itself is the sole producer. In this case, the

---

* In Amtrak's case, the penalty restriction could eliminate incentives if the carrier's performance fell so low that it lost hope of earning any bonuses during the incentive period. Under current VIA Rail contracts, penalties for service degradation cease below some threshhold level of performance.

optimal price schedule offered to the monopolist would represent the optimal subsidy schedule for VIA Rail itself to provide quality. Again, profits that are viewed as "excessive" should be eliminated through lump-sum deductions from fixed payments and not through reductions in the performance incentive schedule itself.

The above case study also presumes that VIA Rail is able to dominate the negotiations between the two parties, or that it can impose the optimal contract on the sole source supplier. In general, of course, this is not the case. Spence (1975) provides a more general analysis in which the purchasing agent (e.g., VIA Rail) and the monopolist behave as duopolists during the negotiations. Spence concludes that both parties will rationally avoid inferior outcomes in which profits and consumer surplus are unnecessarily sacrificed.

Beyond this, however, a plethora of outcomes -- including Cournot and von Stackelberg equilibria -- are possible. In brief, the possible outcomes are so complex, and the informational requirements are so severe, that Spence abandons the model and suggests rate of return regulation as a viable second best alternative to improving service quality. This alternative is described below in the next subsection of this report.

The information required to design the optimum service quality incentive schedule is likely to be difficult to obtain in practice. First, VIA Rail must know the shape of the marginal cost of quality curve in order to establish the optimum incentive schedule. Changes in the railroad's costs as quality varies may be observed -- albeit with some difficulty -- from the carrier's cost records. What is more difficult, however, is to measure the improvement in service quality associated with the change in cost.

On-time performance, for example, is affected by many other factors other than the railroad's direct efforts to assure adherence to schedules (e.g., changes in equipment, improvements in road beds, and weather). These confounding factors are likely to make simple correlations between improvements in on-time performance and marginal costs inappropriate. It should also be noted that, as a general rule, it is necessary for the regulator or VIA Rail to know the shape of the entire marginal cost curve in order to demonstrate the feasibility of the quality improvement (i.e., to assure that the total benefits of the quality improvement are not less than its total costs).

Apart from the difficulty inherent in measuring service quality, the requirement that the negotiator know the firm's marginal cost schedule is not different, in principle, from the information that is required to assure efficient resource allocation in the production of output (i.e., quantity). Informational requirements on the demand side, however, are of a different nature. Specifically, the negotiator must know the valuation of improvement in service quality by inframarginal, as well as marginal, consumers. As was noted earlier, market experiments in the neighborhood in the existing price and quality level are sufficient to identify valuations of marginal consumers.

Such experiments, of course, are not sufficient to identify valuations by inframarginal consumers. As a result, Spence (1975) suggests that surveys be used to identify the social willingness to pay for quality service. It should be stressed, however, that the survey design must account for the complexities involved in obtaining consumer valuations of products which carry with them several quality attributes. Again, in order

to test for the feasibility of the improvement in service quality, one must, in principle, know the entire social willingness to pay curve.

As a practical matter, however, the situation is not hopeless. Only a limited number of viable options can ever be evaluated in applied benefit-cost analyses. Even if the feasibility of providing service quality cannot be evaluated, applied benefit-cost analyses can be employed to evaluate marginal policy changes. In short, any social policy directed to improving service quality for which marginal benefits exceed marginal costs is preferable to the baseline standard. This means that the entire marginal cost and willingness to pay for quality schedules need not be known by the negotiator if the objective is simply to improve resource allocation relative to an existing scenario. Some practical suggestions for achieving this more moderate objective are described in the fifth section of this report.

3.2 Rate of Return Regulation

Several authors, including Kahn (1971), have noted that Averich-Johnson (A-J) effects may serve to mitigate the monopolist's natural tendency to underproduce quality if rate-of-return regulation is employed. In particular, regulated monopolies may have special incentives under rate-of-return regulation to improve service quality if the production of quality is related to capacity (i.e., capital intensive). Kahn suspects that some industries -- most notably, the electric power generating and telephone industries -- may tend to overproduce quality because of A-J affects. In

both of these industries, service quality is related to industry or firm capacity.

Kahn also notes the relative lack of concern about service quality produced by regulated firms, both among regulators and in the professional literature. He suggests three possible reasons for this lack of concern:

- Regulated monopolies are usually assured -- with a regulatory lag -- of receiving the costs of producing quality, and hence, may be less hesitant to improve quality than the unregulated monopolist.
- Service quality often requires an expansion of the rate base, thus providing regulated monopolists an incentive to improve quality (i.e., the A-J affect).
- Regulated monopolists may be especially sensitive to public criticism about service quality.

Of course, none of these factors are likely to cause the regulated monopolist to produce exactly the socially optimum level of quality; rather, they tend to mitigate, to the extent that it exists, the monopolist's tendency to underallocate resources to the production of service quality.

Spence (1975) also suggests rate-of-return regulation as a vehicle for mitigating the monopolist's tendency to underproduce quality. He notes, however, that if the production of service quality is labor-intensive, rate-of-return regulation may induce the monopolist to produce even less service quality than it might otherwise. This occurs because the monopolist regulated under a rate-of-return approach has a natural tendency to substitute capital for labor.

As a "second best" alternative, the rate of return approach is less ambitious than the optimal service quality incentive scheme described

earlier. This being the case, the information requirements for its implementation is also less severe (apart from the usual data required for rate of return regulation). Of course, in order to evaluate the effectiveness of this option in providing the appropriate level of quality, one ideally needs the same data as those described for the design of the optimal incentive quality price schedule.

However, numerous other problems associated with rate-of-return regulation have been well documented in the economic literature. Indeed, many special problems related specifically to the application of rate-of-return regulation to the railroad industry have been noted. A full discussion of these problems is beyond the scope of this study. Nonetheless, the rate-of-return regulatory option is noted here because of its potential impacts on the service quality problem.

3.3 Quality by Fiat

Another VIA Rail option for regulating or assuring the quality of passenger service can be described as "quality by fiat". Under this arrangement, VIA Rail would negotiate with firms a preferred level of service quality, instead of negotiating a contract that merely stipulates implied minimum or baseline service quality levels. Naturally, the price of such an agreement would necessarily be negotiated upward above the price of a contract that implies minimum service quality levels.

In other words, the fixed price of the agreement must necessarily be sufficient to cover the avoidable costs of providing baseline service quality, plus the incremental costs of providing marginal improvements in

service quality that would otherwise be compensated for through service quality incentive payments. Quality is often imposed by such arrangements -- i.e., by fiat -- in the private sector through franchising agreements in cases in which the profit maximizing level of quality for the franchisor are not necessarily the same as that for the franchisee.

Unfortunately, several difficulties, which lie at the heart of the Demsetz-Posner versus Williamson debate on the appropriateness of franchise bidding as a form regulation for natural monopolies, are inherent in this approach.*

These issues involve the following:

- Incentives to produce quality under the "franchise" agreement.
- Negotiating the "franchise" contract.

Both of these issues are discussed below in some detail.

Williamson (1976) notes the difficulty in establishing specific quality standards in a contract, and then agreeing on appropriate measures of such quality standards. But these same problems are inherent in negotiating any type of agreement which attempts to assure service quality. The principle difficulty here is to set up a mechanism which assures that the agreed upon quality standards are enforced. Since the franchise-type agreement represents, de facto, a fixed price arrangement, the franchisee will have economic incentives to shirk on quality in an effort to raise

---

* See Thompson (1980) for a discussion of this problem and an interesting case study on the quality of services at motorway service areas.

profits. The critical question here is whether the franchisor has any recourse to force the franchisee to live up to the standards specified in the agreement.

There are really only two options for enforcing fixed quality standards stipulated in a contractual agreement. These are litigation, and the ability of the franchisor to terminate the contract if it judges that the quality of service is unsatisfactory. The former option seems unsatisfactory, given high litigation costs, and the uncertainty of obtaining reasonable damages through the court system. The ability of the franchisor to terminate the contract is the option that is used most often in the private sector. Reuben (1978) notes that private sector franchise agreements typically give the franchisor a wide degree of latitude in exercising an option to terminate the contract.

It is important to recognize, however, that such an option is viable only if the franchisor -- VIA Rail in this case -- has the ability to obtain services from another competing firm. Since VIA Rail currently negotiates with only two large carriers for the vast majority of all passenger rail services, the option to terminate the contract given unsatisfactory performance does not seem to be an operational alternative. Thus, we conclude that franchise-type agreements in which service quality levels are stipulated at fixed levels during negotiations, are not a viable alternative for VIA Rail unless arrangements are made to promote competitive tendering. As we will argue later in this report, however, options for promoting competitive tendering are somewhat limited.

A second issue related to the franchise-type contract involves problems in negotiating the contract in the first place. There are really two separate sub-issues here. These are:

- Trading service quality against the costs of renegotiating a contract.
- Windfall gains and/or losses occurring to the franchisee as the result of negotiating a favorable contract.

Demsetz (1968) notes that the franchisee may have an incentive to allow service quality to degredate up to the point at which the benefits lost of service quality are equal to the renegotiating costs of the franchisor. That is to say, once the contract is negotiated, the franchisee will have the ability to collect economic rent equal to the cost of negotiating the contract (say with another party). Demsetz argues, however, that this particular problem is not related to allocative efficiency. Instead, he argues that competitive markets, in general, allow economic agents to collect rent of this type. In short, economic rents of this type are unavoidable in competitive markets.

Demsetz also discusses at length the issue of how fixed priced contracts may provide windfall gains (or losses) to the franchisee. These may occur, for example, if input prices increase or decrease more than expected by the parties after the negotiation. He argues, however, that such windfall gains and losses are deemed efficient when conducted in competitive markets in that they are rewards (or penalties) for accepting risk and uncertainty.

It should also be noted that there are vehicles for mitigating such risk and uncertainty. Certainly, the length of the contract agreement can be shortened to reduce the likelihood of such windfall gains or losses. In addition, such contracts often allow for varying payment structures depending on fluctuations in input prices. In short, Demsetz sees no compelling reasons as to why risk or uncertainty should render the franchise-type agreement infeasible or undesirable.

The information required to successfully negotiate a franchise-type agreement at some fixed price is, in theory, similar to the type of information needed to design the optimal incentive structure described earlier in Section 3.1. Obviously, in order to bargain as a representative of rail passengers, VIA Rail must know the average valuation of marginal improvements in service quality for both inframarginal and marginal consumers. It would also seem reasonable to suspect that VIA Rail must also know the marginal cost of quality improvements, in order to negotiate for services at peak effectiveness. This is not an absolute requirement, however. VIA Rail, for example, might be able to learn something about the suppliers expected marginal cost of quality curve during the bargaining session itself. However, when a single supplier is involved, the bid-counterbid scenario reduces the negotiations to the duopoly case described earlier in this section. As we note in the next subsection of this report, one of the primary benefits of extending competitive tendering for services is the ability of VIA Rail to exercise a dominant position in the negotiations for service quality.

It is useful at this point to summarize the above discussions on the franchise-type agreement in which desired service quality levels are stipulated in the contract. The following points should be clear:

- Firms will have economic incentives to reduce quality levels below those specified in the contract, to the extent that it increases their profits.
- This being the case, VIA Rail must have an effective means of enforcing the stipulated quality aspects of service. Enforcement of the contract will be most effective if VIA Rail has the right to terminate the agreement, and negotiate a contract with another supplier.
- Negotiating a franchise-type agreement with a single supplier will cause VIA Rail to become involved in duopoly-type bidding games.
- The information required to effectively implement the franchise-type agreement is similar to that required for constructing the optimal service quality incentive scheme.

Accordingly, the points raised immediately above suggest that the franchise-type agreement will be most effective in a setting in which competitive tendering is feasible.

3.4 Competitive Tendering

As the preceding discussion suggests, competitive tendering can be viewed as a special case of the franchise-type contract. Under this arrangement, VIA Rail would still stipulate the desired level of service quality directly into the contract. The primary difference here is the ability of VIA Rail to negotiate with several suppliers for a given service.

There are three principle, and important, advantages that are afforded by competitive tendering. These are:

- The ability to enforce the service quality aspects of the contract through the threat of terminating the agreement and seeking the services of other suppliers.
- Improved negotiating power which will assure, in the long run, that alternative levels of service quality can be obtained at minimum costs.
- Reduced informational burdens imposed on the negotiator (or franchisor).
- The ability to enforce quality provisions when quality cannot be measured objectively.

The cost of renegotiating contracts when service quality is deemed unsatisfactory is the principle disadvantage inherent in competitive tendering, at least relative to the optimal incentive structure scheme.

Competitive tendering empowers the franchisor to enforce the service quality aspects of the contract through the threat of terminating the contract. As was noted earlier in Section 3.3, the threat of contract termination is only effective if the franchisor has the opportunity to acquire services from other suppliers. This aspect of competitive tendering has been discussed earlier in this section; additional comments here are unnecessary.

Competitive tendering also has the advantage of permitting the franchisor to obtain services -- at any given level of quality -- at minimum cost. There are really two effects here. First, the existence of several suppliers will assure -- absent strategic bidding among suppliers -- that bidders will, in the long run, offer service quality levels at prices that

eliminate monopoly profits. In other words, the existence of several competing suppliers eliminates the need for a negotiator to become involved in the types of duopoly games described earlier. Second, the existence of several competing suppliers will assure that they seek out technologies for providing services that minimize long-run costs. This incentive is less pervasive if only a single supplier dominates the market.

Competitive tendering also reduces the informational burden placed on the negotiator. VIA Rail, for example, would still be required to know how passengers value improvements in service quality. The difficulties described earlier in obtaining such information apply here as well. However, if competitive tendering is a viable option, VIA Rail can be less concerned with estimating the marginal costs of service quality improvements. The reason, of course, is that the costs of quality improvements will become apparent from the alternative bids received by responding firms. That is, to the extent that bids on services reflect a spectrum of price-quality relationships, and suppliers bid at minimum costs, the marginal cost of quality variations will be apparent as they are reflected in competing bids.

The quality incentive price schedule described earlier in Section 3.1 is feasible only if an objective measure of quality can be established. It is difficult, and some impossible, to establish objective measures of many aspects of rail passenger service quality. Factors affecting passenger comfort -- such as car cleanliness, food quality, and the quality of other on-board services -- are examples of these.*

---

* Later in Section 4 of this report, Amtrak's unsuccessful attempt to implement an incentive (penalty) structure for car cleanliness is described.

Competitive tendering can sometimes be a viable alternative when service quality must be measured subjectively. The notion is relatively simple. The franchisor, on behalf of rail passengers, must judge whether the level of service quality is satisfactory. The threat of contract termination, or losing a bid when the contract is renewed, serves as an inducement for the incumbent franchisor to maintain service quality.

Unfortunately, competitive tendering may not be a viable option in all cases. Specifically, competitive tendering may not be viable if bidders incur high sunk costs in providing different rail passenger services.

In the absence of direct performance incentives incumbents can rationally allow service quality to deteriorate to the point at which sunk costs (including the transactions costs associated with negotiating a new contract), just exceed the loss in benefits of service quality to the purchasing agent. Clearly, new bidders will capitalize sunk costs in any bids to takeover services provided by incumbents.

