Improving On-Time Performance for Long-Distance Passenger Trains Operating on Freight Routes

by Carl D. Martland

This paper discusses on-time performance (OTP) for long-distance passenger trains operating over tracks that are owned and operated by freight railroads. OTP is addressed primarily from the point of view of the host railroad. A brief literature review identifies practices that are commonly used by railroads and other modes to develop and implement achievable schedules. Analysis of travel time, train delay and other data for Amtrak trains operating on CSXT's I95 Corridor documents actual levels of reliability and the primary causes of poor OTP. Comparison of performance for passenger trains and various classes of freight trains demonstrates that Amtrak trains operate much faster and more reliably than CSXT's trains. Potential means of improving the OTP of Amtrak trains are discussed. While providing high quality track with sufficient capacity is the long-run solution for upgrading OTP, a short-run solution is to base schedules on past performance ("experience-based scheduling"). After Amtrak increased the schedule of the Auto Train by one hour in 2006, OTP improved from less than 10% in early 2006 to 82% for the first half of 2008. Analysis of the travel time distributions of the other long-distance Amtrak trains operating on CSXT's I95 Corridor from 2004 to 2008 indicates that a similar schedule increase would also have brought these other trains close to Amtrak's goal of 80% OTP. Schedules that reflect track maintenance requirements and other known seasonal and weekly factors would allow further improvements in measured OTP. Additional measures of performance concerning the probability and extent of late arrivals would be beneficial to travelers in planning their trips.

INTRODUCTION

This paper explores the on-time performance (OTP) of long-distance passenger trains operating over freight railroads. From the perspective of the freight railroads, the basic cause of poor OTP is that schedules for long-distance passenger trains are unrealistic. Having schedules that reflect actual operating conditions and capabilities is essential for achieving high levels of reliability. This paper shows that a small increase in the scheduled time of long-distance passenger trains could lead to a dramatic increase in the measured OTP, based upon analysis of data provided by CSXT.

In the United States, long-distance passenger rail services are operated by Amtrak over a rail network that is predominantly owned and used by freight railroad companies. Amtrak was created in 1971 to relieve the private railroads from the losses associated with passenger services (Wilner 1994). Except for the Northeast Corridor, which it owns, Amtrak must negotiate with the railroads for operating passenger services. The basic framework is that railroads give priority service to Amtrak trains, while Amtrak pays the railroads for the use of their track. Amtrak has never been profitable, and it relies on federal appropriations for investments in capacity and for operating subsidies. The appropriations process has at times been very difficult, and there have been numerous calls for restructuring, reducing or replacing Amtrak (Amtrak Reform Council 2001).

The high cost and poor reliability of long-distance trains are continuing problems for Amtrak. The two problems are linked, as the ability to provide reliable service depends upon having a track structure suitable for comfortable passenger service and a railroad that is willing and able to provide the capacity required for reliable passenger train operations.

Compensation for the use of private trackage has long been a matter of dispute among Amtrak and the railroads. Since Amtrak was created at a time when rail capacity far exceeded demand, the major costs imposed by passenger trains concerned the need to maintain the track structure for passenger trains that operated at higher speeds and required better track quality than most freight trains. Amtrak and the freight railroads debated whether or not the methodologies embedded in their contracts properly considered such things as the effects of speed and traffic volume on track maintenance costs (Resor and Smith 1993). By the mid-1990s, the excess capacity in the rail network had largely been eliminated, and the freight railroads were also concerned with costs related to capacity and train delay (Resor 1995). With more trains operating on fewer route-miles, the rail industry suffered from periods of congestion, resulting in service problems, customer complaints and eventually a congressional hearing (House of Representatives 2006). As the system approached capacity, the costs imposed by an additional train – especially a high-priority, high-speed train – became greater for the host railroads, and it became more difficult for the railroads to provide reliable service for long-distance passenger trains.

Train performance is one of the key concerns for rail passengers. In Europe, consistent OTP has been found to bolster rail market share (Steer Davies Gleave 2006). In the US, poor OTP is frequently cited as a short-coming for long-distance passenger trains, and improving the on-time performance of Amtrak trains is a continuing federal objective (Office of Management and Budget 2005). Acceptable levels of train performance are often specified in service contracts between passenger agencies and the host railroads, and Amtrak and other passenger agencies may offer incentive payments to host railroads based upon their ability to meet train schedules. OTP is a measure that has been highlighted in its annual reports since Amtrak was created.

APPROACHES TO SCHEDULING OF TRANSPORTATION SERVICES

There are two ways to ensure that train schedules are feasible. They can be established based upon past experience ("experience-based scheduling") or upon analysis of train capabilities and route characteristics. These two types of scheduling are applicable to two different types of operations, which White (2005) calls improvised and structured. Improvised operations are the norm in North America, while structured operations are common in much of the rest of the world. According to White, improvised scheduling works best when capacity is much greater than demand. Improvised operations are attractive on freight routes because it is often possible to enhance system performance by adjusting train make-up and departure times to reflect current conditions (Kwon 1994; Dong 1997). Structured operations are most necessary on routes that operate close to capacity and routes that are dominated by passenger trains, as it is important for these trains to operate very close to their schedules.

When most trains do not operate on or even close to schedules, past experience will be a more reliable indicator of achievability than the results of models or detailed planning for meets and passes. In the United States and Canada, many trains operate without fixed schedules, and those that have schedules seldom adhere to them. The standard deviation of the actual departure and arrival times of freight trains is typically measured in hours, not minutes (Martland 2008). With such wide variation from schedules, it is necessary to focus on dispatching as much as scheduling in order to manage train operations and performance. If schedules are based upon experience, the time allowed for movement along a route and the time required in terminals can be based upon the actual distribution of times required in the past. With this method, buffers are implicitly built into the schedules and the schedules are easily seen to be realistic. This method does not require careful assessment of meets and passes, because it is assumed that meets and passes that were feasible in the past will be feasible in the future, so long as traffic volumes are similar.

Simulations can be run to determine when growth in traffic will require investment in track or signals, and these simulations can take into account variability in traffic volume, maintenance, and other factors that affect running time (Van Dyke and Davis 1991). White (2005, p. 41) cautions that simulations "may not accurately represent the interaction between traffic and infrastructure" and concludes that "improvised operation still defies a completely accurate determination of the necessary infrastructure."

