MANAGERIAL AND ECONOMIC APPROACHES TO MEASURING FREIGHT RAILWAY PERFORMANCE

Carl D. MARTLAND
Senior Research Associate
Center for Transportation Studies
Massachusetts Institute of Technology
Cambridge - USA

1 INTRODUCTION

Much attention has been given to railway performance by people outside of the rail industry. At an aggregate level, economists are concerned with productivity as it relates to such issues as the deficits of nationally-owned systems, the effects of regulation, and the structure of transportation systems. At an operational level, many operations researchers seek to optimize such things as power management, empty car distribution, and train scheduling. Both levels of concern are also of interest within a railroad.

From the manager's perspective, however, a more complex approach to measuring performance is available and desirable. The economist's estimates of productivity are often too broad to influence strategic decisions, while optimization techniques are too narrowly focussed to have much influence on operations. To improve operating decisions, it is necessary to develop better measures for monitoring internal performance and better information that can be used in a timely fashion by railroad managers. With recent advances in computers and communications, it is possible to design information systems so that managers can understand and react to incremental costs, productivity, and cost/service tradeoffs.

The thesis of this paper is that it is possible to improve freight railway performance by providing more timely, more complete, and more consistent measures of performance to managers at all levels of the company. Areas of chronically poor performance are likely to be those that are influenced by uncoordinated decisions made by managers from different departments. Thus, an integrated set of performance measures is likely to be most important where operations are most complicated, e.g. in managing the service provided for general merchandise traffic.

This paper is based upon insights resulting two decades of research into railroad operations and economics conducted by the MIT Center for Transportation Studies with support from the U.S. Department of Transportation, the Association of American Railroads, and individual railroads. Much of this research dealt with track maintenance planning, freight car utilization, and railroad reliability, which are three areas where performance measurement is critical, as described in Section 3.

2 AN INTEGRATED SET OF PERFORMANCE MEASURES

Railroads usually build their performance measures around financial performance, beginning with accounting systems for both revenue and expenses budgets for operations and capital investment [Martland, 1992]. The origin-to-destination (O-D) movement is the next natural level for performance measurement. The O-D trip provides an obvious basis for measuring prices, trip times, reliability, loss & damage, and equipment availability. Costing is more complicated, since expenses are recorded by functional area, and there are many joint costs to be allocated in order to obtain O-D costs.

Government regulatory bodies define rules and procedures for defining and reporting expenditures. It is these rules and procedures that generally determine the information that is available to outsiders for productivity and other economic analysis. However, the purpose of these rules and procedures is to support a regulatory process, not to provide managers with accurate costs on which to base decisions. Railroads therefore may use other costing systems for routine cost allocation or for special studies of short-run costs.

Most railroad managers are concerned with specific portions of cost or operations, rather than O-D service or profitability. The operating department has an operating plan; it may have service standards; and it has a budget for the costs of operating trains and terminals. The mechanical department has objectives for equipment serviceability and a budget for maintenance and rehabilitation. Likewise, the engineering department has standards for track performance and a budget for maintenance and rehabilitation. In day-to-day operations, managers have very specific concerns, such as moving trains across a line as quickly as possible, giving priority to the most important class of trains. In yards, only a few people worry about movements *through* the yard; most are involved in either classifying or assembling trains.

Only the marketing department has revenue goals; if it has profitability goals, they are likely based upon long-run average costs as computed by the standard cost allocation system. These long-run costs often bear little or no relationship to the actual day-to-day costs experienced at the local level, which vary with traffic volume, weather conditions, and the level of track maintenance activity and equipment availability.

To summarize, the objectives for most managers relate to cost and service standards for their small piece of the overall operation; very few managers are directly concerned with overall cost or service. To have all managers working diligently toward their own objectives is therefore insufficient; their objectives must be well-coordinated for the system to perform at its best. To the extent that managers respond to changes in their budgets and their service standards, providing better objectives will lead to better performance.