It is important to recognize that the existence of high capital costs does not, by itself, render competitive franchising infeasible. Capital equipment can always be sold off at market value to any successful bidder from a losing incumbent (although this is likely to increase transaction and renegotiating costs). However, substantial investments in human capital may pose a more serious difficulty, because it is more difficult for the investor to resell these assets.*

Because of the advantages inherent in competitive tendering, this alternative is generally preferred to other approaches described in this

---

* Labor union agreements may also place constraints on the ability to tender some services competitively. Amtrak, for example, has argued this in the past.

section of the report. This recommendation, of course, is subject to the caveat that large sunk costs do not render competitive tendering infeasible. It should also be noted that competitive tendering combined with direct performance incentive payments should be considered an appropriate means of promoting service quality. Of course, this mixed approach is practical only if objective measures of performance can be established and monitored.

4. QUALITY IN RAIL PASSENGER SERVICE: A REVIEW OF CONTRACTUAL ARRANGEMENTS

A comprehensive description of present and past practices that have been employed to regulate the quality of rail passenger service is provided in this section of the report.* To the extent that they are relevant to the VIA Rail case, contractual agreements negotiated in the airline and bus industries are also reviewed. As the discussion below suggests, numerous complicating issues arise when one attempts to move from theory — i.e., from conceptually appealing approaches to resolving the quality issue -- to implementation. Many of these difficulties, some of which arise from existing institutional arrangements, are not apparent from the conceptual discussions that have been offered in the preceding two sections of this report. In short, there are several important lessons to be learned from history.

Immediately below, in Section 4.1, some important quality aspects of rail passenger service are identified and described. The methods that will employed to evaluate contractual arrangements with respect to passenger service quality are described next in Section 4.2. Following this, detailed discussions of three general classes of rail service quality -- on-time performance, passenger comfort, and schedule quality — are discussed in Sections 4.3, 4.4, and 4.5, respectively.

---

* The investigation in this section was hampered somewhat by the "proprietary" nature of contracts. The discussion in this section is based on secondary reports, numerous discussions with industry officials and observers, and in some cases, on copies of contracts that we were fortunate enough to obtain.

4.1 Quality Aspects of Rail Passenger Service

Naturally, those aspects of rail service quality that are important will depend on the individual priorities of the passenger. Nonetheless, there can be little doubt that three particular aspects stand out as most important in the view of the majority of rail passengers. These are:

- Reliability -- i.e., on-time performance.
- Passenger comfort.
- Schedule quality.

In addition to these, ticket and reservation services, and other services provided at or about terminals and on-board have received some attention as being important to rail passengers.

By way of documentation, Table 4-1 provides a summary of Amtrak passenger complaints for fiscal years 1977 through 1979. These responses are partially a function of the quality of service actually provided to Amtrak passengers, but nonetheless, they do afford some indication of those aspects of rail service that passengers deem important. On-time performance clearly stands out as the service quality aspect of passenger service that has received the most concern. Indeed, as later discussions in this section indicate, this aspect of rail service quality has received most attention in contractual arrangements between rail passenger service corporations and carriers. One important aspect of passenger comfort -- i.e., temperature control -- has consistently ranked second among Amtrak

Table 4-1

AMTRAK PASSENGER COMPLAINTS: SELECTED YEARS

Regulations	Violations		
	1979	1978	1977
Information to be provided	1	0	2
Reservations	941	862	1,265
Reservation-Making	23	22	49
Reservation-Confirming	2	16	10
On-Time Performance	2,654	2,730	4,085
Expeditious Service	27	32	35
Cancellation of Trains	38	34	6
Cancellation En Route	79	108	154
Thru Car Service	1	12	17
Station Hours	21	33	24
Consist of Stations	343	253	343
Checked Baggage	281	259	451
Consist of Trains	500	339	673
On-Board Services	797	478	1,239
Baggage Service	11	5	16
Food & Beverages	397	279	814
Temperature Control	1,667	1,714	3,455
Functioning Equipment	474	445	592
Car Requirements	1,060	739	1,386
Nonsmoking Space	70	74	110
Complaint Procedure	14	20	13
Track Standards	7	12	18
Total Alleged Violations	12,038	8,466	14,757

Source: U.S. Interstate Commerce Commission, Annual Reports.

passenger complaints.* Schedule quality is not included among the list of Amtrak complaints, but the data on Table 4-1 do show passenger concern about reservation services, and other services -- e.g., baggage and on-board services -- as important passenger concerns.

It is fair to say that on-time performance and temperature control are also two important concerns of VIA Rail passengers. Indeed, the Railway Transport Committee, under Section 81 of the National Transportation Act, commissioned a study of VIA Rail on-time performance because of numerous complaints of schedule delays during the 1983-1984 Christmas holidays (see Mozersky, et al., 1984). The report concludes that many of the delays were caused by steam-heating equipment failures. Complaints by passengers suggest that, under these circumstances, poor on-time performance and passenger comfort are directly correlated.

There are two aspects of on-time performance that may be considered important to rail passengers. The first is a simple measure of reliability which considers only whether or not the train arrives at its destination on time. Simply being late may cause the passenger to miss a connection for the next leg of the trip, or to be late for an appointment. A second measure of on-time performance is measured as the length of the delay or degree of lateness. And, as has already been noted, on-time performance also affects passenger comfort. In brief, both being late, and the degree of lateness, may be important to the rail passenger.**

---

* Amtrak officials claim that temperature control has become significantly less of a problem with the installation of new equipment. Unfortunately, the ICC reports upon which Table 4-1 was constructed are not available after fiscal year 1979.

** "Predictability", measured as the variation in tardiness, may also be important to some passengers.

There are numerous dimensions to passenger comfort. As Table 4-1 indicates, temperature control is one that is regarded as important. Car cleanliness, equipment operability, food service quality, and other on-board services are other dimensions of passenger comfort. Although there is no direct evidence that has been identified specifically for the rail passenger industry, the evidence cited earlier in this report for the air line industry suggests that load factor may be another important aspect of passenger comfort. Specifically, it seems reasonable to presume that passengers prefer to ride on uncrowded trains as opposed to those with a high "load factor". The irony of this, of course, is that lower demand for passenger services increases the willingness to pay by remaining passengers.

There are also two important aspects of schedule quality. The first of these is related to trip speed; passengers, of course, prefer shorter trips to longer trips. The second important aspect of schedule quality is trip frequency. Greater trip frequency increases the likelihood that the rail passenger service will satisfy the "appointment" schedule of the passenger.

4.2 Methods for Evaluating Service Quality Contractual Arrangements

In the subsequent three parts of this section, a variety of contractual arrangements for providing passenger service quality are described and evaluated. The most important and general of the evaluation criteria are the following:

- Incentives -- Does the contractual arrangement provide adequate incentives for service suppliers to allocate resources to the production of the appropriate level of service quality?

- Enforceability -- Can the service quality agreement be enforced, and can the specific aspect of service quality be appropriately measured?

- Effectiveness - Does the contractual arrangement actually provide improvements in service quality?

Unfortunately, the most difficult evaluation criteria to assess is the last of the three listed above. This is due to both data problems and a lack of previous studies which have attempted to evaluate incentive effectiveness.

4.3 On-time Performance

Based on our own research, incentives for on-time performance have been far more prevalent than for those of any other quality aspect of passenger service. There are probably two reasons for this phenomenon. First, reliability, as reflected by some measure of on-time performance, appears to be very important to passengers; the evidence cited earlier in this section appears to confirm this conclusion. Second, convenient objective measures of on-time performance can be constructed relatively easily. This latter characteristic is critical for both parties in negotiating and enforcing a performance incentive agreement.

VIA Rail has paid on-time peformance bonuses to its carriers since its inception. In the U.S., Amtrak, which was formed in 1971, did not begin to experiment with on-time performance incentives until 1974. Indeed, as the

figures in Table 4-2 indicate, payments by both rail passenger corporations for on-time performance have been fairly substantial over the years.*

4.3.1 VIA Rail On-Time Performance Incentive Payments —

As was noted earlier, VIA Rail has negotiated on-time performance incentives into its agreements with rail passenger carriers since its inception. Recently, VIA Rail negotiated a two year contract with the Canadian National Railways (CN), and intends to negotiate a similar agreement with the Canadian Pacific (CP).

The newly negotiated contract with the CN again contains provisions for on-time performance. The structure of these incentives is illustrated in Figure 4-1.**

4.3.1.1 The Bonus Pool -- The CN contract provides for a maximum potential bonus pool for on-time performance. This pool is computed as ten percent of the projected annual on-train passenger revenues for all CN trains. The pool defines the on-time incentives that the CN can earn if all of its trains meet their schedules at least 90 percent of the time for each month during the year.

---

* Of course, this is not to say that the incentive payments are large enough. Mozersky, et al. (1984), for example, argue that VIA Rail payments for on-time performance are inadequate. This topic is discussed later in Section 5 of this report.

** VIA Rail regards its contracts as proprietary. The discussion below is based on descriptions provided by VIA Rail officials.

Table 4-2

VIA RAIL AND AMTRAK ON-TIME INCENTIVE PAYMENTS
($ millions)

Year	VIA Rail[a]	Amtrak[b]
1974	-	$ 5.8
1975	-	17.7
1976	-	n/a
1977	-	10.4
1978	$ 0.4	6.8
1979	7.7	n/a
1980	8.9	10.9
1981	10.6	18.8
1982	9.6	20.9
1983	12.2	21.7

[a] Canadian dollars

[b] U.S. dollars

Source: VIA Rail; U.S. Interstate Commerce Commission, Amtrak.

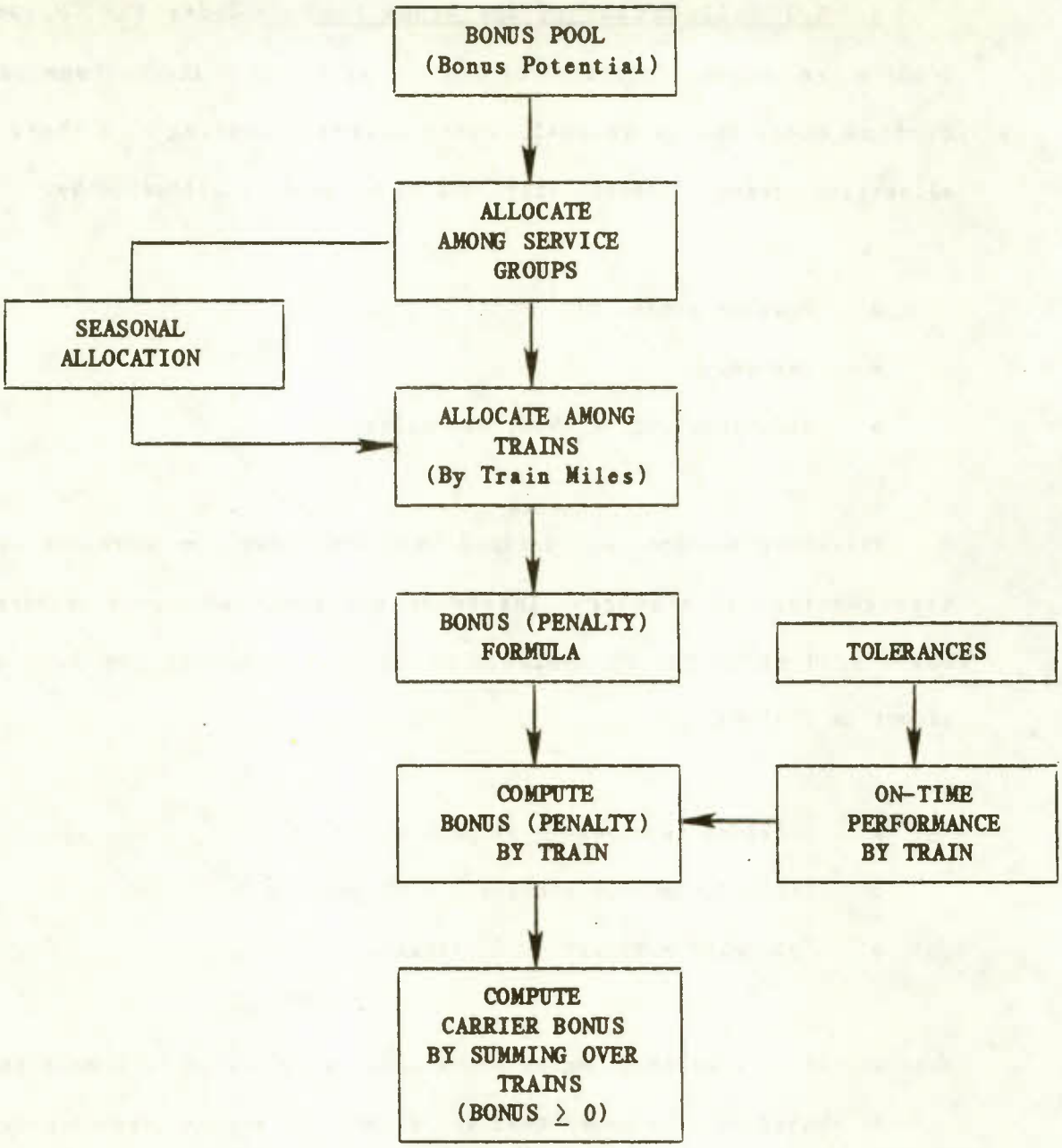


Figure 4-1. VIA Rail On-Time Performance Incentive Structure

4.3.1.2 Allocation of the Bonus Pool -- Under the CN contract, bonuses are computed monthly for individual trains. The maximum potential on-time bonus for individual trains varies according to a three-tiered allocation scheme. Specifically, the bonus pool is allocated by:

- Service group.
- Season.
- Individual train (by train-miles).

Passenger services are defined into four groups -- corridor services, transcontinental services, intercity services, and remote services. The bonus pool under the CN contract is allocated across the four service groups as follows:

- Corridor services -- 83 percent.
- Transcontinental services -- 12 percent.
- Intercity services -- 5 percent.

None of the on-time performance bonus pool is allocated to remote services.

It should also be noted that the allocation across service groups will be different for the CP since the service mix it provides differs from that of the CN. For example, a higher proportion of CP services are transcontinental, and a smaller proportion are offered in the corridor.

The bonus pool is also allocated over seasons. This adjustment is made to account for the fact that ridership, and hence, the total benefits associated with on-time performance, varies over different seasons of the

year. The seasonal weights, which are computed by month, are based on passenger loads for the previous three to four years.

Finally, the bonus pool for a given service and season is allocated across individual trains. This final allocation is done by train mile.

4.3.1.3 The Bonus (Penalty) Formula -- The discussion above describes how the maximum monthly on-time performance bonus for an individual train is computed. The actual bonus (or penalty) earned depends on the actual on-time performance of the train during the month. Bonuses are earned for on-time performance above 85 percent; penalties are assessed for performance below 85 percent.