When operations are structured rather than improvised, sophisticated planning models can be used to develop schedules that take into account track configuration and condition, as well as train characteristics (Harker 1992; Ruffing 2003; Smith 1990). These models produce graphical results (string lines) that illustrate the paths that trains should take through the network. A carefully structured operating plan can specify times and locations for all meets and passes, and these times can be built into the train schedules (Ben-Khedher et al. 1998). If these models are to produce feasible schedules, then they must make reasonable and realistic assumptions about train speeds, the likelihood of delays, and the ability of trains to recover from delays. Allowing extra time in the train schedule will allow trains to make up time lost as a result of delays, whether those delays are related to equipment problems, track problems, congestion, weather, or a lack of crews or power.

How much recovery time to allow is a key issue in developing transportation schedules. Whether it is desirable or feasible to add buffers depends upon the nature of the operation. In rail transit operations, scheduling a few more seconds for each stop would help trains to stay on schedule, but increase the cycle time for the equipment and reduce the frequency of service (Lee 2002). Since transit commuters are very concerned with train frequency and are generally unaware of the schedules, transit operators are likely to minimize buffer time in order to maximize frequency of service and utilization of equipment on high density routes. If trains fall behind schedule, there will be time to recover at the end of rush hour.

Commuter rail operations are similar to transit operations in that there is time for the system to recover after rush hour. They are different in that trains operate less frequently, so that passengers are more likely to plan their trips based upon published schedules. Late arrivals will be noticed, and passengers may be upset. Adding a few minutes of recovery time to the schedule may enhance OTP, improve customer satisfaction, and only cause a minor change in perception regarding the length of the trip.

Airlines, like Amtrak, are under pressure from travelers and from public agencies to provide realistic schedules (Bowen and Headly 2005). In addition to allowing for the actual flight time, airline schedules must add time for moving to and from the gates and for take off and landing to get the equivalent of the "fastest possible time." They must also allow time for queues at both the originating and the terminating airports, as well as for air traffic control delays related to weather or other capacity problems. OTP has been shown to be a factor influencing demand for air travel; poor performance during one period tends to reduce demand during following periods (Susuki 1998). Hence, the time scheduled for a flight may be more than double the fastest possible travel time.

Railroads operating over long distances compete with airlines, but they do not have the option of adding airline-sized buffers to their schedule to ensure reliable operations. Even a buffer that is a small percentage of the total trip time can add hours to the schedule, possibly causing higher crew costs or changing arrival and departure times that were carefully planned to meet the needs of key market segments.

MEASURES OF TRAIN PERFORMANCE

Various measures are commonly used to measure the performance of passenger trains operating over freight rail systems, including OTP, average train speed, and average delays. This section provides some historical perspective on each of these measures, with some recent information concerning the performance of Amtrak trains operating on CSX routes.

On-Time Performance of Amtrak Trains

Amtrak has long sought, but seldom achieved OTP in excess of 80%. Table 1 shows OTP for Amtrak trains during the first 20 years of its existence. OTP bottomed out at 57% in 1978, reached a peak of 80% or more for the three years from 1982-84, then dropped to the mid-70s for most of the following 10 years. OTP was generally above 80% for the shorter distance trains and below 70% for the long distance trains. The period with best performance occurred at a time when freight train miles fell sharply from an average of more than 430 million per year in the late 1970s to a low of 345 million in 1982 and 1983 (AAR 1986, p.33). OTP was also less than the goal of 80% for the period from 1995 to 2005 (Table 2). Performance for short-distance trains was close to the goal, averaging 79% for this 11-year period, but performance for the long-distance trains exceeded 60% only in 1999.

Table 1: Amtrak's On-Time Performance: The First 20 Years (Percent of trains arriving on-time at their final destination)

Year	Total Trips	Trips Under 400 miles	Trips Over 400 miles
1972	75	82	53
1973	60	70	30
1974	75	80	63
1975	74	76	69
1976	62	66	48
1977	62	65	52
1978	57	61	48
1979	69	71	64
1980	77	77	64
1981	79	79	81
1982	82	81	82
1983	80	81	77
1984	81	82	78
1985	74	76	69
1986	74	78	62
1987	71	76	54
1988	75	81	54
1989	76	82	53
1990	77	82	59
1991	77	82	61
1992	72	79	47

Source: Frank Wilner (1994, p. 103)

Recent performance has been similar, as indicated by Table 3, which shows OTP for fiscal 2005, 2006, and 2007. The categories in Table 3 reflect a revised measurement system introduced by Amtrak in 2001. During each of these three years, OTP exceeded 78% for both the Northeast (NE) Corridor and for other corridor services. OTP was 65-70% for the other short-distance services, but only 30-42% for the long-distance services. Overall OTP was just below 70%.

Average Speed of Freight Trains and Passenger Trains

Average train speed is another measure that has long been used as an important indicator of train performance. Freight trains have, on average, seldom operated much faster than 20 mph. Figure 1, which presents data from the Depression years and World War II to the mid-1960s, shows freight train speeds slowly increasing from 15 to 20 mph. This figure also shows average passenger train speeds, which increased from about 35 to nearly 45 mph over this 40-year period. Passenger trains were generally given priority over freight trains and therefore were able to achieve higher average speeds.

Table 2: Amtrak's On-Time Performance: 1995-2005 (Percent of trains arriving on-time at their final destination)

Year	Total Trips	Trips Under 400 miles	Trips Over 400 miles
1995	76	81	57
1996	71	76	49
1997	74	79	53
1998	79	81	59
1999	79	80	61
2000	77	82	52
2001	75	78	55
2002	76	81	46
2003	74	78	50
2004	71	75	41
2005	70	74	42

Source: Bureau of Transportation Statistics

Table 3: Amtrak's Recent On-Time Performance: 2004-2008 (Percent of trains arriving on-time at their final destination)