3 THE ROLE OF PERFORMANCE MEASUREMENT

3.1 HEAVY AXLE LOADS

Increasing axle loads is one way to improve the productivity of transporting coal and other bulk commodities that are transported in unit trains. Coal generally moves in unit trains under long term contracts that specify equipment utilization (turns/per year) and minimum annual volumes. Service *per se* is important only because of the need to minimize the number of cars and locomotives required to move the coal; trip times and reliability mean very little to a power plant with a 6-month stockpile of coal.

Increasing axle loads is a way to achieve a lower net-to-tare ratio, more net tons per length of train, smaller fleets, and greater flexibility in making up trains. The basic issue is whether the resulting savings in fuel, crew, and equipment costs justify more rapid deterioration of the track structure.

In the 1970s, when the U.S. rail industry increased axle loads from 27- to 33-tons, there were marked increases in track failures, more rapid deterioration of marginal facilities, and extensive emergency maintenance. The problems of 33-ton axle loads were eventually overcome through the development of better rail steel, better maintenance practices, and advanced inspection techniques, and the industry began to contemplate still heavier axle loads. In order to avoid the painful experiences of the 1970s, the North American railroads sponsored tests of 39-ton axle loads at the FAST track in Pueblo Colorado and created a committee to investigate the economics of going to heavier axle loads. The economic analysis [Hargrove, 1991] considered increasing axle loads from 33-tons to 36- or 39-tons for several different train configurations operating over terrain varying from flat and level to mountainous. The main conclusions were as follows:

> *The operation of HAL equipment is a question of economics, and not one of technical feasibility. ... total cost savings (including the impact on track costs) of as much as 5%-7% per net ton mile could be obtained under specific HAL operations ... the results are highly route and service specific.*

This example shows the importance of looking at the total picture, not just at operating or equipment or track costs. In the 1970s, axle loads were increased without consideration of the impacts on track; in effect, the railroads believed the costing allocation formulas that they used for regulatory proceedings. In the 1980s and 1990s, the railroads were much more careful in approaching heavier axle loads. They did not use regulatory costing techniques, but instead developed cost functions based upon the underlying engineering relationships.

3.2 EQUIPMENT UTILIZATION

Equipment utilization is important for both bulk and merchandise service. Fleet size, marketing priorities, operating plans, equipment distribution rules, loading and unloading incentives, maintenance efficiency, and rental/usage rules all influence equipment utilization. Hence, a carefully integrated set of performance measures is essential for controlling car utilization [Task Force I-2, 1980].

Where operations are simple, equipment utilization is high. Equipment used in unit train service may make a trip per week, whereas equipment used in general freight service may only make a trip per month. For unit trains, the operation typically involves a loaded move from A to B followed by an empty move from B to A. Utilization is high and control is simple, even if two or more railroads are involved in the move.

For general merchandise, the opposite is true: utilization is low and control is complex. There are many more origins and destinations, many more options for routing the empty equipment, and many more complications introduced by the mixed ownership of the fleet. Managing a general merchandise fleet is a retail, not a wholesale operation.

Table 1
Changes in Freight Car Utilization 1980-1989

	Cycle Time (days)			Fleet Size (1000s)		
	1980	1989	% Change	1980	1989	% Change
Total Flats	17.6	9.8	-44.3%	152.7	129.6	-15.1%
Total Gondola	20.5	14.0	-31.7%	184.9	133.3	-27.9%
Total Hopper	14.7	11.4	-22.4%	347.9	229.8	-33.9%
Total Tank	47.7	43.9	-8.0%	184.0	179.9	-2.2%
Total Box	34.3	35.4	3.2%	431.4	192.2	-55.4%
Total Refrigerator	36.4	39.8	9.3%	79.4	45.8	-42.3%
Total Covered Hopper	26.6	27.0	1.5%	300.0	287.7	-4.1%
Other	47.7	43.9	-8.0%	30.7	15.1	-51.0%
Total	23.2	18.4	-20.7%	1710.8	1213.4	-29.1%

Table 1 shows the car utilization by car types for the period 1980 to 1989, measured as the serviceable car days per load originated, for the fleet as a whole [Economics & Finance Dept., 1990]. Over this period, the overall cycle time improved nearly 20%, from 23.2 days to 18.4 days, but there are clearly three distinct categories of cars. Flat cars, which are used primarily in piggy back service and in transporting automobile, had reductions of roughly 40% in their cycle times. Gondolas and open top hoppers, which are typically used in unit train service, showed reductions in cycle times of about 20%, in part reflecting a reduction in the amount of surplus equipment. Box cars, covered hoppers, refrigerator cars, and tank cars all had slight increases in their cycle times, despite considerable reductions in the number of surplus cars.