Specifically, 20 percent of the maximum monthly bonus is earned for each percentage point that on-time performance exceeds the 85 percent threshhold level (i.e., the full bonus is earned if on-time performance equals or exceeds 90 percent). On the other hand, the carrier is assessed a penalty of 10 percent of the maximum bonus for each percentage point that on-time performance falls below the 85 percent level, down to 75 percent on-time performance (i.e., the maximum penalty, which equals the maximum bonus, is assessed for on-time performance of 75 percent or worse). A hypothetical bonus (penalty) schedule for a train with a maximum monthly bonus of $10,000 is presented in Table 4-3.

The 75-85-90 percent threshholds described above have just been negotiated into the new two-year contract agreed to by VIA Rail and the CN. These same threshholds will apply to all trains without variation. This represents a fairly significant departure from past VIA Rail contracts in which threshholds varied, sometimes substantially, across individual

Table 4-3

VIA RAIL ON-TIME PERFORMANCE BONUS (PENALTIES)
FOR A HYPOTHETICAL TRAIN
($10,000 Maximum Monthly Bonus)

On-Time Performance	Monthly Bonus (Penalty)
75% or Worse	($10,000)
76%	($9,000)
77%	($8,000)
78%	($7,000)
79%	($6,000)
80%	($5,000)
81%	($4,000)
82%	($3,000)
83%	($2,000)
84%	($1,000)
85%	0
86%	$2,000
87%	$4,000
88%	$6,000
89%	$8,000
90% or Better	$10,000

trains. It is our understanding that this is the only substantive change between the new and previous VIA Rail contracts.

4.3.1.4 Schedule Tolerances -- On-time performance for each train is computed at the final destination of a run. A train is considered "on-time" if it arrives at its final destination within five minutes of its scheduled arrival time. Partial credit for being on-time is credited, however, if the train arrives up to thirty minutes late. The on-time tolerance-credit schedule is reported in Table 4-4.* No credit is earned if the delay exceeds 30 minutes. The schedule displayed in Table 4-4 applies to all trains, regardless of trip length.

4.3.1.5 Computing the Total Monthly Bonus -- The total monthly bonus for a carrier -- i.e., the CN or CP -- is computed by summing the monthly bonuses of all individual trains, and then subtracting the sum of assessed penalties. The net monthly bonus for a carrier, however, must be nonnegative; that is, no penalties are assessed if the monthly sum of penalties exceeds the sum of bonuses.

4.3.1.6 Recouped Time -- VIA Rail contracts with passenger carriers also provide for credits for recovered time. Recall that on-time performance is based on the arrival time at the final destination of a trip. If, however, a train recovers delay time on the return leg of a trip, an 80 percent credit for the first leg of the trip is awarded. Sixty percent

---

* Partial credits apply to the computation of the on-time performance percentage figure upon which bonuses and penalties are based.

Table 4-4

VIA RAIL SCHEDULE TOLERANCES

Schedule Delay (Minutes)	On-Time Credit (Percent)
0 to 5	100
5 to 15	70
15 to 30	50
> 30	0

Source: VIA Rail

credits are earned if delayed time is recovered in subsequent round-trips during the day.

4.3.1.7 Comments on VIA Rail Incentives -- Many of the provisions of VIA Rail on-time performance incentives are similar to those specified in Amtrak contracts. A more complete evaluation of these and other contracts is provided later in this section of the report.

At this point, however, two comments are noteworthy. First, because of the 90 and 75 percent threshholds, passenger carriers have little incentive to improve on-time performance above 90 percent, and to prevent performance degradation below 75 percent. Second, because on-time performance is computed at the final destination, incentives to be on-time at intermediate points are less strong than they might be otherwise. Carriers might be late during the middle of a trip, but may be able to recover some lost time before the train arrives at its final destination.

4.3.2 Amtrak On-Time Performance Incentive Payments —

Amtrak's experience with on-time performance incentives is longer and more varied than that of VIA Rail. These are described below in a fair amount of detail. There are several potentially important lessons in this experience for VIA Rail. Many of these lessons are negative in nature; that is, they identify mistakes or pitfalls that should be avoided in the future if VIA Rail considers a restructuring of its own on-time performance incentive payments.

Since 1971, Amtrak has negotiated four different types of contracts with its carriers. In each of these, the structure of on-time performance incentive payments has varied, sometimes considerably. These four types of contracts are referred to as:

- The Original Agreements.
- First Amendment Agreements.
- Second Amendment Agreements.
- Third Amendment Agreements.

Throughout the history of these contracts, Amtrak has been subjected to a fair amount of criticism. It is fair to say, however, that at least with respect to on-time performance incentive payments, the third Amendment Agreements represent substantial improvements over the earlier agreements.

4.3.2.1 Amtrak's Original Agreements -- In 1970, Amtrak signed what are generally referred to as the "Original Agreements" with twenty railroads to provide passenger service. Like VIA Rail, Amtrak inherited a national rail passenger service in considerable disarray. This was due, in large part, to the fact that private rail carriers had little incentive to upgrade passenger service upon which they were incurring substantial losses.

Unfortunately, the original agreements signed by Amtrak and twenty passenger carriers provided little incentive for upgrading passenger service. Under the original agreement, railroads had agreed to provide the following services:

- Services and personnel to operate Amtrak trains (including train and engine crews).
- Station personnel.
- Equipment maintenance and cleaning.

The agreements granted Amtrak the right to operate its passengers over the carriers' lines. Amtrak also had the option to purchase or lease passenger cars from the railroads. In return for the services provided by the carriers, Amtrak agreed to reimburse the railroads for "solely related and/or avoidable costs" plus 5 percent of these costs to cover both unidentifiable costs associated with passenger service and a share of system-wide common costs.

The original agreements, however, provided no incentives for service quality whatsoever. In addition, many of the carriers argued that the 5 percent markup on avoidable costs was not sufficient to cover the unidentifiable costs associated with passenger service. To the extent that this was true, they had incentives to allow passenger service to deteriorate even further. By the accounts of most observers, this is precisely what happened.

4.3.2.2 First Amendment Agreements -- By June of 1974, Amtrak had signed the First Amendment Agreements with ten of the twenty rail passenger carriers. These First Amendment Agreements contained numerous incentives for enhancing the quality of passenger service, but the discussion

immediately below focuses on those aspects of the agreements that were directed to on-time performance.*

Several aspects of the First Amendment Agreements focus either directly or indirectly on the reliability of rail passenger service. These features of the First Amendment Agreements include:

- Schedule adherence.
- Excessive delay.
- Recovered time (i.e., time made up on a trip if a train was delivered late from another carrier).
- Equipment operability.

Under the terms of the First Amendment Agreements, schedule adherence was defined as the percent of on-time arrivals at the final destination for all trains operated by the carrier. Incentive payments were based on monthly performance, and were determined by fixed payments based on the number of percentage points that on-time performance exceeded some baseline standard. Payments (per percentage point above the baseline standard) ranged from $200 to $35,000 per month, depending on the carrier. The baseline standard was set at 65 percent in most cases, and in a few cases as high as 75 percent. Thus, if the baseline performance was established at 65 percent, and the incentive payment (per percentage point above the baseline) was $20,000, a carrier with an on-time performance of 75 percent

---

* For detailed descriptions of First Amendment Agreements between Amtrak and the 10 rail carriers, see Baumol (1975) and GAO (1977).

would earn $200,000 per month in performance incentive bonuses for schedule adherence.

The First Amendment Agreements also define tolerances within which arrivals would be considered on-time or late. These tolerances, which varied with trip length, are reported in Table 4-5. As this table indicates, the tolerance varied from 5 minutes for trips up to 150 miles to 30 minutes for trips exceeding 550 miles.

Passenger carriers signing First Amendment Agreements were also penalized for "excessive delays", and were rewarded for making up time when trains were delivered late from other carriers. Excessive delays were defined by trip length which varied from lateness of 15 minutes or greater for trips up to 250 miles, and high as 60 minutes for routes exceeding 450 miles. Penalties were assessed based on a flat rate (per minute late) and varied considerably across different routes. First Amendment Agreements also specified the maximum number of minutes that carriers could be penalized for excessive delays. These maximums varied between 60 minutes and 180 minutes, again depending on the length of the route. In most cases, penalties for excessive delay were relatively small; for example, one carrier's penalty was $2.50 per minute of excessive delay.

Bonuses for recovered time were structured along lines similar to the penalties for excess delays. These bonuses were based on a flat rate of (per minute) time recovered. Like the penalties for excess delay, incentives for recovered time were, for the most part, relatively small.

It is also noteworthy that penalties could only be assessed against previously earned bonus credits; that is, no single rail carrier could receive payment less than its avoidable costs (plus a 5 percent management

Table 4-5

SCHEDULE TOLERANCES IN FIRST AMENDMENT AGREEMENTS

Trip length in miles	Tolerances in minutes
0 to 150	5
151 to 250	10
251 to 350	15
351 to 450	20
451 to 550	25
551 or more	30

Source: GAO (1977).

fee). Incidentally, under the First Amendment Agreements, avoidable costs were redefined to include some carrier costs that were not included in the original agreements. In other words, the definition of avoidable costs was made more liberal.

Rail carriers were also paid incentives for keeping equipment "operable" above a baseline level which was defined as a percent of the passenger cars assigned to them. These payments were intended to improve both schedule adherence and passenger comfort. Passengers cars were defined as inoperable if the following occurred:

- The car had to be cut from the train and another car substituted.
- The car caused the train to reach its destination late.
- The train caused passenger discomfort because of either low or high temperature, or inadequate lighting.

Bonuses of $150 for each car above the baseline that operated for a month without failure were paid to carriers signing First Amendment Agreements. For example, if a rail carrier had been assigned 1,000 passenger cars, and all operated without failure for a single month, the rail carrier would be paid a bonus of $7,500 for that month for the 50 cars exceeding, say, a 95 percent baseline. In general, bonus payments for equipment availability were relatively small under the First Amendment Agreements, at least in comparison to the payments that were received for schedule adherence.

Despite Amtrak's good intentions in signing First Amendment Agreements, several observers criticized some of the provisions that were nego-

tiated in an effort to improve on-time performance. Some of these criticizms were valid; others were less valid.

Two concessions made by Amtrak in negotiating First Amendment Agreements were criticized in ICC (1976) and GAO (1977) reports. These concessions were:

- A liberalization of the definition of on-time.
- Setting the baseline on-time performance level below that achieved by passenger carriers prior to the First Amendment Agreement.

Under the First Amendment Agreement, Amtrak agreed to more liberal definitions of on-time than had been previously accepted. Prior to the First Amendment Agreements, a train was defined as being late if it arrived at its final destination more than 5 minutes after the scheduled arrival time. As was noted earlier in Table 4-5, trains could arrive as much as 30 minutes after the scheduled arrival time and still be considered on-time if the trip exceeded 550 miles in length under First Amendment Agreements. It should be recognized, however, that because payments to passenger carriers under original agreements had no relationship with on-time performance, the prior definition of on-time performance is a moot issue.

The important point here is that the First Amendment Agreement provided no incentives for carriers to reduce delays between the interval before which they were considered on-time and after which the delay was considered excessive. For example, a train with a trip length of 550 miles or greater would be considered on time if it arrived at its final destination by as much as 30 minutes late, but the delay would not be considered

excessive unless it was 60 minutes late. Thus, the carrier had no incentive to improve performance within a 30-60 minute delay. In addition, since the carrier receives the incentive credit even if it arrives 29 minutes late, it has little incentive to reduce the delay within the tolerance. Accordingly, the main problem with this aspect of the First Amendment Agreement was not the liberalization of on-time definitions per se, but rather, the "lumpiness" in the performance incentive structure.

Amtrak was also criticized for setting the baseline on-time performance level (i.e., from 65 to 75 percent) below performance levels that had been achieved by carriers prior to the First Amendment Agreements. Some of this criticism, however, has been somewhat misdirected. The First Amendment Agreements provided no direct penalties for degredation of on-time performance. This being the case, it was appropriate to set the baseline below performance standards achieved previously. In effect, this arrangement did provide for penalties associated with performance degredation; namely, the opportunity cost of losing incentive payments.

There were a number of other criticisms that were directed to Amtrak's First Amendment Agreements with rail passenger carriers. These include:

- On-time percentages were computed by carrier, instead of by individual train; moreover, schedule adherence incentives were based only on arrival times at the final destination and not at intermediate points.
- The incentive structure for schedule adherence conflicted with the incentive provisions in First Amendment Agreements intended to improve schedule quality.
- Bonuses for made-up time included payments for arriving at the final destination ahead of schedule.
- The First Amendment Agreements should have included additional penalties for excessive delays.

It should be noted that some of the deficiencies in the First Amendment Agreement noted above were corrected in later contracts negotiated between Amtrak and passenger carriers.

Under the First Amendment Agreements, on-time percentages were computed by carrier instead of by individual train. Thus, delays on long and short trips were given equal weight and the carrier could offset poor performance on one route with good performance another route. In addition, on-time performance at intermediate stops along the route had no effect on the schedule adherence bonus. GAO (1977) notes that a high percentage of Amtrak passengers disembark at intermediate stops, and that such passengers were frequently inconvenienced by lateness, even though the train was eventually able to make up the time and arrive at the final destination on schedule.

Even though the First Amendment Agreements provided incentives for carriers to improve schedule quality, (i.e., shortened the duration of the trip) the incentives were so small relative to those provided for schedule adherence that a great many carriers signing First Amendment Agreements were successful in negotiating relaxed schedule times. For example, GAO (1977) notes that one carrier would have received a one-time payment of $10.51 for a one hour reduction in running time on one particular route. However, this same carrier, at the time, was receiving monthly on-time performance payments in excess of $150,000. Clearly then, rail carriers had substantial incentives to increase running times on routes, thus reducing the likelihood of late arrivals.

As was noted earlier, the First Amendment Agreements also provided incentive bonuses for carriers that made up lost time on trains that were

delivered late by another carrier. Apparently, however, a loophole in the provision permitted trains to bill Amtrak for arriving at the destination ahead of time; that is, some carriers were able to not only make up lost time, but arrive at the terminal ahead of schedule, thus billing Amtrak for the entire reduction in travel time. Since rail passengers receive few (if any) benefits from arriving at the destination ahead of schedule, the bonus for made up lost time was eliminated from the subsequent Amtrak agreements.

Baumol (1975) notes that the penalties for excessive delays were probably inadequate in First Amendment Agreements. He argues that passenger inconvenience probably increases as a function of delay time. As a result, Baumol suggests that the penalty should be structured as an increasing function (e.g., quadratic or exponential) of the number of minutes that the excess of delay lasts.

4.3.2.3 Second Amendment Agreements -- By 1977, Amtrak had signed Second Amendment Agreements with several passenger carriers. These agreements changed several aspects of the schedule adherence incentive payment structure. The most significant of these changes are:

- On-time performance was computed by individual train, instead of as the average on-time performance for the entire carrier.
- The baseline on-time performance level was set at 80 percent in most cases, and penalties were assessed for on-time performances below 70 percent.
- Bonuses for made-up lost time were eliminated.
- Bonuses for equipment operability were eliminated.