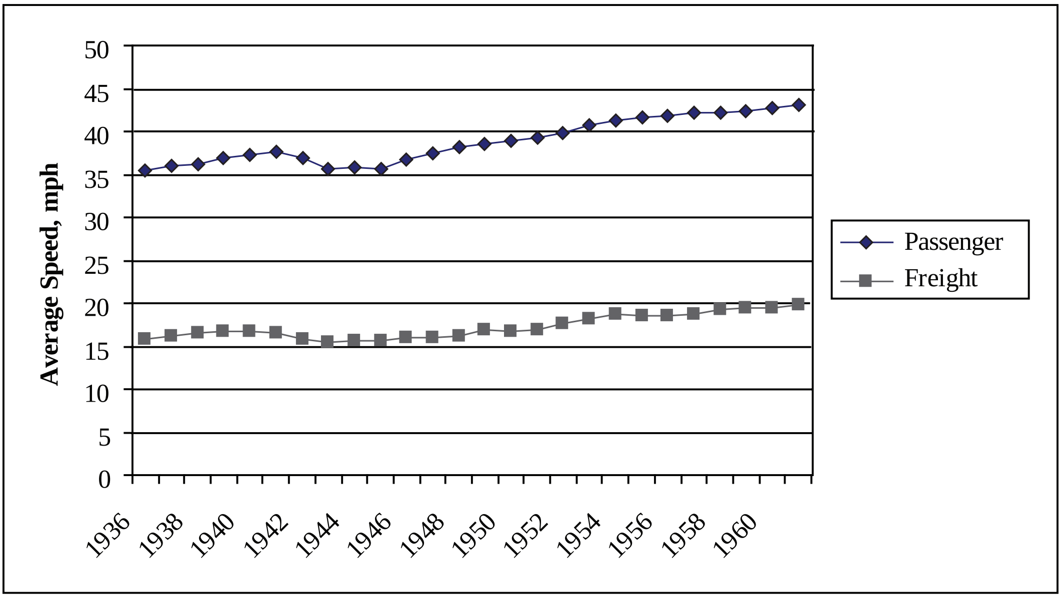
Amtrak Service	System	Northeast Corridor	Corridor	Short Distance	Long Distance
2004-05	69.8	78.2	78.1	70.4	41.4
2005-06	67.8	84.7	79.7	67.3	30.0
2006-07	68.6	87.8	80.3	66.6	41.6
October 2007 – June 2008	72.6	84.6	81.1	69.7	58.5

Source: Amtrak Host Railroad Performance Reports, September 2006, September 2007 and April 2008

Figure 2 shows the same measures for the 10 years following the creation of Amtrak. Freight train speeds continued at or near 20 mph, while passenger train speeds ranged from the low 50s at the beginning of the period to the mid-40s at the end of the period. Amtrak trains were able to operate much faster than freight trains because they continued to receive preferential treatment in dispatching.

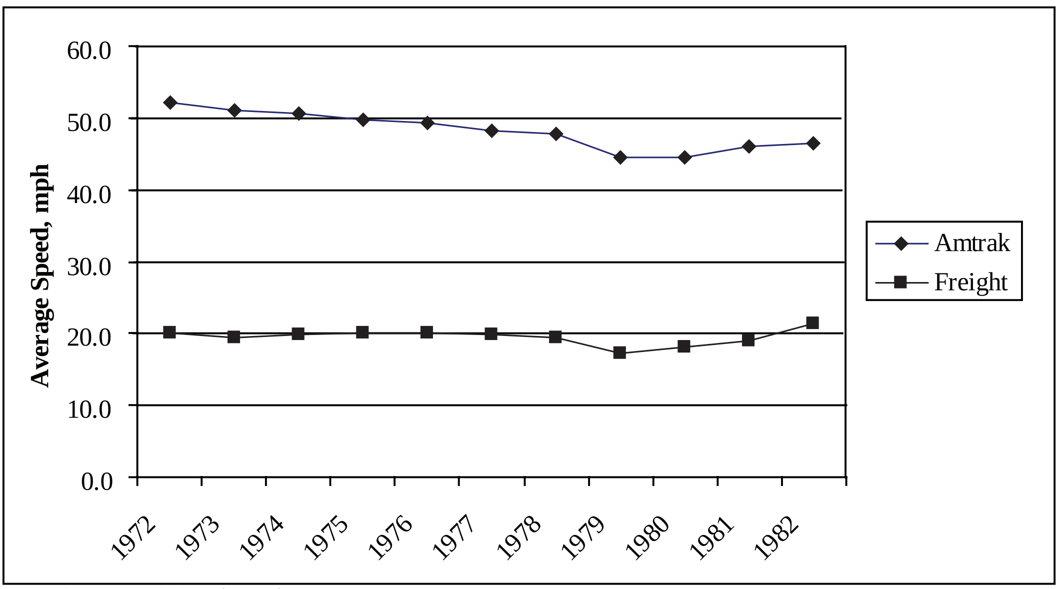
Average train speed is not used as a measure of performance for Amtrak trains. Amtrak does not emphasize – or even mention – train speed in its advertising or in its routine performance reports. The fact that passenger trains operate at much higher average speeds than freight trains indicates that

Figure 1: Average Train Speeds in the Pre-Amtrak Decades



Data Source: AAR (1962)

Figure 2: Average Train Speeds in the Decade Following the Creation of Amtrak



Data Source: AAR (1983)

passenger trains receive priority over freight trains, although it does not mean that passenger trains operate at a level anticipated by or acceptable to Amtrak.

Train Delays

OTP and average speed are just two of many ways of measuring train performance. Another approach to monitoring and managing train reliability is to focus on train delays. Delays can be categorized by type and ultimate cause of delay. Delays related to weather are different from delays related to accidents or delays related to maintenance or congestion. Unanticipated delays are likely to be a greater concern to travelers than delays that are known in advance.

Amtrak's Schedules and Measurement of On-Time Performance

In addition to the time required for unimpeded train operation and station stops, Amtrak's schedules include recovery time that is designed to allow the train to remain on schedule despite minor delays. Trains also have an allowance (30 minutes for long distance trains) that allows trains that arrive close to their scheduled arrival to be considered on time. The stated goal of 80% OTP allows the host railroads to achieve satisfactory OTP even if one train in five encounters serious delays. Amtrak schedules are generally not adjusted to reflect changes in traffic volumes, capacity, or maintenance activities on the host railroads. In some cases, Amtrak has adopted weekend schedules (Amtrak 2007). Amtrak monitors OTP and publishes various monthly reports showing OTP by train and by railroad.

CSXT analysis (CSXT 2006a) indicates that some portions of Amtrak's schedules are very close to the limits of operating performance. CSXT compared Amtrak's schedules to results from a train performance simulator for the four trains operating from Washington D.C. to Richmond. The simulation results indicated that the fastest possible run time for this route would be 2:02, based upon the average tonnage and horsepower of a typical Amtrak train. The scheduled times over this route allowed at most eight additional minutes (7% of the fastest possible run time) for recovery, and one train's schedule was actually seven minutes shorter than the simulated time. The average time for all Amtrak trains over this segment was 2:16 for the entire year, which was six minutes over the longest scheduled time for any of these trains, but just 11% over the fastest possible run time for the route.