Development of car management systems, improved organizational structures, and improved institutional arrangements have as yet had a limited effect on the utilization of general purpose equipment. The fundamental problem has been that too many cars were purchased in the late 1970s at a time when the merchandise traffic was shifting to faster, more reliable modes, i.e. truck and intermodal [Sloss and Martland, 1984].

Although empty car distribution might appear to be a good area for applying OR techniques, success has been limited. Dejax and Crainic [1987] found only three models that were implemented successfully, and all of these were simple linear programming techniques that had been implemented prior to 1975. Railroads seem to have avoided more advanced OR techniques, not because of any lack of interest in or fear of the techniques, but because general improvements associated with centralizing (and thereby controlling) car distribution proved so much more effective.

A more recent study examined the overall car cycle for boxcars using a random sample of 1% of the movements from the AAR's TRAIN II car cycle data base [Little et al, 1991]. The results from this study can be compared to the results for two earlier studies. The 1990/91 cycle for box cars was slightly above the 1977/78 car cycle for plain box cars, but somewhat below the cycle obtained for plain box cars in an earlier Reebie study [Reebie Associates, 1972]. The average loaded trip time was lowest in the first study, and worst in the second study; the empty time was longest in the most recent study. Customer time improved steadily.

Table 2
Comparison of Cycle Time for Box Cars, 1977/8 and 1990/1

Cycle Component	Mean Time, 1970	Mean Time 1977/78	Mean Time 1990/91
Shipper	3.1	2.0 days	2.2 days
Loaded Trip	10.4	11.3	8.3
Consignee	2.7	1.7	1.6
Empty Trip	12.2	9.3	13.7
Total	28.4	24.3	25.7

A couple of caveats should be noted. The 1990/91 study includes both plain and equipped box cars, and equipped box cars tend to have slightly shorter trip times and cycles. Also, traffic conditions affect the overall car cycle. In 1970, traffic levels were at intermediate levels for the 1969-71 period, and it is unlikely that either equipment surpluses or congestion were major problems. In 1977/78, rail ton-miles on the Class I railroads reached a record high of 858 billion, there was a severe winter, and yard congestion and car shortages were major problems in many locations [AAR, 1979]. In 1990/91, yard congestion was not a major problem, and there were surpluses of box cars.

Also, the approaches to measuring the car cycle used in Table 2 ignore any cars that were stored serviceable and therefore give lower estimates than that used in Table 1, i.e. total serviceable car-days divided by total loads originated in box cars.

Despite these caveats, the striking result is how similar the cycle time components were over this 20 year period. Dramatic breakthroughs are not evident, as they were in Table 1 for bulk and intermodal traffic. The very significant changes in track and vehicle technology and in control technology have not yet resulted in major improvements in trip times or cycle times for box cars.

3.3 SERVICE RELIABILITY

Service reliability is influenced by many operating, engineering, and marketing factors. Improving service reliability therefore requires close attention to the performance measures used to support and evaluate decisions made in each railway department. Studies of freight service reliability during the 1970s and 1980s generally emphasized the importance of control systems and downplayed the importance of track and equipment reliability. The Industry Task Force on Reliability Studies [1977], after reviewing the FRA-sponsored studies and supervising a dozen case studies involving most of the largest railroads, concluded that:

> *Few railroads presently are organized and have the data systems necessary to provide reliable service on other than a limited amount of selected traffic. ... development of an effective operating/service plan and an organization geared towards successfully executing that plan are key elements in fulfilling a commitment to improve service reliability and car utilization.*

This task force then supervised a case study of the Boston & Maine in which B&M adopted an integrated set of standards for train connections, yard times, O-D trip times and O-D reliability. An inter-departmental service committee used a network model to examine alternative operating strategies, implemented an improved operating/service plan, and monitored performance relative to the new service standards. B&M also established a separate department for equipment management that allocated responsibility for various aspects of car utilization to the other departments. The net result of these efforts was an improvement of over $3 million annually (3% of freight operating costs), most of which was attributed to the improved planning process and interdepartmental coordination

[Martland, Marcus, and Raymond, 1986]. This study clearly demonstrated that improvements in the control system can lead directly to improvements in performance.