The Second Amendment Agreements provided that on-time performance incentive payments be calculated by individual train. This represented an improvement over the First Amendment Agreements, since it provided the carrier's with a direct incentive to improve performance on individual routes. Recall from the earlier discussion provided in Section 3 of this report that the socially optimum level of service quality is achieved at the point where the average valuation of quality improvements equals the marginal cost of such improvements. Since conditions vary sometime substantially across different routes (e.g., because of track conditions and freight traffic density) it is reasonable to expect that the marginal cost of improving schedule adherence likewise varies across routes. Only by providing direct incentives on an individual train basis is it possible to achieve the socially optimum on-time performance level, given that costs on the margin also vary.

Based on criticism of the First Amendment Agreements, Amtrak negotiated upward the baseline on-time performance level to 80 percent. That is to say, bonuses were paid by month for each percentage point of on-time performance exceeding the 80 percent baseline. In addition, Second Amendment Agreements provided for penalties on performance below 70 percent. However, the Agreements also stipulated that penalties could be assessed only against bonus credits earned in the preceding 12 months.

Some observers (GAO, 1981) applauded this change in the Second Amendment Agreements in that it required a greater level of achievement by carriers in order to earn any incentive payments. It is not clear, however, that this aspect of the Second Amendment Agreements represented an improvement. The basic problem is that carriers had no incentives whatso-

ever to improve on-time performance if they could not expect to exceed the 80 percent threshhold, but had no fear of falling below the 70 percent threshhold, below which the penalties were assessed. This is a generic problem with "lumpy" incentive payment structures. Note that under the First Amendment Agreements, carriers faced "continuous" incentive structure for on-time performance, unless they were unable to achieve at least a 65 percent on-time performance rate.

The Second Amendment Agreements also eliminated bonuses for equipment operability. In their place, Amtrak instituted a preventive maintenance program. The demise of the incentives for equipment operability is probably attributable to two factors: 1) difficulty in policing the First Amendment Agreements; and 2) Amtrak had taken over many of the maintenance services that were previously provided by the carriers.*

4.3.2.4 Third Amendment Agreements -- Recently, Amtrak has successfully negotiated Third Amendment Agreements with several passenger carriers. With respect to on-time performance incentive payments, Third Amendment Agreements carry over most of the changes that occurred in the Second Amendment Agreements. For example, on-time performance is still measured by individual trains, and bonuses and penalties are assessed for performances above and below the 80 and 70 percent baselines. As was the case with Second Amendment Agreements, penalties can only be assessed against previously earned bonus credits earned in the preceding 12 months.

---

* GAO (1977) documents the difficulty that Amtrak had in identifying "inoperable" cars. It seems that Amtrak consistently overpaid carriers for equipment operability because of its inability to police this aspect of First Amendment Agreements.

There is, however, one noteable change in the Third Amendment Agreements. Specifically, Amtrak has successfully negotiated into these agreements the requirement that on-time performance be measured at all intermediate points along a given route. This change represents an improvement over the previous Amtrak agreements in that it finally provides passenger carriers with increased incentives for providing service quality to passengers that disembark at intermediate points along the route.

4.3.3 On-Time Performance Incentives in Other Transport Industries —

Contractual arrangements related to on-time performance in both the airline and bus industries were explored to determine if any lessons applicable to the VIA Rail case study could be learned. The investigation focused on the relationship commuter lines feeding major airlines in the airline industry. In the bus industry, the investigation was focused on contractual arrangements between private bus lines and local mass transit authorities (MTA's).

We were unable to identify any contractual relationships in the airline industry in which commuter airlines, feeding major hubs, were paid incentives for on-time performance. The general rule in the industry appears to be quality by fiat. Commuter lines and major airlines often sign into agreements whereby commuter lines design their schedules to fit those of the major carriers. Generally, in return, the commuter lines acquire certain advertising privileges, and in one case at least, an extension of credit. Industry spokesmen have explained to us that on-time performance is generally not the problem in the industry because it is in

the vested interest of the commuter lines to provide service quality, since they retain revenues from their own ticket sales. Translated, of course, this means that the profit maximizing levels of quality for the commuter lines and the major carriers are not sufficiently different to present significant conflicts.

In addition, industry spokesmen told us that the contracts engaged between commuter lines and major carriers typically allow the major carrier to terminate the agreement if performance is unsatisfactory. Indeed, the single copy of a contract between a commuter and a major carrier that we were able to obtain confirms this. In fact, the contract stipulated that either party could terminate the relationship if the arrangement proved unsatisfactory.

Our casual survey of spokesmen in the bus industry can hardly be termed a random sample. Nonetheless, it appears that contractual arrangements stipulating on-time performance between private bus lines and MTA's are fairly common, although relatively new.* Because of obvious policing difficulties, on-time performance is not measured at intermediate points. Instead, on-time performance appears to be measured at exchange points where buses drop off passengers for connections with other transport modes (e.g., train or subway service). Spokesmen have explained to us that this is what really matters in commuter service; that is, passenger inconvenience arises when the bus arrives late at a railroad or subway terminal and causes the passenger to miss the connection.

---

* Competitive tendering is also employed by MTA's to assure schedule adherence. Specifically, incumbents providing poor performance face the risk of losing contracts.

One MTA spokesman described a rather unique penalty system for poor on-time performance. Specifically, on-time performance for private bus lines feeding the MTA's rail system is measured at rail terminal drop-off points. If the bus is late such that the passengers miss their rail connection, the prorated payment for that bus trip is deducted from the bus line's service payment. In short, if the bus is late, the service was not delivered, and no payment is made.

Large mass transit authorities also sometimes contract with privately owned railroads to deliver passengers to their transit system. Generally, the on-time performance arrangements here seem to mirror those described earlier for VIA Rail and Amtrak. Specifically, bonus payments are paid for on-time performance exceeding a given baseline level, and penalty points are assessed against performances below a minimum performance level. It is interesting to note that, like VIA Rail and Amtrak, it appears that penalty assessments are generally deducted from previously earned performance credits, but never against the fixed payments made to cover the rail carrier's avoidable costs.

4.3.4 Evaluation of On-Time Performance Incentive Contracts —

At this point, it is instructive to evaluate the various characteristics of the on-time performance incentive contracts that have been reviewed above. As was stated previously, these arrangements are evaluated in terms of: 1) incentives for providing service quality; 2) the degree to which the provisions are enforceable; and 3) the effectiveness of the contracts in improving performance. It should be stressed at the outset, however,

that the contracts described earlier in this section have typically resulted from difficult negotiating sessions. Accordingly, any of the improvements suggested here may be difficult to implement in view of the complex strategic games that are involved in contractual negotiations.*

4.3.4.1 Inducements for On-Time Performance -- Perhaps the most important conclusion here is that, in principle, contractually negotiated performance incentives are an appropriate means for dealing with the reliability aspect of rail passenger service quality. Several valid measures of service reliability can be easily constructed, and, at least relative to other aspects of service quality, easily enforced. In theory, so long as the incentive schedule is set at an appropriate level, rail passenger carriers will respond by providing the desired level of on-time performance.

This is not to say, of course, that the incentive payments described earlier have been set at appropriate levels in the optimal sense. Presumably, negotiators attempt to set the on-time performance payment schedules based on their own view of the benefits of improved service quality, and the incremental costs of achieving such improvements. The degree to which these incentive structures approximate those that are optimal, however, will depend on the quality of information available to the negotiator, and the appropriateness of the negotiator's objectives.

---

* The discussion that follows is abstracted from institutional settings within which contracts are negotiated and enforced. Institutional constraints on VIA Rail have considerable impacts on the effectiveness of its contracts. These are discussed in detail in Section 5.

The contractual arrangements reviewed earlier in this section are also appropriate in two other respects. First, with the exception of some schedule "lumpiness" problems, the incentive structures take the appropriate form. Specifically, they generally represent a fixed price per unit of quality provided by the rail carriers. So long as these prices per unit of quality are set at appropriate levels, rail carriers will have incentives to provide on-time performance up to the point at which the marginal costs of improvement are equal to this price. Second, the incentive payments are substantial enough so that carriers are likely to be able to earn profits in producing quality.* As was noted earlier in Section 3, any profits that are deemed excessive should be "taxed" away in the form of a lump sum deduction from the fixed portion of the negotiated contract.

There are, however, some suggestions that are appropriate for this case study. These include:

- Lumpiness in incentive schedules should be avoided. This includes different threshholds for computing bonuses and penalties, as well as tolerance intervals. Carriers caught in between these performance intervals have no incentives to improve performance. We suspect that the on-time tolerance interval is the lesser of the two problems.

- On-time performance should be measured by individual train, and it may be appropriate to vary incentive payments across trains as well. This allows the carrier to account for differences in the marginal cost of improving on-time performance over various routes. Performance payments should be higher on routes with high load factors to account for the higher social benefits of improved service quality.

- On-time performance should be measured at intermediate stops in order to provide incentives for carriers to accommodate passengers who disembark before the train reaches its final destination.

---

* Again, this does not imply that the size of the incentive payments are optimal.

4.3.4.2 The Enforceability of On-Time Performance Incentives -- Fortunately, the fact that on-time performance is relatively easy to measure allows for relatively easy enforcement of the contractual agreement. This, of course, is critical if the incentive structure is to be effective. Because of the large amount of payments that are involved, it is important that an accurate accounting of the actual performance of the rail carrier can be made.

This is not to say, however, that there are no problems in enforcing the terms and conditions of the incentive agreement. The basic problem is that policing the agreement is not costless to the contractor. Amtrak, for example, attempted to minimize its own enforcement costs during its experience with First Amendment Agreements. Apparently, the incentive payments themselves were based on statistics supplied by the carriers. Of course, the contractee is placed in a situation of moral hazard, since the incentive payments themselves depend on recorded performance.

GAO (1977) conducted a study on the reliability of on-time performance data submitted by Amtrak's rail carriers. In brief, they claim to have found substantial discrepancies -- i.e., the rail carrier's own figures for on-time performance substantially overstated actual performance -- when compared with some of their own figures, and Amtrak internal audits.

Nonetheless, given the relatively simple objective measures of on-time performance that are available, the question is not so much one of disputing the quality of the service, so much as it is finding ways of reducing the contractor's cost of policing the agreement. While some check on the rail carrier's own performance statements seems necessary, it is not necessary for the contractor to monitor all stops. Standard sampling

procedures for audits are available. In other words, the contractor can mitigate enforcement costs by conducting appropriately structured sample audits of carriers' on-time performances, comparing these against the rail carriers' own submitted records, and adjusting the incentive payments accordingly. It should be feasible to negotiate such an arrangement into contracts, because the savings in monitoring costs can be shared by both the contractor and the contractee.

One MTA representative interviewed by us described a relatively unique approach to the enforceability problem. Specifically, the MTA relied for the most part, on on-time performance records submitted by the contractee. However, if the contractee failed to report a delay that was discovered either through a customer complaint, or through the MTA's own investigation or audit, the penalty for the delay was doubled. Thus, the MTA was able to negotiate a contract which internalized to the contractee some of the costs of enforcement.

4.3.4.3 The Effectiveness of On-Time Performance Incentives -- Ultimately, the questions regarding on-time performance incentives center on their effectiveness. The question is: Exactly how much service quality is the contractor purchasing with the incentive payments? Another way of stating this is: What is the marginal cost of an improvement in on-time performance?

Unfortunately, there is precious little information upon which to answer this question. We have been unable to identify any systematic study that has attempted to measure the elasticity of on-time performance with respect to incentive payments.

Simple comparisons between incentive payments and on-time performances -- either using time series or cross sectional data -- are likely to be inconclusive. There are several complicating factors here that must be accounted for, some of which have been noted earlier in this report. These factors include:

- Roadway conditions.
- Equipment conditions.
- Traffic density.
- Weather conditions.

These caveats notwithstanding, Table 4-6 does provide some interesting statistics on the effects of incentives on on-time performance for Amtrak passenger carriers. The first column in this table shows overall Amtrak on-time performance statistics -- i.e., on a system-wide basis -- for fiscal years 1973 through 1983. The second column reports the on-time performance for carriers which have signed on-time performance incentive contracts. These figures indicate that those rail carriers under incentive agreements have consistently out-performed those carriers that have refused to sign any of the amended agreements that carry incentives for performance.*

In addition to the caveats regarding simple comparisons made above, it should be noted that the simple comparisons provided in Table 4-4 suffers in two other regards. These are:

---

* The figures displayed in Table 4-4 are carrier-wide averages, and are not based on the same statistics that Amtrak employs to compute incentive bonuses.

Table 4-6

AMTRAK ON-TIME PERFORMANCE

Fiscal Year	On-Time Performance (%)	
	Overall	Carriers Under Incentive Contracts
1973	60.2	61.2
1974	75.4	80.8
1975	77.4	90.3
1976	74.4	84.6
1977	62.0	72.2
1978	62.0	75.3
1979	57.0	68.7
1980	69.0	80.9
1981	77.0	88.0
1982	79.1	83.2
1983	81.5	86.6

Sources: Amtrak Annual Reports, U.S. Interstate Commerce Commission.

- Self-selection bias -- i.e., those carriers for which it is difficult to improve on-time performance are less likely to have signed any of the amended contracts. It is interesting to note, however, that in 1973, prior to any of the amended agreements, the system-wide figure of 60.2 percent is roughly comparable to the 61.2 percent obtained by carriers who subsequently sign incentive agreements.

- Many of the carriers included in the second column in Table 4-6 were successful in negotiating liberalized running times over routes prior to negotiating amendment agreements.

The last statement warrents further comment. The substantial increase in on-time performance among carriers signing incentive contracts in Table 4-6 coincides with Amtrak's successful negotiation of First Amendment agreements with 10 of its passenger carriers. However, many of these rail carriers also were able to negotiate increased running time over many of their important routes. In fact, both GAO (1977) and ICC (1976) conclude that much of the improvement in the on-time performance statistics is attributable to liberalized running times on schedules and on-time definitions. Although these studies do not account for the conditional effects of all of the complicating factors identified above, it would be difficult to dispute that their conclusions are reasonable.

We also asked many of the industry spokesmen that we interviewed for their views on the effectiveness of incentive payments on on-time performance. Their responses were somewhat mixed, but they generally agreed that the incentive payments were effective in that they did increase on-time performance. As can be expected, however, none would venture a guess as to how much of the improvement in performance was attributable to performance incentives, and other complicating factors.

It is also important to note that, with a given level of capital, rail carriers are constrained in their ability to improve reliability. Specifically, both the condition of roadways and rolling stock affect on-time performance. Many rail carriers that have refused to negotiate incentive contracts with Amtrak have claimed that they have refused to do so because they have little or no control over on-time performance. Rather, they argue, the condition of Amtrak equipment -- which Amtrak now maintains largely by itself -- is the primary cause of most schedule delays.

The extent to which this is true is unclear. However, it is safe to conclude that, given equipment condition, there is only so much that rail carriers can do to achieve improved on-time performance. In short, improving performance beyond some threshhold level may require investments in capital.

4.4 Passenger Comfort

As far we can determine, Amtrak has been the only passenger service to experiment with incentives for improving passenger comfort. In particular, several clauses in the First Amendment Agreements provided incentives (or penalties) for services related to passenger comfort. Unfortunately, the results of the experiment were so dismal that Amtrak eliminated these clauses in subsequent amendments to its agreements with rail carriers. Amtrak concluded that the expense of policing the agreements resulted in costs in excess of benefits.