OTP suffers because a wide variety of delays are encountered, as documented in Amtrak's monthly Host Railroad Performance Report. In this report, Amtrak defines a delay as anything that prevents a train from operating at its maximum allowed speed. Delays are separated into various categories that are attributed to the host railroad, to Amtrak, and to third parties. For example, in February 2007, the top causes of delay were as follows:

- Amtrak (19.6% of total delay minutes)
    - a. Engine failure (3.1%)
    - b. Passenger holds (2.9%)
    - c. Crew related delays (2.6%)
    - d. Other (10.9%)
- Host railroad (75.7%)
    - a. Freight train interference (24.4%)
    - b. Slow orders (16.9%)
    - c. Communications & Signal Work due to defect (10.6%)
    - d. Other (23.8%)
- Third Party (including weather) (4.7%)

Train delays do not automatically result in late trains, because the scheduled recovery time and the leeway included in the definition of on-time will be sufficient to cover minor delays. The vast majority of delays are related to freight train interference, track work and other host railroad problems.

VARIATIONS IN TRAIN PERFORMANCE BY CLASS OF TRAIN ON CSX

The speed and reliability of rail service is largely a result of the infrastructure that is available (e.g. single-track or multi-track), the track class and condition of the track (i.e. speed limits and slow orders), traffic volume, and dispatching policy. In single track territory, it is necessary to plan for meets and passes, and delays will be greatest for the lowest priority trains. As traffic volume approaches capacity, delays will increase, eventually affecting even the highest priority trains.

CSXT compared performance by class of train for three different operating environments in what they call their I95 corridor between Albany, New York, and Miami, Florida. The first segment, a 32-mile portion of the double-track route between Albany and New York City, is dominated by passenger trains (26 Amtrak trains, commuter service, and four or five freight trains per day). The second segment, 144 route-miles between Washington, D.C., and Richmond, Virginia, is a high-density double-track route with substantial volumes of both passenger and freight trains (18 Amtrak trains, commuter service along a portion of the route, and 25-30 freight trains). The third segment, 126 route-miles between Richmond and Rocky Mount, North Carolina, is a single-track segment dominated by freight trains (10 Amtrak trains, no commuter service, and 25-30 freight trains per day).

Not unexpectedly, the OTP was best for the first segment (82%), because of the high-capacity infrastructure and the minimal level of freight traffic. OTP was lowest for Amtrak on the third segment (40%), because of the difficulty of operating passenger trains on a single-track route that is operating close to capacity. OTP was 52% for the middle segment, which benefited from having double tracks, but also experienced the highest overall level of traffic of these three segments, especially where commuter trains shared the right-of-way with Amtrak and the freight trains.

CSXT documented the mean and standard deviation of travel times for the major classes of trains over each segment for the entire year of 2005 (Tables 4-6). The average speed and the coefficient of variation (the standard deviation divided by the mean time) are shown in order to make the results more comparable for routes of different lengths. Amtrak trains were always faster and more reliable than any category of the freight trains. The commuter trains, which made frequent stops, naturally could not match the average speed of Amtrak trains, but their travel times were less variable. Intermodal trains were faster and more reliable than merchandise trains, which in turn were generally faster and more reliable than the unit trains.

Table 4: Train Performance in 2005 for a 32-mile Segment Between Albany and New York City

Class of Train	Average Speed (mph)	Average Time	Standard Deviation	Coefficient of Variation
Amtrak	58	0:33	0:10	0.30
Merchandise	30	1:04	0:34	0.53
Unit	19	1:39	3:11	1.93

Source: CSXT (2006b)

Table 5: Train Performance in 2005, Washington D.C., to Richmond, Virginia

Class of Train	Average Speed (mph)	Average Time	Standard Deviation	Coefficient of Variation
Amtrak	47	2:16	0:26	0.19
Commuter (Washington to Fredericksburg)	34	1:35	0:16	0.17
Intermodal	32	4:48	1:01	0.21
Merchandise	21	6:46	3:01	0.45
Unit	21	6:48	2:51	0.42

Source: CSXT (2006b)

Table 6: Train Performance in 2005, Richmond to Rocky Mount, North Carolina

Class of Train	Average Speed (mph)	Average Time	Standard Deviation	Coefficient of Variation
Amtrak	46	2:45	0:33	0.20
Intermodal	33	3:47	1:12	0.33
Merchandise	15	7:48	4:10	0.53
Unit	18	6:27	3:58	0.61

Source: CSXT (2006b)

IMPROVING ON-TIME PERFORMANCE

This section is based upon discussions at a workshop sponsored by CSXT on "Passenger Operations on Freight Railroads" that was held in Tampa, Florida in June 2007. The workshop participants included officials from Amtrak, CSXT, BNSF, UP, and five commuter rail agencies. The discussions addressed various ways to measure and improve on-time performance (OTP), along with other issues of mutual interest such as maintenance planning and liability. The options for improving OTP were categorized as follows:

- Develop more achievable schedules
  - Experience-based scheduling: base schedules on past performance in order to ensure achievability
  - Use dispatching models to ensure that schedules are feasible
  - Ensure that sufficient time is allowed for loading and unloading trains
  - Adjust schedules for track maintenance programs
  - Schedule substantial recovery time at regular intervals (e.g. provide a 30 minute buffer every three hours for long-distance trains
- Increase operating discipline: ensure that all trains operate within their proper slots
- Improve coordination among inter-city, commuter, and freight operators
- Enhance the infrastructure
  - Short term: maintain or upgrade existing facilities so as to reduce slow orders
  - Long term: expand facilities so as to provide more capacity

The first category in this list includes various ways to improve the achievability of the schedules, beginning with experience-based scheduling. Using this method, the scheduled trip time for the coming period could be defined to be, say, the 75th or 80th percentile of the actual trip times for the previous period, and schedules could be adjusted periodically to reflect prior performance. Experience-based schedules should be achievable so long as there are no dramatic changes in traffic volume, no unusual and continuing weather-related disruptions, and no significant increases in the amount of infrastructure maintenance that is required. What was achievable in the past would remain achievable in the future, especially with annual or semi-annual adjustments in schedules. The underlying assumption is that train dispatching is – and would continue to be – effective, so that the highest priority trains would continue to get the best service, and actual service would not deteriorate if additional buffer time was added to the schedules. This is the approach that has been used by airlines to enable them to provide reasonable OTP despite the marked increase in congestion at major airports and along major flight paths. CSXT officials have hypothesized that "minor adjustments to the current schedules would improve performance, reliability, and customer satisfaction" (CSXT 2006a).