A follow-on study of the Santa Fe, a much larger railroad, documented the cost-service trade-off for general merchandise freight [Martland, Clappison and Van Dyke, 1982]. The study took place during an extremely congested period for Santa Fe and other railroads. As a result, the operating plan was seldom followed and service reliability was poor. The difference between actual base case performance and any of the alternatives studied was found to be greater than the differences among the alternatives. Therefore, this study also emphasized the importance of having a feasible plan and providing the resources to operate according to the plan.

In 1990, the AAR initiated a new series of studies concerning railroad reliability, prompted by the importance attributed to this topic by senior railroad officials and by shippers [Barenberg and Wormley]. One of the primary objectives of this study was to examine the causal relationships underlying reliability, including the cascading effects of engineering and mechanical failures. With more unit train traffic and higher main-line densities, it was felt that engineering and mechanical reliability might be more important now than when the earlier studies were done. In fact, that does appear to be the case for lines operating near capacity or for high priority intermodal trains, where reliable train arrivals are a key component of service.

However, for general merchandise traffic, the engineering and mechanical issues continue to take a back seat to operational and institutional issues. In particular, the typical level of reliability for general merchandise traffic in 1990 was little different from what was seen in the 1970s. For example, the Industry Task Force on Reliability Studies [1977] documented service for 61 O-D pairs for three large shippers; the overall average trip time was 6.8 days, while 57% of the cars arrived within a two-day window. The recent MIT study of box car traffic [Little, Kwon, and Martland, 1992] found that the average trip time for all box car shippers was just over 8 days, while the average for the 100 largest shippers was just under 6 days. The average percentage of trips arriving within a two-day window ranged widely, with an average of about 60%. Discussions with railroad officials highlighted the same problems of control: the need to develop and use a feasible operating plan that is designed to provide the required levels of service and the need to develop better performance measures to guide operating managers.

With significant advances in control capabilities over the past decade, the potential for gaining control over service is more promising now than in the 1970s:

a. Computer assisted dispatching: railroads have taken advantage of communications and computer technologies to consolidate their dispatching centers
b. Car cost budgets: car costs, once largely ignored, are now routinely incorporated into most terminal and operating budgets and traded off against crew and other costs.
c. Car scheduling: most of the largest railroads now have the ability to generate trip plans for any shipment and to monitor progress against the trip plan
d. Interline service management: several committees are investigating ways to manage interline service, e.g. to allow a customer to call any participating railroad to complete information concerning shipment status and estimated time of arrival [Ad Hoc Committee, 1991].
e. Customer service centers: customer service has been centralized, in order to ensure more consistent response to customer inquiries and to allow better supervision.
f. Coordination between operations and engineering: many railroads are using or developing tactical planning systems to assist in allocating track time to maintenance forces without disrupting priority trains.

In some respects, though, the old problems are still visible. Despite high level support for improving service, costs are still apt to be more visible to operating managers than service objectives. A group of industry officials working to promote better interline service emphasized the importance of changing the culture of the typical railroad organization:

> *Management will have to instill a service priority philosophy across the railroad and in the field. Consciousness of the importance of service reliability and of meeting the ETA (estimated time of arrival) must be integrated directly into the line operating decision process. That in turn means evaluating individual and unit performance first on the basis of adherence to the operating plan to achieve ETAs or ETIs and only secondarily on the basis of minimizing costs and controlling budgets. [Ad Hoc Committee, 1991]*

There is still no system for reserving capacity, so that congestion during peak periods continues to cause many cars to be delayed. While average operating costs are much better understood, systems have yet to be developed that provide timely information concerning *today's* costs in order to guide *today's* decisions. Critical operating decisions concerning such things as train consolidation and the operation of extra trains still receive relatively little support from most information systems. While priority traffic does receive priority service, the costs of providing such service are still poorly understood. Also, service priorities, whether in yard or line operations, are seldom adjusted based upon whether cars are ahead of or behind schedule or whether cars are susceptible to missed connections at the next terminal. More generally, railroads are just beginning to come to grips with market segmentation and service differentiation. Marketing departments are more influential in setting priorities, but it is still not clear how best to integrate marketing and operating concerns. In short, there are still major opportunities for improving the performance measures and information that are made available to the railroad officials responsible for general merchandise traffic.