4.4.1 Passenger Comfort and Amtrak's First Amendment Agreements —

The First Amendment Agreements had two provisions that focused on passenger comfort. These provisions were related to equipment operability and car cleanliness. As was noted earlier in this section, these agreements provided bonuses of $150 for each car that carriers maintain in operable condition above some baseline percentage of all cars assigned to them. The provisions in this clause directed to passenger comfort included temperature control and lighting adequacy. Cars were defined inoperable if the temperature fell outside the range of 62 to 82 degrees Fahrenheit for more than 1 hour during the trip (unless the outside temperature exceeded 96 degrees Fahrenheit). Car lighting was defined inadequate if it was too poor for passenger reading for more than one hour during the trip.

Passenger carriers were also assessed a $50 penalty for each car that was found to be unclean. Car cleanliness was defined by twenty-five different criteria including the following:

- Trash removed.
- Floors and seats vacuumed.
- Ashtrays emptied.
- Towels, sheets, and headrests changed.
- Wall surfaces washed.
- Sinks and toilets cleaned and disinfected.
- Windows cleaned.

Cars were inspected periodically by Amtrak by employees at the origin of the trip.

4.4.2 Evaluation of Passenger Comfort Incentives —

Many of the problems that Amtrak encountered in attempting to negotiate incentives for improved passenger comfort were the result of difficulties in policing and enforcing the agreements. Nonetheless, it is not clear that the incentive payments were sufficiently large to induce carriers to improve service quality in the first place.

Payments for equipment operability were never substantial under First Amendment Agreements. In addition, the $50 penalty for unclean cars was the purportedly established at the cost of cleaning itself. If this is true, it is not clear that rail carriers had any incentives under the agreements to promote car cleanliness.

In addition, it should be noted that the agreements failed to cover many important aspects of passenger comfort in any event.* For example, one passenger carrier refused to pay penalties for a broken window, and in another case, for a water cooler with no water, because these were not defined as failures in Amtrak contracts, even though they clearly affected passenger comfort. Examples of other failures potentially affecting passenger comfort, but not included in the First Amendment Agreements include inoperable doors and toilets, broken seats, and water leaks. These problems all point out the difficulty of specifying detailed conditions in

---

* Baumol (1975) warned of this problem at the time the First Amendment Agreements were being negotiated.

passenger cars that potentially affect passenger comfort, and then negotiating them into a contract.

Perhaps the greatest difficulty encountered by Amtrak in its experiment with passenger comfort incentives, however, dealt with enforcing the terms of the contract. By the end of June 1976, Amtrak had assessed penalties for only 479 unclean cars (GAO 1977). Difficulties in interpreting cleanliness standards, and the amount of manpower that was required for inspections are offered as the principle reasons for the difficulties. GAO (1977) for example, found 130 violations in 343 cars that it inspected; yet, Amtrak had assessed penalties for none of the violations (indeed, GAO claimed that its standards for cleanliness on its inspection were less strict than those specified in the agreements).

4.4.3 Passenger Comfort: Some Conclusions —

As the preceding discussion suggests, it appears that incentive payments directed to improving passenger comfort are likely to fail for two reasons: 1) it is difficult to define objective measures of many aspects of passenger comfort; and 2) even where objective measures are available, considerable costs may be borne by the contractor in enforcing the agreement, especially given that the incentive structure is based on a per unit violation.

We suggest that quality by fiat may be a viable alternative in contracts where passenger comfort is involved. Indeed, this appears to be the most common arrangement employed by mass transit authorities in their dealings with private bus lines. Buses are periodically inspected, and if

unsatisfactory conditions are discovered, violations are noted to the contractee, and the contractee is expected to respond accordingly. If the contractee's performance continues to be unsatisfactory, either litigation or termination of the contract can be exercised.

Given on our earlier comments, however, quality by fiat is expected to be effective only if competitive tendering is possible. It occurs to us that, absent institutional constraints, competitive tendering for the cleaning and light maintenance services necessary for passenger comfort can be obtained through competitive tendering. It is doubtful that significant scale economies apply to the production of these services, and, at least for some of these services, highly specialized human capital is not required. We suspect that the same conclusions may hold for other on-board services such as food and beverage preparation.

This is not to say, however, that institutional barriers here are not substantial. Many of these services are currently provided by VIA Rail (and Amtrak).* Private sector markets must be established to render competitive tendering feasible. Contracts with labor unions may also impose substantial institutional barriers.

Improving some aspects of passenger comfort may require investments in capital. Passenger comfort is likely to be affected by the age and condition of passenger cars. To the extent that passengers prefer low load factors, additional seating capacity is required. This is not to say, of course, that improvements in these services are warranted in terms of economic efficiency. The point here is: the ability to improve the

---

* How efficiently such services are currently provided is an empirical issue, but their provision is not subjected to the rigors of competitive markets.

quality of passenger comfort is constrained, at some point, by available capital.

4.5 Schedule Quality

As was noted earlier, there are two aspects of schedule quality that may be important to rail passengers. First, passengers are likely to prefer shorter travel times between any given origin and destination pair. The value to the passenger will equal his or her value of time, weighted by the reduced running time.

Second, passengers are likely to prefer greater trip frequencies. Greater trip frequency increases the likelihood that the trip schedule will match the appointment schedule of the passengers. In other words, the availability of a wider trip menu will decrease the inconvenience to the passenger of having to arrive early at the destination in order to arrive on-time for an appointment.

Typically, trip schedules are determined through negotiations between the contractor and the contractee. Other things being the same, reduced running time reduces the likelihood that the carrier will be able to achieve a high level of on-time performance. Thus, it is reasonable that negotiations over trip schedules and on-time performance incentives are conducted jointly. Indeed, Amtrak's contractual experiment on providing direct incentives for reduced travel times failed because of the inter-dependency between schedule quality and on-time performance.

4.5.1 Schedule Improvement Incentives Under Amtrak's First Amendment Agreements —

Amtrak attempted to induce rail passenger carriers to improve their schedules through reduced travel times in its First Amendment Agreements. More specifically, the schedule improvement bonus, B, was computed as:

$$B = N\left(\frac{T_1 - T_2}{T_1}\right) R$$

where, N = the number of annual trips for the train;

T_1 = scheduled trip time in minutes;

T_2 = the trip time in minutes for the improved schedule; and

R = bonus rate negotiated by contract.

4.5.2 Evaluation and Summary of Schedule Improvement Incentives —

Unfortunately, the incentives in the First Amendment Agreements for schedule improvements were so small relative to on-time incentives that most carriers opted to forego these bonuses, and instead, negotiated increased running times for many of their trains.* No doubt, the fact that the schedule improvements were one-time only incentives added to the carriers reluctance to take advantage of these bonuses.

In principle, however, schedule improvement is one aspect of the quality of rail passenger service that is especially amenable to direct

---

* Baumol (1975) warned of this flaw in the agreements.

incentive payments. Clearly, the performance standard is easily defined and enforced. Indeed, no additional policing of the agreement (other than that required anyway for schedule adherence) is required once the contract is signed.

The problem, of course, is that the incentives must be sufficient to provide some inducement to the contractee relative to on-time performance incentives. This also poses a difficult information problem for the negotiator, however. Specifically, the negotiator must weigh the passengers' preferences for reliability against travel time, and then against the relative marginal costs of obtaining the two service quality characteristics.

It is also apparent that schedule improvements, in the short-run, will be limited to capital condition and availability. Clearly, improvements in travel time are limited by the condition and capacity of both roadways and rolling stocks. Similarly, the ability to increase trip frequency is eventually limited by the capacity of the system.

5. RAIL PASSENGER SERVICE QUALITY IN CANADA: ISSUES, PROBLEMS, AND RECOMMENDATIONS

This section of the report is focused specifically on issues related to rail passenger service quality in Canada. First, the importance of considering service quality within a framework of overall VIA Rail objectives is discussed in Section 5.1; methods for evaluating alternative rail passenger quality levels are also described in this section. Several important institutional issues are discussed next in Section 5.2. The primary focus here is on how institutional constraints affect VIA Rail's ability to control service quality. The evaluation of existing incentives for Canadian rail passenger service quality is provided in Section 5.3. Finally, Section 5.4 describes methods for obtaining the data necessary for implementing a rational policy toward rail passenger service quality. Specific recommendations are offered, as they are appropriate, in each of these sections.

5.1 Service Quality, VIA Rail Objectives, and the Evaluation of Alternatives

Earlier, in Section 2.1 of this report, it was noted that policy options directed to promoting service quality among government enterprises must be established within a framework which considers the overall objectives of the enterprise. Here, we attempt to sort out several issues related to rail passenger service, and possible objectives for VIA Rail as either a broker, or a producer, of rail passenger service. The attempt here is not to suggest what these objectives should be, or what weights

should be attached to different subobjectives; rather, the discussion is intended to organize a given set of objectives within a systematic framework. In addition, the discussion also identifies those cases in which subsidies may be appropriate.

A method for evaluating the desirability of alternative levels of rail passenger service quality in Canada is also offered. In particular, we suggest that such evaluations be conducted within a framework of applied benefit-cost analysis. The suggested approach here is relatively modest in recognition that it may not be practical to identify the optimal level of service quality for rail passenger service. Instead, we suggest that a limited number of alternatives to the existing service quality level be identified and compared to a baseline performance level. Much of the information required for the applied benefit-cost analysis will be difficult to obtain in practice; nonetheless, we suggest that this approach is useful for thinking about alternatives to existing performance standards.

5.1.1 VIA Rail Objectives —

The schematic displayed in Figure 5-1 provides a useful method of organizing the impacts of service quality changes in terms of the objectives of the enterprise. The first box is labeled "VIA Rail objectives." The implied presumption here is that VIA Rail objectives are consistent with Canada's overall objectives for rail passenger service.

If VIA Rail is considered a pure enterprise, its primary objective is to maximize the net present value of the firm. In general, this objective will not likely be consistent with Canada's overall objectives for rail

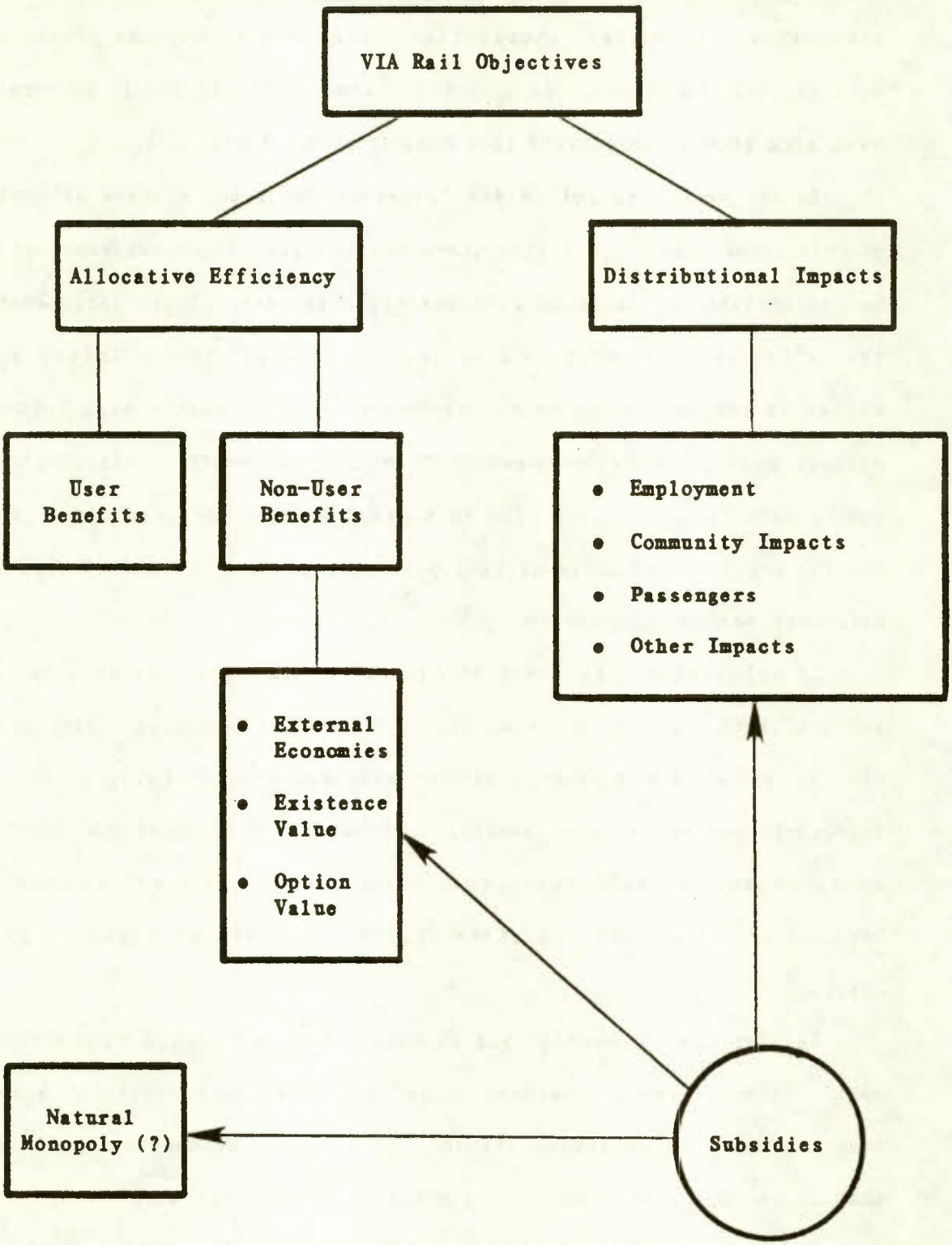


Figure 5-1. VIA Rail Objectives and Service Quality

passenger service. To the extent that VIA Rail is regarded as a true enterprise, appropriate institutional constraints must be placed on VIA Rail so that its behavior as a profit maximizer is limited to be consistent with more general objectives that society deems worthwhile.

In any event, we follow the framework developed earlier in Section 2 of this report and suggest that there are two generic objectives that might be appropriate for Canadian rail passenger service. These include allocative efficiency and welfare distribution. Recall that allocative efficiency is achieved when no one member of society can be made better-off without making some other member of society worse-off. Distributional or equity considerations give rise to another generic objective if it matters how the benefits and costs of rail passenger service are distributed across different members of society.

If allocative efficiency is considered as a goal for rail passenger service, both user and non-user benefits must be evaluated. User benefits are enjoyed directly by rail passengers, and are reflected by the willingness-to-pay for service quality. Absent distributional considerations, improvements in rail passenger service quality need not be subsidized because of direct user benefits, unless the natural monopoly problem exists.

The "Natural Monopoly" box in Figure 5-1 is flagged with a question mark. In order for the natural monopoly problem to be relevant here, the long-run costs of improving service quality must be decreasing. Although this is an empirical issue, it is not likely to be the case.

Non-user benefits related to allocative efficiency include those created by external economies, existence value, and option value. An

external economy is created when some member of society other than the passenger receives a benefit from the consumption and production of rail passenger services. Possible reduced community congestion, improved safety and pollution are examples of external economies. It is important to note here, however, that the issue is whether improvements in rail passenger service quality (and not output) generate external economies. It is quite likely that the external economies generated by changes in output are substantially different from those generated by quality improvements.