A similar approach would be to base schedules on models of operations. By using models, it is possible to take into consideration the effects of changes in traffic volumes, maintenance schedules, changes in the track structure, and other factors that affect train performance. With this approach, it is essential that the models be well-calibrated to ensure that the model results are in fact indicative of likely performance. As discussed above, the model-based approach is superior to the experience-based approach when operations are highly structured (as is the case for many commuter operations) or when major changes in traffic volumes or operating conditions are expected.

Commuter operators emphasized the importance of allowing enough time for loading and unloading trains, especially during rush hour. Since track maintenance and rehabilitation programs may disrupt normal operations over an extended period of time, it can make sense to adjust schedules on a temporary basis to reflect what is likely to happen. This is relevant both for commuter operations and for corridor or long-distance operations. The final suggestion was to take a new approach to scheduling long-distance passenger trains. By adding a large buffer every several hours, these trains would have a better chance to catch up if there are delays; the buffer could be added in locations where there are high volumes of passengers boarding so as to provide on-time departures for the highest number of people. This approach would increase OTP at the expense of higher crew costs and lower utilization of equipment.

The next two categories discussed at the workshop call for improved management, first within the railroad operating the passenger trains, and second among all of the agencies involved in scheduling and operating freight, commuter and inter-city passenger trains. Relating performance to schedules is useful for the operator and the transportation agency as well as for the travelers. The agency that provides the service may view the schedule as the product – a product that will be compared to the services offered by competitors. The schedule and various delay or OTP measures can be incorporated into the contracts between agencies and operators, who will have to negotiate a balance between shorter schedules and higher OTP.

The final category addresses what was felt to be the fundamental problem underlying poor OTP of passenger trains operating over freight railroads, namely the lack of sufficient capacity to handle actual traffic reliably. The long-run solution would be to expand capacity, while the short-run solution would be to adjust schedules to reflect the capabilities of the existing infrastructure. The ability to invest in capacity will depend upon the public and private benefits from, and commitment to the services provided along any particular route.

There was also discussion of other measures of train performance. If schedules are known to be unreliable, then travelers will need information about typical trip times in order to make plans that will not be disrupted by the normal variations in service. Travelers will also be concerned about the possibility of a major disruption, e.g. cancellation of all or a portion of the trip or the possibility of extremely long delays that will cause serious disruptions and inconvenience. Amtrak officials indicated that they are developing other measures to complement their measures of OTP and train delay, and they have an internal reporting system that measures the percentage of trains that arrive no more than 30, 60 or 90 minutes late.

PERFORMANCE OF AMTRAK TRAINS RUNNING ON CSX

This section uses data provided by CSX to illustrate two concepts. First, many different measures can be used to monitor the performance of long-distance passenger trains. Second, the historical performance of Amtrak trains can be used to estimate how much improvement in on-time performance (OTP) could be achieved by implementing experience-based scheduling. In 2007, adding time to the schedule for the Auto Train, in fact, achieved very significant improvements in performance. Adjusting schedules to provide additional time for predictable daily or seasonal factors, such as planned track maintenance, will also enhance OTP.

Performance of Long Distance Trains in the I95 Corridor

Table 7 compares the median and the 75th percentile of travel times for the long distance Amtrak trains operating in the I95 corridor for two recent 12-month periods. In each case, the 75th percentile was on the order of an hour longer than the median trip time, which gives an easily understood measure of the typical variability to plan for in using this service. The performance of these trains was stable, as illustrated by the last column, which shows the percentage change in the 75th percentile of the trip time distribution from one year to the next. For each train, this change was less than 2%.

Table 7: Performance of Long Distance Trains in the I95 Corridor
April 2005 to March 2006 vs. April 2006 to March 2007

Service	Train	Median Time 05/06 (Minutes)	Median Time 06/07 (Minutes)	75th Percentile 05/06 (Minutes)	75th Percentile 06/07 (Minutes)	Change in 75th Percentile
Auto Train	P052	1043	1026	1098	1085	-1.3%
	P053	1063	1043	1112	1096	-1.5%
Palmetto	P089	720	735	772	784	+1.6%
	P090	720	726	766	769	+0.3%
Silver Star	P091	1612	1586	1668	1651	-1.0%
	P092	1549	1564	1640	1635	-0.3%
Silver Meteor	P097	1341	1321	1397	1384	-1%
	P098	1358	1332	1406	1398	-0.5%

Data source: data provided by CSXT

For the three year period from April 2004 to March 2007, the Silver Star and the Silver Meteor generally had mean times and median times slightly better than their schedules, while the Palmetto generally had mean and median times slightly longer than its schedule. To get a normalized view of the variability of these trains, the difference between the median and the 75th percentile was expressed as a percentage of the median trip time. Since the median times were roughly the same as the scheduled trip time, the ratio indicates approximately how much longer the schedules would have to be to achieve OTP of 75%. The difference between the 75th percentile and the median averaged 4.3% of the median time for these long distance trains in 2004, 4.8% in 2005, and 5% between April 2006 and March 2007. This measure varied from as low as 2.8% (for the Silver Star operating from Washington to Miami in the winter of 2004) to a high of 8.4% (for the Silver Star operating from Miami to Washington during the summer 2005). The average for all four trains over the three-year period was less than 5%. Thus, an increase of about 5% (i.e. an hour) in the scheduled time for these trains would have been sufficient to bring OTP close to Amtrak's goal of 80% OTP.

The Auto Train Experience: Experience-Based Scheduling Improved OTP

The Auto Train, which operates between metropolitan Washington D.C., and Sanford, Florida (near Orlando), was of great interest to Amtrak and CSXT because its OTP was very low. To test whether experience-based schedules would in fact improve OTP (rather than lengthen actual travel times), Amtrak agreed to increase the scheduled trip time of the Auto Train by 6%, from 16.5 to 17.5 hours, effective April 1, 2007. This increase represented a large (57%) increase in the scheduled recovery time from 105 to 165 minutes or 16% of the new schedule. The impact on measured reliability was immediate, decisive, and continuing, as is evident from examining Amtrak's monthly Host Railroad Performance Reports. In April 2007, the average delays attributable to CSX as the host railroad for the Auto Train were less than the scheduled recovery time – the first time in more than 2.5 years that any CSX train had achieved this level of performance. OTP averaged 64% for the three months beginning in April 2007 compared to less than 7% for the same period in 2006. For the 12 months ending September 2007, OTP improved from 17% in 2006 to 62% in 2007. For the period from October 2007 to April 2008, OTP increased to 82%. In this case, a small increase in scheduled travel time did, in fact, lead to a dramatic improvement in OTP.