4 CONCLUSIONS

This paper presented examples from three areas to illustrate the importance of having an integrated set of performance measures. These examples show that providing managers with more timely, more complete and more consistent performance measures can improve performance, while failing to do so hinders performance. Where operations are least complex and where railroads are most competitive, namely in the high volume movement of bulk commodities and intermodal freight, railroads have achieved their greatest improvements in productivity. Where operations are most complex, and where performance is most difficult to measure, railroads have had the least success. General merchandise service and equipment utilization in the 1990s seems to be little different from service in the 1970s, despite the great advances in railroad technology.

To improve service for general merchandise traffic, it will be essential to gain control over operations and to make the proper trade-offs between cost and service. It will be more beneficial to eliminate gaps and inconsistencies in performance measurement than it will be to improve vehicle and track technology or to develop specific optimization routines for routing and scheduling. Specific opportunities include the refinement of priority schemes for line and terminal operations, the development of better measures to support operations control, and, more generally, better measures concerning the incremental costs and benefits of operating decisions.

5 REFERENCES

Ad Hoc Committee to Develop ETA and Trip Plan Capabilities Among Railroads, *A Proposal for Systems to Support Interline Service Management*, AAR Report R-776, Association of American Railroads: Washington, D.C., 1991

Barenberg, E.J. and D.N. Wormley, *Survey of Research Needs in the Railroad Industry*, Association of American Railroads: Washington, D.C., 1990

Dejax, P.J. and T. G. Crainic, *A Review of Empty Flows and Fleet Management Models in Freight Transportation*, **Transportation Science,** Vol. 21, 1987 pp. 227-246

Economics and Finance Department, *Average Serviceable Car Days per Load Originated,* Association of American Railroads: Washington D.C., 1990

Hargrove, M.B., *Economic Assessment of Increased Axle Loads Based on Heavy Axle Load Tests at the AAR Transportation Test Center - Pueblo*, AREA Bulletin No. 732, American Railway Engineering Association: Wash. D.C., 1991, pp. 227-244

Little, P., O.K. Kwon, and C.D. Martland, *An Assessment of Trip Times and Reliability of Boxcar Traffic*, MIT/AAR Affiliated Laboratory Working Paper 91-3, Center for Transportation Studies, MIT, Cambridge, MA, Oct. 1991

Martland, C.D., *Rail Freight Service Productivity from the Manager's Perspective*, **Transportation Research,** 1992 [in publication]

Martland, C.D., P.A. Clappison, and C. Van Dyke, *A Case Study of the Santa Fe's Northern and Southern Divisions,* **MIT Studies in Railroad Operations and Economics**, Volume 35, Center for Transportation Studies, MIT, Cambridge, MA 1982

Martland, Carl D., Henry S. Marcus, and George B. Raymond, Jr., *Boston & Maine Achieves Control Over Railroad Performance*, **Interfaces**, Vol. 16:5, September-October 1986

Reebie Associates, *Toward a More Effective Demurrage System*, Report No. FRA-OE-73-1, Federal Railroad Administration: Washington, D.C., 1972

Sloss, James and Carl D. Martland, *Government Intervention in Railroad Freight Car Hire: An Historical Perspective*, **Transportation Journal**, Summer 1984

Task Force I-2, **Freight Car Utilization: Definition, Evaluation, and Control**, Freight Car Utilization Program, Association of American Railroads: Washington D.C., 1980

Task Force on Reliability Studies, **Freight Car Utilization and Railroad Reliability: Case Studies**, AAR Freight Car Utilization Program, AAR Report R-442, Association of American Railroads: Washington D.C., 1980