Existence value is present when some members of society receive benefits from rail passenger service apart from its use or potential use. As was noted earlier in Section 2, existence value is often used to describe the benefits of maintaining the viability of rare species threatened by environmental conditions. It is not implausible to suspect that some members of society receive existence value because of the presence of rail passenger service. In part, existence value may explain Canada's "love affair" with rail passenger service. Again, however, the relevant issue here is whether improvements in service quality create additional benefits associated with existence value. That is to say, existence value must depend not only on the presence of rail passenger service, but with the quality of services.

Option values exist if some members of society receive benefits from the opportunity to use rail passenger service at some future date. The possibility of future events such as emergencies or higher fuel prices may generate benefits associated with option value. It is indeed likely that the availability of rail passenger service generates benefits associated with option value. The heavier than normal use of rail passenger service

during bad weather (e.g., when airports close) suggests the presence of option value. Similarly, the heavy use of Amtrak services in the U.S. Northeast Corridor during the era of high gasoline prices in the 1970's provides additional evidence of the existence of option value. Again, however, it is important to recognize that the issue here is whether improvements in service quality generate additional benefits associated with option value.

A number of distributional impacts may be associated with changes in rail passenger service. Direct (e.g., VIA Rail, CN, and CP employees) and indirect employment effects are examples. The availability of rail passenger service may also generate community impacts apart from those that directly affect rail passengers. Property values and wage rates, for example, may be higher in those communities for which rail passenger service is available.

Obviously, rail passengers themselves are affected by changes in services. To the extent that rail passengers represent lower income classes, improved rail service transfers the benefits to this cohort group. In short, improvements in rail service will tend to transfer benefits to those demographic cohorts that take advantage of the opportunity to travel by rail.

The list of distributional impacts provided in Figure 5-1 is certainly not exhaustive. If distributional impacts are considered as objectives in formulating Canada's rail passenger service policy, identification of all significant economic impacts is required. The standard tools of economic impact analysis can be applied to assess the distribution of benefits and

costs of alternative rail passenger service levels across different members of society.

Figure 5-1 also identifies those cases in which subsidies to rail passenger service may be appropriate. Subsidies may be an effective means of promoting allocative efficiency if the marginal cost of providing service quality decreases as service quality improves (e.g., the natural monopoly case) or if improvements in service quality generate non-user benefits that are not appropriable by VIA Rail through passenger revenues. As was noted above, it was not likely that the natural monopoly case is applicable to service quality.

External economies and existence value can never be appropriated through passenger revenues. Option value can be appropriated by VIA Rail only if it requires potential future passengers to buy passes (i.e., options) that permit them to purchase a ticket at some future date. In the absence of such a formal market, subsidies may be appropriate to address the option value problem.

Subsidies may also be appropriate if distributional impacts are included in VIA Rail objectives. The degree of subsidization here depends on the implicit value that the decision-maker (or society) places on the distribution of benefits and costs associated with rail passenger service. The appropriate structure of subsidization will depend on how specific aspects of service quality affect allocative efficiency and distribution of benefits and costs. In general subsidies should be structured such that they provide direct incentives for improving service quality. As a result, lump sum subsidies should be avoided.

5.1.2 Evaluation of Service Quality Alternatives —

In theory, benefit-cost analysis can be employed to identify the best of all possible alternatives. This, however, requires knowledge of the entire shape of both the benefit and cost curves associated with a continuum of alternative levels of rail passenger service quality. Given the existence of several categories of benefits -- some of which are very difficult to measure in practice -- numerous distributional impacts, and the inherent difficulty in estimating the costs of service quality, such information is not likely to be available in practice.

Because of the severe informational burdens, we suggest that a less ambitious approach be adopted in evaluating alternative levels of rail passenger service quality. This approach, which we refer to as applied benefit-cost analysis, is illustrated in the schematic displayed in Figure 5-2.

As has been stressed in the previous discussion, the first critical step is to specify objectives for Canadian rail passenger service. Next, a limited number of alternatives to existing rail passenger service must be specified. As a first step, these alternatives should be defined in view of previously determined objectives. In addition, the range of alternatives considered should reflect a relatively wide variation in service quality to permit subsequent analyses to discover information about how marginal costs and benefits move with variations in service quality. An evaluation of alternatives clustered in the neighborhood of existing service quality levels will not permit the analysis to learn much about the

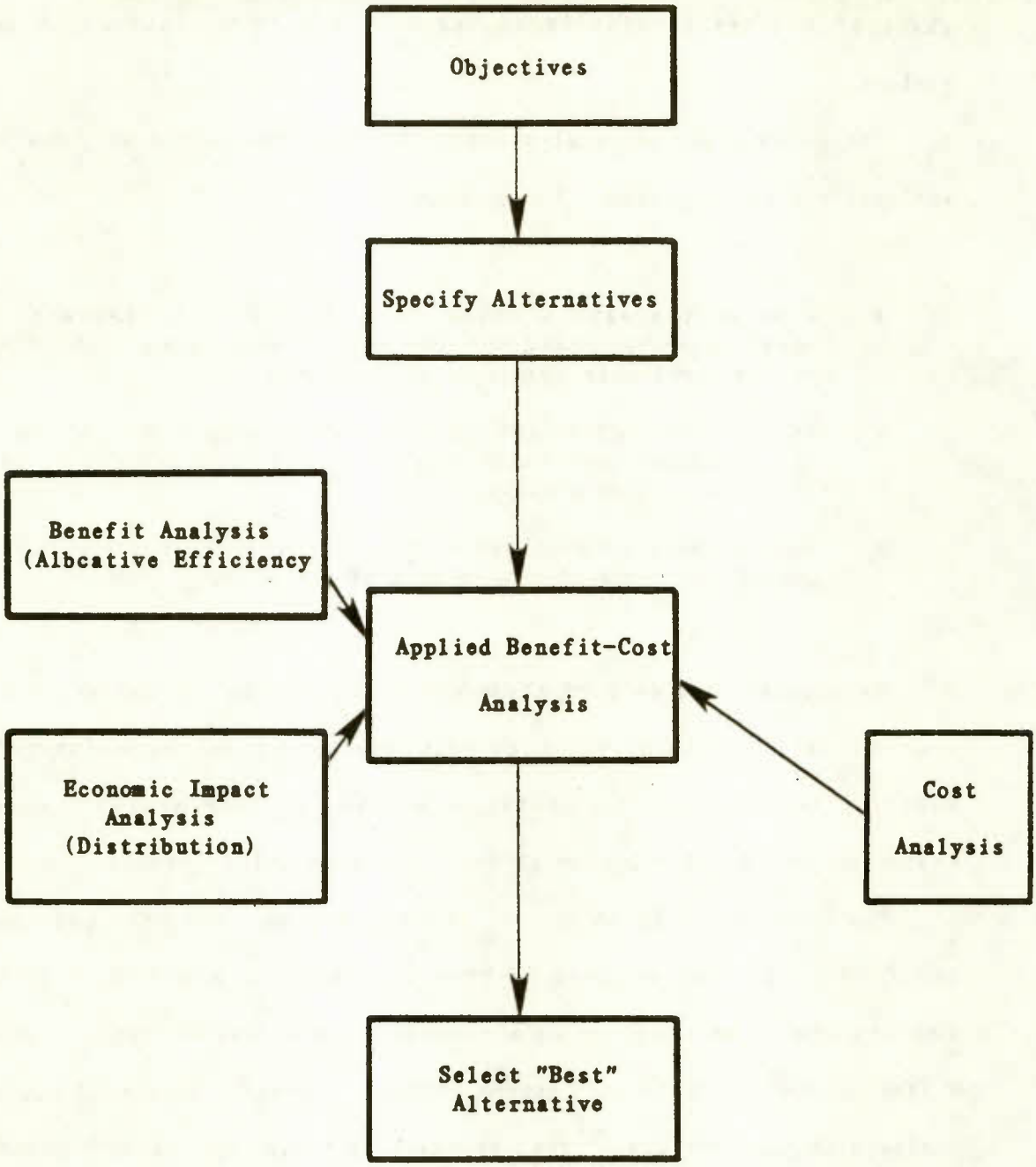


Figure 5-2. Evaluation of Service Quality Alternatives

shape of the net benefit curve associated with variations in service quality.

Three important sub-analyses must be conducted before the benefit-cost analysis can be completed. These include:

- A benefit analysis which is conducted to estimate the benefits -- both user and non-user -- associated with each of the previously specified alternatives.
- An economic impact analysis which is conducted to assess the distributional impacts associated with alternatives to existing service quality.
- A cost analysis which provides estimates of costs -- both direct and indirect -- for each of the alternatives.

We suggest that each of these three analyses be structured to measure incremental impacts of policies relative to a baseline which reflects current levels of service quality. Generally, incremental impacts are easier to evaluate for policy alternatives than total impacts.

The benefit analysis should distinguish between user and non-user benefits as defined earlier in this section. In addition, additional impacts should be analyzed on a separate track, and separate impacts on different demographic cohort groups should be traced. This will permit the analysis to identify those areas in which subsidies may be appropriate.

The final analytical step is the applied benefit cost analysis itself. In this analysis, the total incremental costs and benefits associated with each alternative level of service quality are evaluated and net benefits relative to the baseline standard are computed. At this point, the analyst or decision-maker must place implicit value on the distributional impacts estimated through the economic impact analysis. The best alternative among

those considered is the service quality level that generates the largest net benefits relative to existing service quality levels.

It is important to conduct separate applied benefit-cost analyses for the four major VIA Rail service groups. The benefits of rail passenger service and service quality may vary substantially across communities and passenger groups. In short, a rationalization of the passenger rail service requires independent evaluations of its separable parts. Certainly the availability and quality of alternative transportation modes plays a critical role here.*

Even this modest evaluation procedure may require information that is difficult to obtain in practice. Certainly, estimates of benefits, costs, and distributional impacts will be made with error. Nonetheless, we suggest that this procedure is useful for the purpose of evaluating policy toward rail passenger service quality. The technique is analytically sound, and can be adapted to make use of limited information that may be available to the decision-maker. Sensitivity analyses can be conducted to evaluate the potential impacts of benefit and cost categories for which estimates are unavailable or are known with error.

5.2 Institutional Issues

Institutional problems facing VIA Rail and Canadian rail passenger service in general have received considerable attention in recent studies.** Concern over institutional arrangements spans a broad spectrum

---

* See McQueen (1984) for a detailed discussion of this and related issues.

** See McQueen (1984), and Cubukgil and Soberman, (1984).

of issues related to Canadian rail passenger service. Some of these institutional problems are discussed below insofar as they create barriers to establishing appropriate incentive structures for providing rail passenger service quality at minimum cost.

In particular, institutional arrangements related to the following issues are discussed:

- VIA Rail as an enterprise.
- Cost efficiency control.
- Equipment problems.

5.2.1 VIA Rail as an Enterprise —

It has been noted earlier in this report that appropriate policy toward rail pasenger service quality must be established within a framework of the overall objectives of the enterprise. The paradoxical nature of the government enterprise and broad social objectives has been noted previously in this report. These issues are a logical starting point for a discussion of institutional problems confronting VIA Rail and its ability to effectively control service quality.

Clearly, VIA Rail does not and cannot behave as an enterprise in the traditional sense of a profit-seeking firm. First, it is not likely that rail passenger service can be profitable in Canada in any event; thus, the presumption here is that VIA Rail will require some subsidization regardless of institutional arrangements and the level of services it offers. Second, VIA Rail lacks, to a large degree, control over the very

services for which it has responsibility. This second issue will be discussed later.

Given the presumption that VIA Rail does not (and should not) behave as a profit-maximizer, the appropriate question is: What are the goals of VIA Rail and by what standards should its performance be measured? Earlier in this section, we have suggested methods for organizing objectives and evaluating alternatives, but have not stated what those objectives should be.

McQueen (1984) concludes that VIA Rail has never received a definitive statement of what its goals and objectives should be. Not surprisingly then, there are no definitive criteria upon which its performance is judged.

Accordingly, it is suggested here that establishing a clear set of objectives is the first institutional issue that must be resolved so that VIA Rail can adopt appropriate strategies to promote service quality. Again, the appropriate level of service quality will depend on specific objectives defined for Canadian rail passenger service. Clearly defined objectives are also necessary to determine the degree and nature of subsidization.

5.2.2 Cost Efficiency and Control —

It is taken as given that an objective of VIA Rail is to obtain the appropriate level of service quality at minimum cost. Presently, however, VIA Rail has little or no control over the costs of services it purchases

from the two rail carriers, the CN and the CP. This creates two distinct problems for VIA Rail:

- It cannot assure that its services are delivered to passengers at minimum cost.
- It is unable to identify the incremental costs of specific aspects of services and service quality; as a result, it cannot make intelligent policy decisions which require an evaluation of the benefits and costs of alternatives.

By some observers, the principle culprit here is CTC Costing Order R-6313.*

The Canadian Transport Commission (CTC) is charged with the responsibility of establishing the basis upon which the railroads charge costs for services provided to VIA Rail. Under R-6313, VIA Rail is theoretically required to pay carriers the long-run avoidable costs associated with rail passenger service.

In practice, VIA Rail makes monthly payments to the CN and CP based on the estimated costs of rail passenger services. At the end of the year, VIA Rail receives a bill for the balance. Costs are then audited by the CTC, but only to assure that they comply with the cost accounting procedures stipulated in the Commission's Railway Costing Regulations. Moreover, it has been argued that the carriers often bill VIA Rail for charges on sunk investments, thus violating the principle of long-run avoidable costs.**

---

* See Cubukgil and Soberman (1984), and McQueen (1984).

** See Cubikgil and Soberman (1984).

Clearly, the carriers have no incentive to reduce costs under this arrangement, given that they are reimbursed on a cost-plus, ex post basis. Nor, for that matter, does VIA Rail itself. Subsidies to VIA Rail are set at VIA Rail's own costs plus railway charges less passenger revenues. It appears that cost accountability passes from one tier in the system to the next.

The problem of cost control is exacerbated by the fact that VIA Rail is not permitted to audit railway charges. As was noted earlier, railway charges are audited by the CTC. In short, VIA Rail cannot determine in advance what its costs will be (because of ex post billing by railways); in addition, it has great difficulty in determining what actual costs were for specific aspects of service quality.

This latter problem imposes significant constraints on VIA Rail's ability to assess policy directed to service quality. As was noted earlier in this section, information on both incremental benefits and costs are necessary to evaluate projects embodying alternative levels of service quality.

It is instructive to contrast the arrangement between VIA Rail, the CN, and the CP, to contractual agreements negotiated between Amtrak and many of the railways from which it purchases services. In particular, Amtrak has been able -- in many cases -- to negotiate flat rate agreements for the services it purchases.

Although these agreements have been criticized by some observers as being too generous,* they do afford Amtrak the following advantages:

---

* See GAO (1977) and GAO (1981).

- Absent unanticipated inflation, Amtrak knows in advance what its charges for basic services will be (the contracts do make allowances for some inflationary cost increases).

- The fixed price agreements provide incentives to railways to provide services efficiently, since they are permitted to keep any net revenues if their actual costs are less than the flat charges.