Seasonal Effects

Looking at performance for an entire year masks important seasonal differences in performance, notably the fact that rail track maintenance disrupts service primarily during the summer months from April to September. Table 8 shows seasonal variations in service for the Amtrak trains operating in the I95 corridor. All of these trains performed better in the winter than in the summer. The 75th percentile of trip times increased by 4% to 7% in the summer, while the absolute differences in the trip times for these trains ranged from about 30 to 90 minutes. The observed seasonal differences were more than double the less than 2% year-to-year variations in performance described above.

Table 8: Seasonal Variations in Performance for the Silver Services (minutes)

Service	Train	Median Time Winter	Median Time Summer	75th Percentile Winter	75th Percentile Summer	Change in 75th Percentile
Auto Train	P052	1002	1059	1044	1111	7%
	P053	1020	1064	1068	1111	4%
Palmetto	P089	711	752	760	806	6%
	P090	705	737	753	792	5%
Silver Star	P091	1554	1625	1600	1671	4%
	P092	1533	1604	1608	1697	5%
Silver Meteor	P097	1288	1357	1334	1417	6%
	P098	1299	1381	1342	1430	7%

Data source: data provided by CSXT

Day-to-Day Variations: the Palmetto

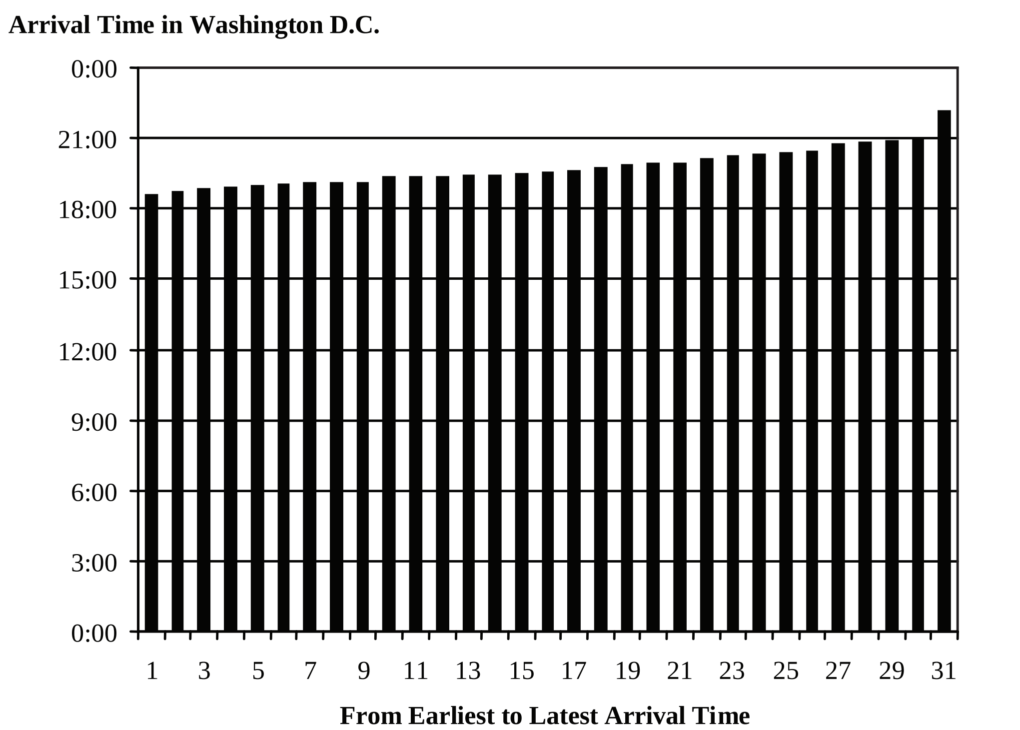
Variations in performance were measured in more detail for the Palmetto, which operates 829 miles between New York City and Savannah, Georgia. These trains move along the Northeast Corridor between New York and Washington D.C. and on CSX between Washington and Savannah. During the first half of 2006/07, the Palmetto's OTP of 43% was between that of the Silver Star (34%) and the Silver Meteor (63%). January 2007 was a typical month for the Palmetto: 26 trains arrived on time as defined by Amtrak, 19 trains arrived late, and 17 trains that were terminated before they reached their destination (in order to accommodate track rehabilitation, passengers were provided bus service on weekdays between Dillon and Savannah). For purposes of measuring OTP, southbound trains that were terminated at Dillon were reported as not completing their run, while northbound trains originating at Dillon were reported as being on time or late depending upon when they reached Washington D.C.

Figure 3 shows the arrival times of the northbound trains in Washington D.C., sorted from the earliest to the latest arrivals. The train almost always arrived between 6:30 and 8:30 P.M., and only once was delayed after 9 P.M. Table 9 shows various measures of performance for the northbound trains. OTP of 61% was above average for Amtrak's long distance trains, but somewhat below the goal of 80%. The 80th percentile of the arrival times was 8:21 P.M., 76 minutes later than the scheduled arrival time of 7:05 P.M.

Amtrak's measurement of OTP does not charge the host railroad for delays caused by Amtrak or third parties. Hence the OTP of 61% was greater than the 52% of trains that arrived less than 30 minutes late. OTP would have been greater than 68% if an hour had been added to the schedule, and adding 1.5 hours would have increased OTP to more than 84%. There was also a notable difference in OTP by day of the week: arrivals were nearly always on-time on Sunday and Monday, always late on Friday, and mixed on the other days. Adding time to the schedules for Tuesday through Thursday would be another way to achieve better OTP.

The data could be further broken down for the major O-D pairs served by this train. During the week of January 1, this train carried 1590 passengers. At least 100 passengers boarded the train at each of six stations: Savannah, GA; Charleston, SC; Kingstree Lane, SC; Florence, SC (Myrtle Beach); Fayetteville, NC; and Richmond, VA. Most (971) were destined to points beyond Washington, 278 disembarked at Washington, and 342 disembarked prior to Washington (including 80 at Richmond, VA, 73 in Alexandria, VA, and 63 in Wilson, NC). On-time arrival at Washington D.C. was therefore the primary concern of only about a quarter of the passengers, since the great majority were headed further north or had already left the train before it reached Washington. It is likely, however, that OTP at Washington was indicative of train performance along the way, while late arrivals at Washington would lead to late arrivals further along on the Northeast Corridor. OTP for long-distance trains is more an indicator of overall performance than something of direct interest to travelers.

Figure 3: Distribution of Arrival Times of the Palmetto in Washington D.C., January 2007



Source: CSXT analysis

SUMMARY AND CONCLUSIONS

There is a very long history of passenger and freight trains operating over the same routes. Records going back more than a half century show that passenger trains have been given priority so as to achieve much faster average speeds than freight trains. Following the creation of Amtrak, passenger trains were required to be given preference when operating on freight railroad lines, and the historical data confirm that Amtrak trains have, in fact, operated at much higher average speeds than freight trains.

Recent studies by CSXT went into considerable detail to demonstrate that passenger trains operating on the I95 corridor between New York and Florida have faster speeds and greater reliability than freight trains. These results suggest that Amtrak trains are indeed given preference over freight trains in this corridor. Moreover, the divergence in performance between passenger and freight trains seems similar to what persisted for decades prior to the creation of Amtrak.

While Amtrak trains operate faster and more reliably than freight trains, that does not mean that Amtrak trains are reliable or that their performance is acceptable to Amtrak or to travelers. According to measures used by Amtrak, average on-time performance (OTP) is very low for essentially all trains other than the premium services offered on the Northeast Corridor and Amtrak's other corridor services. The measure used by Amtrak to document performance is the percentage of trains that operate on time, where on time is defined as arriving within a specified time (30 minutes for the long distance service) of their scheduled arrival. Many of the long distance trains have OTP less than 50%, and some have OTP less than 20%.

The major factors that limit OTP for Amtrak trains operating in this corridor appear to be the capacity of the infrastructure, the density of freight and passenger operations, and the amount of time

Table 9: Various Measures of Performance for the Northbound Palmetto, January 2007

Scheduled arrival time	19:05
On-time within 30 minutes (Amtrak's measurement of host railroad performance)	61.3%
Average arrival time	19:45
Standard deviation of arrival time	0:47
Median arrival time	19:32
75th Percentile	20:17
80th Percentile	20:21
Scheduled trip time	11 hours, 5 minutes
Scheduled recovery time	58 minutes (8.7% of total scheduled trip time)
(80th Percentile-Scheduled Arrival)/Scheduled Trip Time	5.5%
% less than 30 minutes late	52%
% less than 60 minutes late	68%
% less than 90 minutes late	84%
% less than 120 minutes late	96%
% less than 2.5% late	53%
% less than 5% late	76%
% less than 7.5% late	91%
% less than 10% late	98%

Data sources: Schedules from Amtrak (2007); recovery time from Amtrak(a); arrival data from CSXT

devoted to track maintenance, rehabilitation, and expansion of capacity. The long-term solution to service problems would be expansion of the infrastructure, so as to reduce the need for meets and passes and to reduce the disruption caused by maintenance and other track work. However, such expansion will be feasible only if justified financially and politically. On-going programs for improving track quality and eliminating bottlenecks will enhance OTP by reducing slow orders, maintenance requirements, and congestion delays, thereby helping both freight and passenger operations. Better coordination among Amtrak, the host railroads, and commuter agencies will help improve operations in the most congested areas, especially when major track work is underway.

The most promising short-term solution for improving OTP is to adopt experienced-based schedules, i.e. to increase the recovery time that is included in the schedule. Because the recovery time included in Amtrak's schedules is small (8-16% for the trains operating in CSXT's I95 Corridor), a substantial increase in recovery time would have only a small increase in the scheduled time for long-distance trains while immediately producing a significant increase in OTP. After Amtrak added an hour to the Auto Train's schedule in 2007, OTP for the three month period April-June improved from less than 10% in 2006 to 64% in 2007 and 82% in the first half of 2008. Analysis of the trip time distributions of the other long-distance Amtrak trains operating in CSXT's I95 Corridor indicates that increases of about 5% in scheduled time would be sufficient to increase OTP close to Amtrak's goal of 80%.

OTP of Amtrak trains varied by an average of 5.5% between winter and summer, more than the year-to-year variations of less than 2% in CSXT's I95 Corridor. Measured OTP therefore could be improved by adjusting schedules semi-annually to reflect the effects of planned track work and other seasonal factors on expected travel times. In some cases, there may also be significant day-of-week variations in performance that should be reflected in the schedules.

It would be very useful to have additional measures of train performance available to potential riders and to the public at large. The OTP and delay measures currently used by Amtrak are designed for the contractual relationship between Amtrak and the host railroads or for Amtrak's reports to Congress. These measures can provide a useful means of understanding and then dealing with the most important problems, e.g. the most common causes of delays. It may be good management for Amtrak to resist relaxing the objectives included in contractual incentive clauses, and it may be worthwhile to preserve a long record of performance based on stable schedules and criteria. However, the details of contracts between Amtrak and host railroads and the clarity of historical performance records should not drive public and congressional perceptions of Amtrak's current service levels. Additional measures – or new ways of presenting the available information – would help Amtrak and the host railroads to present a better image of rail passenger service to prospective customers, the public at large, and to funding agencies.

It is not helpful to travelers to have schedules that are known to be unrealistic, and it is necessary to go beyond saying that OTP is low. Potential passengers would like to consider arrival and departure times, total trip time, reliability of trip times and many other factors when contemplating a trip on Amtrak. They need a reasonable estimate for the expected trip time, plus a reasonable estimate of the likelihood that the train will be late – and how late it might be. Amtrak should make additional measures available to their customers, such as the 80th percentile of the trip time distribution or the percentage of trains that are less than 30, 60, or 90 minutes late. Any of these measures could be normalized relative to the scheduled time or to the median time (e.g. the measure could indicate that the 80th percentile is 4% greater than or 50 minutes longer than the scheduled time). It would also be useful to passengers to provide measures of arrival time performance for all of the major city-pairs served by trains, not just for major trip segments and the final destination.