- Since the contracts stipulate -- in a fair amount of detail -- charges for specific services, Amtrak has at least some information on the variable or "avoidable" costs of specific aspects of passenger services.

- The flat rates for basic service render Amtrak's performance incentive payments more effective, and provide some information to Amtrak on the incremental costs of service quality improvements.

This last point listed above is discussed in more detail later in this section.

In addition to the problems created by R-6313, it can be argued that the union-backed work rules contribute to higher railway charges to VIA Rail. The appropriateness of the long-run avoidable cost benchmark can also be debated.* There is, no doubt, merit to these arguments. While these factors may inflate railway charges to VIA Rail, they pose less of a control problem than the ex post billing practice.

Another important factor that reduces VIA Rail control over the services for which it has responsibility is the fact that it relies on sole-source suppliers for many critical inputs to the production of rail passenger services. Most importantly, these include train operations and

---

* For a comparison of VIA Rail and Amtrak costing approaches, see Canadian Transport Commission (1982).

maintenance services. Both of these services have critical impacts on service quality, especially reliability (on-time performance) and passenger comfort.

There are several institutional changes that may improve the efficiency of providing service quality and increase VIA Rail control over the services for which it is responsible. These include the following:

- VIA Rail should be permitted to negotiate flat-rate agreements with the CN and CP. This will encourage railways to reduce costs and afford VIA Rail some degree of control over costs. If the flat rates are set too high, they can be adjusted in subsequent contracts.

- If VIA Rail is to be responsible for policies directed to rail passenger service quality, it must have direct access to railway cost records pertinent to passenger service. This information is necessary if VIA Rail is to make rational decisions regarding alternative service quality projects.

- In addition to controlling railway costs, VIA Rail's own costs should be controlled. Accordingly, subsidies to VIA Rail should be based on a set of clearly defined performance criteria. This arrangement will be possible only if: 1) VIA Rail is given a set of clearly defined objectives; and 2) VIA Rail is given greater control over the services for which it is responsible.

- VIA Rail should attempt, as a long-run objective, to reduce its reliance on sole-source suppliers for critical inputs to rail passenger services.

Regarding the last point, it is our understanding that VIA Rail has plans to take over some of the maintenance services currently provided by the two railways. Some comments on these plans are offered below.

5.2.3 Equipment Problems —

There can be little doubt that equipment-related problems impose serious constraints on VIA Rail's ability to improve rail passenger service quality. Equipment condition affects all important aspects of service quality including reliability, passenger comfort, and schedule quality. Current institutional arrangements make it difficult for VIA Rail to deal with equipment-related problems.

A recent study by Mozersky et al. (1984) concludes that a high percentage of schedule adherence problems are related directly and indirectly to equipment problems resulting from poor performance of VIA Rail's rolling stock. This study notes that over 80 percent of VIA's cars and locomotives are 30 years old. The study also notes, however, that VIA's new LRC coaches and locomotives are also unreliable and expensive to maintain.

At present, VIA Rail has minimal control over equipment-related problems. The government provides funding for equipment purchases. As a result, the Minister of Transport determines the type of equipment operated by VIA Rail through the railways. Moreover, equipment maintenance services are currently provided by the CN and CP. As was noted earlier, VIA Rail currently has little control over these activities and the associated costs.

VIA Rail on-time performance incentive arrangements with the CN and CP were described earlier in Section 4 of this report. Problems created by equipment failures and institutional arrangements, however, render these incentives ineffective:

- First, it is important to note that delays caused by VIA equipment problems do not count against CN and CP on-time performance measures upon which bonuses are computed. As a result, there are no incentives for carriers to reduce delays caused by equipment-related problems, a major cause of poor on-time performance.

- Second, even if the equipment problem loophole was closed in the incentive agreements, it is not clear that on-time performance would improve. In the past, several U.S. railways refused to sign on-time performance incentive agreements with Amtrak. They argued that there was nothing they could do to reduce delays caused by Amtrak equipment failures.*

- Third, VIA Rail has little control itself over equipment-related problems. As was noted earlier, VIA does not directly make decisions on equipment purchases, nor does it have control over maintenance activities. The upshot of all this is that the three parties most responsible for delivering Canadian rail passenger service -- the CN, the CP, and VIA Rail -- have neither the incentives nor the control to mitigate the principle cause of poor on-time performance.

In addition, it is important to recognize that equipment problems lead to degradation in other aspects of rail passenger service quality. Many VIA Rail equipment problems are caused by temperature control systems; passenger comfort is directly affected here. Passenger comfort is also affected indirectly by delays, especially if the train breaks down or if the delays are excessive. Finally, the ability to improve schedule quality by shortening running times is certainly limited by the capabilities of rolling stock.

Among other recommendations, the Mozersky et al. (1984) study suggests that VIA Rail replace its aging fleet with new equipment. This recommendation is somewhat controversial, but it seems reasonable to conclude that

---

* See GAO (1977).

Canadian rail passenger service cannot be significantly improved without upgrading -- in some fashion -- the condition of its rolling stock. The decision to purchase new equipment, however, should be preceded by a thorough examination of goals for Canadian rail passenger service.

As was noted earlier, VIA Rail apparently intends to take over some of the maintenance services currently provided by the CN and CP. To the extent that maintenance activities and equipment failures are related, such an arrangement will improve VIA Rail's control over the problem. Given current institutional settings, however, it is not clear that such an arrangement will lead to the provision of maintenance services at minimum costs.

In the U.S., Amtrak has been directly responsible for most maintenance services for several years. By simple measures, Amtrak maintenance costs appear to be substantially less than those currently charged to VIA Rail.* However, there are several important and confounding factors that must be considered in assessing the cost savings that VIA Rail might realize under a similar arrangement.

First, and perhaps most importantly, Amtrak's fleet is substantially newer than VIA Rail's. Second, Canada's climate, on average, is more harsh than that to which Amtrak equipment is exposed. Many equipment maintenance problems are caused by Canadian winters. Finally, differences in work rules must be considered. Certainly, some cost savings will be realized if VIA Rail, as it assumes direct responsibility for maintenance services, increases labor productivity.

---

* See Mozersky et al. (1984).

One major reservation about changing responsibilities for maintenance activities is the current lack of incentives that VIA Rail itself has to minimize costs. Earlier, it was noted that current subsidies to VIA Rail are essentially based on the residual of its own costs and passenger revenues. Incentives for VIA Rail to minimize maintenance costs can be created if subsidies to VIA Rail are based on performance. Of course, such an arrangement is feasible only if definitive objectives and performance standards are established.

5.3 Comments on Existing VIA Rail Performance Incentives

Some specific comments of current VIA Rail performance incentives (or lack thereof) are offered below. In particular, incentives for each of the principle aspects of service quality -- reliability, passenger comfort, and schedule quality -- are discussed. The details of current VIA Rail performance incentives have already been described earlier in Section 4 of this report. The evaluation here considers explicitly the institutional framework within which these incentives are implemented.

5.3.1 Reliability —

Perhaps the most significant shortcomings of VIA Rail's existing on-time performance incentives are related to institutional problems that have been noted previously. First, railways are not penalized if delays are

caused by equipment failures.* Thus the CN and CP have no incentive to reduce delays caused by equipment problems, a major cause of poor on-time performance. As was noted earlier, however, it is not clear that significant improvements in reliability can be achieved without upgrading equipment. In short, eliminating the equipment failure loophole without upgrading equipment may be akin to pushing on a string.

The second major problem is created by the cost-plus billing arrangement. If railways are permitted to charge VIA Rail all costs anyway, (subject to cost accounting conventions) it is not clear what additional incentives are created by the bonuses. In all likelihood, there is some double-subsidization occurring here -- that is, railways receive payment once for "avoidable costs," and a second time in performance bonuses. This effect complicates the already difficult problem that VIA Rail faces in ascertaining the true incremental costs of improvements in reliability. Again, flat rates are required to correct this problem.

Apart from these issues, several specific comments on the structure of VIA Rail's current on-time performance incentive arrangement are appropriate. These comments include:

- Railways have little incentive to improve on-time performance above the 90 percent threshold, or to prevent service degradation below the 75 percent threshold.

---

* Amtrak does not allow exceptions because of equipment failures in its agreements with U.S. carriers. In fact, the only allowed exceptions are: 1) a train delivered late from another train; and 2) Amtrak itself holds up a train. Of course, Amtrak's fleet is newer than VIA Rail's, and presumably, less troubled by maintenance problems.

- Because on-time performance is not measured at intermediate points, railways have reduced incentives to provide reliability to passengers disembarking at points other than the final destination.

- The performance incentives do not provide penalties for excessive delays. In all likelihood, passengers care about the degree of lateness, as well as whether the train arrives on time or not. It is noted, however, that the ability of railways to earn credits for recovered time does provide some incentive for avoiding excessive delay.

Given the wide variation in historical on-time performance for VIA Rail passenger service, the 75 and 90 percent thresholds may indeed create incentive problems. On-time performance for selected VIA Rail trains during the 1980 to 1984 period are reported in Table 5-1. It is important to note that the on-time percentages reported in this table do not coincide with the performance figures employed by VIA Rail to compute on-time performance bonuses; accordingly, these figures are only illustrative. Nonetheless, the figures do show a wide variation in on-time performance -- well beyond the 75 to 95 percent range. It is interesting to note, that with few exceptions, Amtrak concedes no lower bound on the range for its on-time performance penalties. Of course, as is the case with VIA Rail contracts, Amtrak agreements stipulate that penalties cannot exceed bonuses.

This raises another issue. Some observers have commented on the provisions in both VIA Rail and Amtrak contracts that prevent penalties from exceeding some measure of current or previously earned bonuses. In short, the implication is that railways cannot lose money because of service degredation. As was noted earlier in Section 4, however, the important point is not that payments to railways cannot go below some

Table 5-1

VIA RAIL ON-TIME PERFORMANCE: SELECTED TRAINS, 1980-1984

ON TIME TRAIN PERFORMANCE
(PERCENTAGE - %)

TRAIN SERVCICES	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	AVG	
TRANSCONTINENTAL	1980	23%	57%	45%	75%	59%	35%	17%	36%	30%	73%	68%	36%	46%	
NO.1-2 CANADIAN	1981	78%	79%	93%	85%	89%	80%	62%	55%	75%	91%	93%	81%	80%	
	1982	20%	57%	61%	85%	84%	67%	74%	91%	68%	97%	92%	84%	73%	
	1983	87%	91%	97%	98%	84%	88%	71%	68%	89%	92%	95%	59%	85%	
	1984	80%	88%	92%	94%	92%	90%	70%	63%	99%	89%			86%	TO DATE
MONTREAL-HALIFAX	1980	81%	85%	82%	89%	89%	86%	87%	79%	92%	95%	78%	42%	82%	
NO.14-15	1981	39%	76%	74%	94%	94%	91%	70%	63%	83%	95%	91%	75%	79%	
	1982	16%	38%	67%	73%	92%	94%	75%	76%	98%	96%	99%	78%	75%	
	1983	76%	50%	73%	82%	81%	90%	87%	73%	97%	88%	90%	28%	76%	
	1984	50%	71%	63%	84%	89%	82%	70%	74%	90%	97%			77%	TO DATE
MONTREAL-STE FOY	1980	94%	78%	80%	93%	71%	82%	95%	92%	98%	86%	88%	74%	86%	
NO.20-21-22-23-24	1981	65%	71%	89%	81%	68%	80%	77%	71%	92%	84%	91%	89%	80%	
25	1982	66%	61%	86%	73%	78%	89%	90%	91%	92%	90%	92%	90%	83%	
	1983	93%	83%	90%	85%	56%	73%	87%	93%	96%	79%	87%	60%	82%	
	1984	76%	65%	82%	95%	87%	81%	90%	81%	85%	89%			83%	TO DATE
MONTREAL-OTTAWA	1980	91%	92%	83%	90%	88%	70%	72%	68%	84%	82%	86%	71%	81%	
NO.28-29-30-31-32	1981	63%	73%	74%	82%	59%	83%	89%	89%	94%	83%	85%	91%	80%	
33-34-35-36-37	1982	78%	79%	78%	91%	77%	51%	74%	76%	70%	64%	68%	72%	73%	
38-39-130	1983	82%	82%	84%	83%	79%	75%	62%	53%	59%	66%	79%	61%	72%	
	1984	68%	79%	58%	78%	69%	67%	65%	74%	77%	71%			71%	TO DATE
MONTREAL/OTT-TOR	1980	97%	90%	81%	95%	91%	81%	84%	74%	97%	93%	92%	81%	88%	
NO.43-44-45-46	1981	65%	84%	92%	86%	76%	87%	87%	83%	81%	88%	91%	86%	84%	
53-56	1982	68%	64%	83%	87%	83%	86%	84%	82%	76%	87%	84%	76%	80%	
	1983	77%	82%	92%	91%	82%	77%	80%	80%	83%	71%	83%	64%	80%	
	1984	65%	87%	58%	60%	60%	73%	87%	38%	59%	59%			65%	TO DATE
MONTREAL-TORONTO	1980	100%	100%	99%	99%	99%	97%	97%	98%	98%	98%	100%	91%	98%	
NO.58-59	1981	86%	95%	98%	100%	99%	97%	98%	98%	94%	94%	100%	99%	97%	
	1982	78%	88%	98%	97%	97%	99%	100%	98%	98%	98%	98%	98%	96%	
	1983	97%	93%	100%	96%	100%	98%	100%	95%	98%	99%	100%	83%	97%	
	1984	92%	95%	99%	99%	96%	94%	99%	97%	100%	97%			97%	TO DATE
MONTREAL-TORONTO	1980	94%	92%	87%	94%	78%	66%	63%	63%	90%	90%	91%	79%	82%	
NO.60-61-62-63-64	1981	67%	63%	82%	81%	71%	80%	86%	71%	66%	79%	53%	61%	72%	
65-66-67-68-69	1982	42%	44%	58%	74%	66%	67%	60%	57%	23%	32%	43%	49%	51%	
	1983	46%	46%	70%	70%	56%	54%	48%	49%	59%	72%	74%	38%	57%	
	1984	52%	71%	66%	78%	63%	89%	87%	80%	86%	81%			75%	TO DATE
TORONTO-WINDSOR	1980	88%	76%	72%	86%	69%	53%	64%	56%	78%	82%	88%	79%	74%	
	1981	81%	79%	83%	58%	43%	72%	83%	66%	70%	73%	78%	82%	72%	
NO.70-71-72-73-74	1982	59%	41%	70%	73%	75%	77%	79%	60%	82%	84%	70%	81%	71%	
75-76-77-78-79	1983	83%	84%	86%	66%	56%	57%	60%	68%	77%	71%	60%	71%	70%	
	1984	76%	63%	66%	84%	64%	80%	87%	83%	74%	87%			76%	TO DATE
TORONTO-SARNIA	1980	93%	89%	80%	88%	84%	80%	78%	63%	82%	93%	94%	81%	84%	
	1981	78%	92%	90%	81%	61%	66%	67%	68%	65%	69%	89%	89%	77%	
NO.80-81-682-83	1982	71%	59%	80%	70%	59%	59%	80%	71%	90%	83%	82%	84%	74%	
84-85-86-687	1983	83%	85%	86%	64%	64%	64%	68%	69%	81%	79%	73%	79%	75%	
	1984	85%	75%	86%	87%	65%	84%	93%	85%	84%	86%			83%	TO DATE

Source: VIA Rail

predetermined benchmark -- i.e., flat rates for Amtrak and avoidable costs for VIA Rail -- but that railways have no incentives to permit service degredation below some threshold point on specific trains. In strict economic terms, penalties (i.e., deductions from payments) and the loss of an opportunity to earn a bonus are equivalent. Accordingly, the threshold provisions, but not penalties per se, are what create disincentives.

Issues related to the size of the bonus payments have also been raised. Some observers have noted that Amtrak on-time performance payments constitute a much higher percent of total payments to railways than VIA Rail's. Table 5-2 reports VIA Rail and Amtrak on-time incentive payments as a percent of total payments to railways since 1974. It is clear that, in recent years, Amtrak payments as a ratio of total payments are significantly higher than VIA Rail payments. For example, Amtrak on-time payments represented almost 10 percent of total payments to railways in 1983, while the corresponding figure for VIA Rail is about 3.3 percent.

Much of this difference, however, is attributable to substantial differences in the scope of services that railways provide to the rail passenger corporations. The most significant difference is that Amtrak assumes direct responsibility for virtually all maintenance activities, while these services are provided by the CN and the CP for VIA Rail. A large part of this effect can be observed in the significant increase in the percent of on-time performance incentives in Amtrak in 1981 (relative to figures for 1980 and earlier). This is about the time that Amtrak began to assume maintenance responsibilities.

Total on-time performance incentive payments by VIA Rail were approximately 12.2 million Canadian dollars in 1983, while Amtrak payments totaled

Table 5-2

VIA RAIL AND AMTRAK ON-TIME INCENTIVE PAYMENTS
AS A PERCENT OF TOTAL RAILWAY PAYMENTS

Year	On-Time Incentives as a Percent of Total Payments	
	VIA Rail	Amtrak
1974	--	2.0
1975	--	5.8
1976	--	n/a
1977	--	4.2
1978	2.3	2.7
1979	3.0	n/a
1980	2.6	4.3
1981	2.6	8.0[b]
1982	2.6	8.9[b]
1983	3.3[a]	9.9

[a] Estimate based on pre-audit VIA Rail payments.

[b] Estimate based on assumed total Amtrak payments of $235 million.

Sources: Via Rail; U.S. Interstate Commerce Commission, Amtrak.

21.7 million U.S. dollars, or roughly 28.2 million in Canadian dollars. These figures are roughly comparable on a passenger-mile basis. For example, under the assumption that Amtrak passenger miles are at least two and a half times VIA Rail passenger miles, VIA Rail payments would have been 30.5 million in Canadian dollars if its passenger miles were equivalent to Amtrak's (i.e., 12.2 x 2.5).

It has been argued earler in this report that on-time performance incentive payments should be set at the point at which the marginal benefits and marginal costs of improvements in service quality are equated. This, of course, is difficult to achieve in practice, given the uncertainty in benefits and the difficulty of measuring marginal costs (especially for VIA Rail). Given this problem, it seems advisable that direct incentive payments should be increased as a proportion of total payments to railways, especially since fixed or cost-plus payments do not provide direct performance incentives. Nonetheless, the appropriate benchmark upon which to measure incentive payments is against the cost directly associated with the provision of service quality, and not total payments. For this reason, direct comparisons between the VIA Rail and Amtrak situation are inappropriate.

5.3.2 Passenger Comfort —

As was noted previously, many aspects of passenger discomfort are related to equipment problems. Failures of temperature control systems have already been noted; these have direct effects on passenger comfort. On-time performance also affects comfort, especially if passengers must be

re-routed or delays are excessive. The point here is that many important aspects of passenger comfort can only be improved within the limits of VIA Rail's rolling stock.

Nonetheless, VIA Rail contracts currently make no provisions for incentives to improve passenger comfort. Indeed, it is fair to say that VIA Rail currently has minimal control over this aspect of service quality. Accordingly, VIA Rail may wish to consider some options.

Direct incentive payments (or penalties) can be tried, but Amtrak's unhappy experience during the 1970's should be considered. As was noted earlier, the reasons most often cited for the failure of the experiment are difficulties in establishing objective measures of passenger comfort, and problems in monitoring and enforcing the agreements. Still, at least some observers argue that Amtrak abandoned direct incentives for passenger comfort too quickly.*

The recommendation here, however, is that VIA Rail considers, as a long-run goal, tendering many services related to passenger comfort competitively.** Competitive tendering is not feasible if large sunk costs are present, but it is not clear that this would be a problem for many services related to passenger comfort. It seems reasonable to suspect that at least cleaning and light maintenance services can be tendered competitively. These services affect many aspects of passenger comfort. It is also worth repeating that incentives may also be employed along with competitive tendering if objective performance standards can be established, and enforcement costs are not prohibitive.

---

* See GAO (1981).

** VIA Rail currently employs competitive tendering in purchasing food and beverages served on-board to passengers.

5.3.3 Schedule Quality --

Schedule quality -- both in terms of trip frequency and running times -- is currently determined through negotiations. Given Amtrak's failure in its experiment with direct incentives for schedule quality, there is some reluctance to recommend this approach here.

It can be argued that the Amtrak experiment failed because the incentives to reduce running times were far too low relative to on-time performance incentives. Theoretically, of course, it is possible to construct the optimum mix of schedule quality and schedule adherence incentives. In practice, however, this may be difficult to achieve.

If schedule quality continues to be negotiated with the railways, additional improvements are likely to depend on the negotiating leverage held by VIA Rail. This being the case, the earlier comments on VIA Rail control apply here as well.

5.4 Data Requirements for Service Quality Policy

The preceeding discussion has focused on structures and mechanisms for creating incentives for improving rail passenger service quality. The following discussion reviews the data requirements and identifies methods of obtaining such data necessary for implementing a rational policy toward establishing the appropriate level of service quality. In other words, the question here is: How much service quality is optimum, given a set of objectives for Canadian rail passenger service? Data requirements are also

described with a view to information required to establish subsidization policy.

It should be noted that VIA Rail's current policy toward service quality is, to some degree, consistent with a set of implied objectives. The schedule adherence bonus pool, for example, is allocated across the major service groups (e.g., transcontinental, corridor, etc.) based on some set of implied goals. In addition, the pool is allocated by train based on seasonal factors and passenger miles thus accounting for higher benefits associated with greater passenger loads. Although this allocation scheme can probably be improved, there are advantages associated with simplicity.

In any event, the data required for service quality policy must include, at a minimum, information on both the incremental benefits and costs associated with quality changes. The data requirements for estimating incremental benefits are described first.

5.4.1 Estimating Benefits of Service Quality —

It is important at the start to recognize that benefits should be estimated by distinct categories. The important categories are:

- User benefits.
- Non-user benefits.
- Equity or distributional impacts.

These categories are important both in terms of objectives and subsidization policy.

User benefits apply to goals related to allocative efficiency; absent the natural monopoly problem, subsidies here are inappropriate. Non-user benefits also apply to allocative efficiency; subsidies may be appropriate if VIA Rail cannot appropriate these benefits. Distributional impacts apply to goals related to equity considerations; subsidies may be appropriate here as well.

Perhaps the easiest case is user benefits. Information on existing passengers' willingness to pay for service quality is sufficient to estimate user benefits. Recall, however, that estimates of willingness to pay for both marginal and infra-marginal buyers are required. As a result, survey data is necessary.

Baumol (1975) suggests a simple approach. Specifically, Baumol suggests that surveys of passenger complaints be employed to identify those aspects of service quality that are important to passengers. Naturally, standard sampling rules should be applied, and complaint forms should be readily available, comprehensive, and easy to use. The relative mix of complaints across items will reveal those aspects of quality that passengers deem most important to improve. Absolute increases in the number of complaints -- appropriately adjusted for changes in passenger miles and possible design effects -- will reveal changes in consumer preferences and possibly, service degradation as perceived by passengers.

Baumol's approach is simple, relatively inexpensive, and will provide useful information on which services passengers prefer improvements. That is to say, the information can be used to establish the direction of service quality policy. But, because this approach does not provide

estimates of passengers' willingness to pay, it does not afford the information required to determine the optimal level of service quality.

There are, however, standard tools available in the economics and marketing literature that can be employed to estimate willingness to pay. In particular, attribute analysis can be employed to reveal the marginal preferences of passengers for trips embodying alternative levels of quality.* Specifically, survey subjects are asked to rank by order, their preferences for services embodying several different attributes. In order to reveal willingness to pay, price (i.e., fare) must be one of these attributes. Subjects are placed in a moral hazard in revealing willingness to pay, but techniques for adjusting for this bias are available.

Unfortunately, obtaining estimates of non-user benefits is somewhat more difficult. The least expensive approach is to conduct sensitivity analyses about reasonable ranges of values for non-user benefits.

Formal techniques for measuring existence and option values have been developed. These techniques, which are described in the environmental economics literature, generally require the use of survey data. It should also be noted that such studies may produce estimates with considerable error. Accordingly, sensitivity analysis is appropriate.

The standard tools of economic impact analysis can be employed to estimate distributional impacts and the benefits associated with external economies. Community and employment impacts, and equity effects across economic and demographic cohorts are effects to be considered if these are included in the set of objectives for rail passenger policy.

---

* This type of analysis is often used as a marketing research tool in the private sector.

To the degree possible, benefits should be estimated by train and major service group. There are two reasons for this suggestion. First, the benefits associated with different aspects of service quality may vary depending on the mix of passengers. Second, the rationalization of the rail passenger system requires an evaluation of the net benefits associated with separable parts of the system.

5.4.2 Estimating the Costs of Service Quality —

In principle, at least, measuring the direct costs of improvements in service quality is more straightforward than measuring benefits. This statement, of course, is based on the assumption that institutional changes are made that permit VIA Rail to negotiate flat rates for payments other than direct incentives and that VIA Rail be granted direct access to all railway cost records related to rail passenger service.

Incremental costs must be defined and measured as dollars per change in quality. The difficult tasks here are to measure changes, and to define units of quality. Quality is relatively easy to define for reliability and schedule quality, and relatively difficult to measure for passenger comfort.

If direct incentive payments for on-time performance are employed, the marginal cost of a given percentage improvement, on the margin, is equal to the payment (per unit of on-time percent) itself since the producer will maximize profits by producing quality up to the point where the unit

incentive payment is equal to its own marginal costs.* In theory then, the shape of the marginal cost of quality curve can be measured by observing how on-time performance varies with changes in the incentive payment structure (i.e., as the bonus per unit of on-time percent varies).

The complicating factor here, however, is that other factors affect on-time performance, especially weather and equipment operability. A number of techniques are available to adjust for these factors, but the necessary data must be collected (e.g., measures of weather severity and equipment failures). It should also be noted that the marginal cost of reliability will likely vary substantially across trains because of variations in roadbed conditions and traffic flows.

Measuring the cost of passenger comfort is somewhat more complicated, but it should be possible to obtain rough estimates. The preceding comments apply if direct incentives are employed (since presumably, objective measures of performance are available).

If passenger comfort is negotiated or imposed by fiat (i.e., competitive tendering), then the following options are available:

- Identify specific resources and costs devoted to passenger comfort from railway (or other firms) records.
- Impute costs based on variations in costs and quality received from competitive bids.

---

* Note that direct costs obtained from railway records will tend to understate costs because they will not include indirect costs such as those associated with impacts on freight traffic.

If objective measures of passenger comfort are unavailable, responses from passenger complaints can be used to measure historical changes in service quality. In this case, perceived changes by passengers matter.

A final note is appropriate here. In Section 5.4.1, we noted the possible presence of non-user benefits. External or indirect costs may also be present. If allocative efficiency is a goal, these indirect costs should also be considered. As is the case for non-user benefits, external costs are somewhat difficult to measure. These can at least be identified through economic impact analysis.

REFERENCES

Baumol, W.J. (1975). "Payment by Performance in Rail Passenger Transportation: An Innovation in Amtrak's Operations," The Bell Journal of Economics, Vol. 6, No. 2, pp. 299-313.

Canadian Transport Commission (1982). A Comparison of Amtrak and VIA Rail Costing Approaches. Staff Report, 30-3025-77.

Cubukgil, A. and R.M. Soberman (1984). The Cost of Rail Passenger Service in Canada: An Examination of Institutional Problems. Transportation Development Centre, TP 5823E, Montreal.

Demsetz, H. (1968). "Why Regulate Utilities," Journal of Law and Economics, Vol. II, pp. 55-65, April.

DeVany, R.S. and T.R. Saving (1983). "The Economics of Quality," Journal of Political Economy, 91:6, pp. 979-1000, December.

DeVany, A.S. (1976). "Uncertainty, Waiting Time, and Capacity Utilization: A Stochastic Theory of Product Quality," Journal of Political Economy, 84:3, pp. 523-541, June.

DeVany, A.S. (1975). "Capacity Utilization Under Alternative Regulatory Restraints: An Analysis of Taxi Markets," Journal of Political Economy, 83:1, pp. 83-94, February.

General Accounting Office (GAO) (1981). "Further Improvements Are Needed in Amtrak's Service Contracts, But They Won't Come Easily," CED-81-35, Washington, DC, January.

General Accounting Office (GAO) (1977). "Amtrak's Incentive Contracts with Railroads -- Considerable Costs, Few Benefits," CED-77-67, Washington, DC, June.

Hendler, R. (1975). "Lancaster's New Approach to Consumer Demand and Its Limitations," American Economic Review, Vol. 65, pp. 194-199, March.

Ippolito, R.A. (1981). "Estimating Airline Demand with Quality of Service Variables," Journal of Transport Economics and Policy, 15:1, pp. 7-15. January.

Kahn, A.E. (1970). The Economics of Regulation: Principles and Institutions, Vols. I & II, John Wiley & Sons, Inc., New York.

McQueen, D. (1984). Aspects of Rail Passenger Policy in Canada. Canadian Transportation Commission, WP 10-84-09, Toronto.

Mozersky, K.A., M. Parry, and S. Pepin (1984). Report of the Inquiry into the On-Time Performance of VIA Rail Canada, Inc. Canadian Transport Commission, Ottawa.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy, Vol. 82, No. 1, pp. 34-55, January/February.

Rubin, P.H. (1978). "The Theory of the Firm and the Structure of the Franchise Contract," Journal of Law and Economics, 21:1 pp. 223-244, April.

Spence, A.M. (1975). "Monopoly, Quality and Regulation," The Bell Journal of Economics, Vol. 6, No. 2, pp. 417-429, Autumn.

Thompson, R.S. (1980). "Quality Control Under Regulation: A Comment on the Prior Report," Applied Economics, 12:3, September, pp. 357-362.

Trapani, J.M. and C.V. Olson (1982). "An Analysis of the Impact of Open Entry on Price and the Quality of Service in the Airline Industry," Review of Economics and Statistics, 64:1, pp. 67-76, February.

Williamson, O.E. "Franchise Building for Natural Monopolies: In General and with Respect to CATV," Bell Journal of Economics 7, pp. 73-104.