For experience-based scheduling to work effectively, it will be essential to ensure that revised schedules do not lead to lower priorities for Amtrak trains and a downward spiral in service. CSXT railroad officials maintained that their dispatchers indeed give Amtrak trains highest priority and would continue to do so if the schedules were revised. Amtrak officials and Amtrak riders are probably not convinced. Conceivably, contractual incentives or penalties could be devised to encourage host railroads not to allow service to decline. Monitoring performance before and after changes in schedules, as was done by CSXT and Amtrak with the Auto Train, would make it possible to identify situations where performance was slipping.

In situations where traffic volumes are growing faster than capacity, it is likely that Amtrak performance will suffer along with the performance of freight and commuter trains. A better understanding of what levels of performance are achievable at what cost would facilitate public-private partnerships aimed at boosting rail capacity and improving rail service.

Future research could help determine what aspects of performance are most related to the demand for and cost of long-distance rail service. Performance measures could then be devised that more clearly relate to the needs of customers, Amtrak and the host railroads. Having realistic measures of performance would also help the public debate concerning the future of long-distance intercity passenger service in the US by providing a more accurate and not necessarily so dismal a portrayal of current service levels of long-distance trains.

References

Association of American Railroads. *Railroad Transportation: A Statistical Record, 1921-1961*. Washington, D.C. 1962.

Association of American Railroads. *Statistics of Railroads of Class I in the United States, Years 1972 to 1982.* Washington, D.C. 1983.

Association of American Railroads. *Railroad Facts*. 1986.

Amtrak. *Host Railroad Performance Reports*. Published monthly on-line. (www.amtrak.com)

Amtrak. *Amtrak System Timetable*. Spring-Summer 2007.

Amtrak Reform Council. *Intercity Rail Passenger Service in America: Status, Problems, and Options for Reform. Executive Summary*. The Second Annual Report of the Amtrak Reform Council. Washington, D.C., 2001.

Ben-Khedher, Nejib, Josephine Kintanar, Cecile Queille, and William Stripling. "Schedule Optimization at SNCF: From Conception to Day of Departure." *Interfaces* 28 (1), (1998): 6-23.

Bowen, Brent and Dean Headley. *2005 Airline Quality Rating*. April 2005.

Bureau of Transportation Statistics. *Amtrak On-Time Performance Trends and Hours of Delay by Cause*. (www.bts.gov/cgi-bin/).

Bureau of Transportation Statistics. *Amtrak Trains Arriving On Time: 1995-2000 and 2001-2005* (www.bts.gov/cgi-bin/).

CSXT. *Amtrak On-Time Performance: An Evaluation of CSX's I-95 Corridor*. PowerPoint Presentation. March 27, 2006.

CSXT. *Passenger Train Performance on CSXT*. PowerPoint Presentation. August 30, 2006.

Dong, Yan. "Rail Freight Operations Under Different Operating Strategies." Ph.D. Dissertation. M.I.T. Department of Civil and Environmental Engineering, May 1997.

Harker, Patrick. "Use of Advanced Train Control Systems in Scheduling and Operating Railroads: Models, Algorithms, and Applications." *Transportation Research Record 1263*, (1990): 101-110.

Kwon, Oh-Kyoung. "Service Differentiation in Railroad Freight Service." Ph.D. Dissertation. M.I.T. Department of Civil and Environmental Engineering, May 1994.

Lee, Derek Sze-Ming. "Understanding Capacity and Performance of Urban Rail Transit Terminals." Master of Science in Transportation Thesis. M.I.T., June 2002.

Martland, Carl D. *"Railroad Train Delay and Network Reliability."* Association of American Railroads Research Report R-991. Transportation Technology Center, Inc. Pueblo, CO, 2008.

Office of Management and Budget. "Detailed Information on the Amtrak Assessment." 2005. (www.whitehouse.gov/omb/expectmore/detail/10004000.2005.html) (accessed 5/11/207)

Resor, Randolph. "Issues in Rail Passenger Service Costing and Operations." *Journal of the Transportation Research Forum* 35 (2), (1995): 44-58.

Resor, Randolph, and Michael E. Smith. "Track Maintenance Costs and Economies of Density: An Analysis of the Speed Factored Gross Tonnage Model." *Journal of the Transportation Research Forum* 33 (1), (1993): 24-41.

Ruffing, James A. "An Analysis of the Scheduling of Trains on High Density, Heavy Haul Networks." *Proceedings Implementation of Heavy Haul Technology for Network Efficiency, Specialist Technical Session*. International Heavy Haul Association. Virginia Beach, Virginia. 2003: 5.42-5.45

Smith, Michael. "Keeping Trains on Schedule: On-Line Planning Systems for the Advanced Railroad Electronics Sytstem (ARES)." *Journal of the Transportation Research Forum* 31 (1), (1990): 17-24.

Steer, Davies Gleave. *Air and Rail Competition and Complementarity*. Final Report for European Commission DG TREN. August, 2006 (www.steerdaviesgleave.com).

Suzuki, Yoshinori. "The Relationship Between On-Time Performance and Profit: An Analysis of U.S. Airline Data." *Journal of the Transportation Research Forum* 37 (3), (1995): 36-43.

U.S. House of Representatives. "The U.S. Rail Capacity Crunch." Hearing before the Subcommittee on Railroads, April 26, 2006.

Van Dyke, Carl D. and Llewellyn C. Davis, "Railroad Capacity Planning: A Case Study of the Beijing-Shanghai Rail Corridor." *Journal of the Transportation Research Forum* 32 (1), (1991): 86-102.

White, Thomas. "Alternatives for Railroad Traffic Simulation Analysis." *Transportation Research Record* 1916 (2005): 34-46.

Wilner, Frank. *The Amtrak Story*. Simmons-Boardman Books, Inc., Omaha, Nebraska, 1994.

Acknowledgements

This paper was prepared with support from CSX Transportation.

Carl D. Martland *is a transportation consultant in Sugar Hill, NH. He graduated from MIT with a B.S. in Mathematics in 1968 and the Civil Engineer degree in 1972. In July 2007, he retired from his position as senior research associate in the MIT Department of Civil and Environmental Engineering, where he had been actively engaged in transportation research and teaching since 1971. A specialist in rail transportation, Martland has studied service design, costing and control, equipment utilization, preventive maintenance, terminal operations, intermodal transportation, productivity, and technology assessment. He has taught undergraduate and graduate subjects at MIT, including freight transportation management, project evaluation, engineering system design, and transportation demand & economics. In 1997, Martland received the Distinguished Transportation Researcher Award from the Transportation Research Forum, recognizing his lifetime contributions to rail research.*