Office of Inspector General

ANALYSIS OF THE CAUSES OF AMTRAK TRAIN DELAYS

Federal Railroad Administration

Report Number: CR-2012-148

Date Issued: July 10, 2012

Memorandum

U.S. Department of Transportation
Office of the Secretary of Transportation
Office of Inspector General

Subject: INFORMATION: Analysis of the Causes of Amtrak Train Delays
Federal Railroad Administration
Report No. CR-2012-148

Date: July 10, 2012

From: Mitchell Behm
Assistant Inspector General for Rail, Maritime, and Economic Analysis

Reply to Attn. of: JA-50

To: Federal Railroad Administrator

Amtrak train delays outside the Northeast Corridor (NEC) reduce the value of Amtrak service as an option for travelers and increase the railroad's need for subsidies. Consequently, they have long been the subject of congressional concern and industry debate. Amtrak points to freight railroads' dispatching practices as the cause with the greatest impact on Amtrak train delays, while the freight railroads contend that capacity limitations, or insufficient infrastructure[1] for rail traffic levels, contribute more heavily.

In 2008, we issued two audit reports on Amtrak delays in response to requests from the Surface Transportation Subcommittee of the Senate Committee on Commerce, Science, and Transportation, and the Senate Appropriations Subcommittee on Transportation, Housing and Urban Development, and Related Agencies. In the first report, we concluded that delays outside the NEC substantially reduced Amtrak's net earnings.[2] In the second report, we identified causes of these delays but did not assess their relative importance.[3]

The impact of Amtrak delays took on a new significance with the passage of the Passenger Rail Investment and Improvement Act of 2008,[4] and the American Recovery and Reinvestment Act of 2009.[5] Together, these acts provided

---

[1] Rail infrastructure includes track, sidings, signals, switches, and yards.
[2] OIG Report Number CR-2008-047, "The Effects of Amtrak's Poor On-Time Performance," March 28, 2008. OIG reports are available on our Web site: www.oig.dot.gov.
[3] OIG Report Number CR-2008-076, "Root Causes of Amtrak Train Delays," September 8, 2008.
[4] P.L. No.110-432, Div. B.
[5] P.L. No. 111-5.

$12 billion to create and sustain a High Speed Intercity Passenger Rail (HSIPR) program. A substantial portion of the HSIPR program is geared towards improving the speed and reliability of existing Amtrak services. The Administration's current budget proposal would add $47 billion more to the HSIPR program over the next 6 years. Stakeholders responsible for determining the distribution of these funds need information on the causes of Amtrak service delays.

We conducted this analysis to determine significant causes of Amtrak delays and to quantify their relative impacts. In our analysis, we (1) identified statistically significant causes of delays—those whose impacts were greater than would be expected to occur by chance—outside the NEC and under the control of either Amtrak or a freight railroad; (2) assessed the degree of influence each cause had on Amtrak delays system-wide[6] and by individual route; and (3) examined delay determinants at locations of consistent, substantial delay.

To conduct our work, we constructed statistical models of Amtrak delays using data from Amtrak, the Surface Transportation Board's Carload Waybill Sample, and Oak Ridge National Laboratory's national rail network model.[7] Our sample covered fiscal years 2002 through 2007.[8] Since the marked reductions in freight traffic caused by the economic recession that began in late 2008 resulted in different usage patterns, we did not include post-fiscal year 2007 data in our analysis. However, our findings remain relevant given freight rail traffic's recent strong growth.[9] We analyzed on-track delays, which are the delays incurred by a train between the time it leaves one station and arrives at another, and did not analyze delays occurring in stations.[10] Furthermore, we analyzed delays with respect to pure runtime—the time it takes a train to travel between two stations in the absence of interruptions—rather than delays with respect to schedule. Schedule delays differ from delays with respect to pure runtime by the amount of schedule padding negotiated between Amtrak and the freight railroads, which can differ for reasons unrelated to underlying delay causes. Additional information on our scope and methodology is provided in Exhibits A and B.

---

6 System-wide delays were calculated as the average delays across all non-NEC routes.
7 We also used data from the National Climatic Center to control for some causes beyond stakeholder control.
8 Amtrak's fiscal year begins October 1.
9 The numerical estimates of different causes' delay contributions are specific to the sample period, but the results concerning the relative importance of different categories of causes remain relevant. For example, the estimated seconds of delay attributable to each host railroad holds only for the sample period, while the findings concerning the relative importance of host railroad effects to delays remains relevant.
10 On-track delays accounted for 77 percent of total delays in our sample, and in-station delays for the rest.

BACKGROUND

The Passenger Rail Investment and Improvement Act of 2008 directed the Secretary to establish a HSIPR program, and authorized approximately $4 billion over a 5-year period for HSIPR investments. Four months after that act's passage, the American Recovery and Reinvestment Act of 2009 appropriated another $8 billion for HSIPR to be allocated over a much shorter time span. Under its current budget proposal, the Administration would allocate $47 billion to HSIPR over the next 6 years. Responsibility for implementing the HSIPR program has been delegated to the Federal Railroad Administration (FRA). Under the program, State transportation departments, often in collaboration with freight railroads and Amtrak, apply to FRA for investment funds.

Outside the NEC, Amtrak trains operate almost exclusively on infrastructure owned and dispatched by freight railroads. Freight railroads' dispatchers are required by law[11] to give Amtrak's trains preference when freight trains and passenger trains operate on the same routes. However, Amtrak and the freight railroads disagree over whether preference is actually always given to Amtrak. Further, the relative importance of freight train interference in Amtrak delays is a particularly controversial issue for Amtrak and the freight railroads. In our 2008 study, we found that freight train interference was one of the primary causes of Amtrak delays, but that study relied entirely on data constructed from Amtrak conductors' observations. Amtrak's conductors have limited abilities to separate freight train interference from other delay causes. Hypothetically, if a conductor observes a freight train blocking his Amtrak train but cannot determine the reason for the freight train's failure to move, the conductor may record the cause of the train's delay as freight train interference, even though something else may be impeding the freight train's movement.

Amtrak services on routes outside the NEC generally exhibit worse on-time performance (OTP)[12] than Amtrak services on the NEC—where Amtrak largely owns the tracks and dispatches rail traffic. OTP on non-NEC long-distance routes is particularly poor. During the study period for this analysis, OTP on the NEC ranged between 75 and 90 percent, while on the 15 long-distance routes outside the NEC, it was consistently below 55 percent. During the same period, OTP on

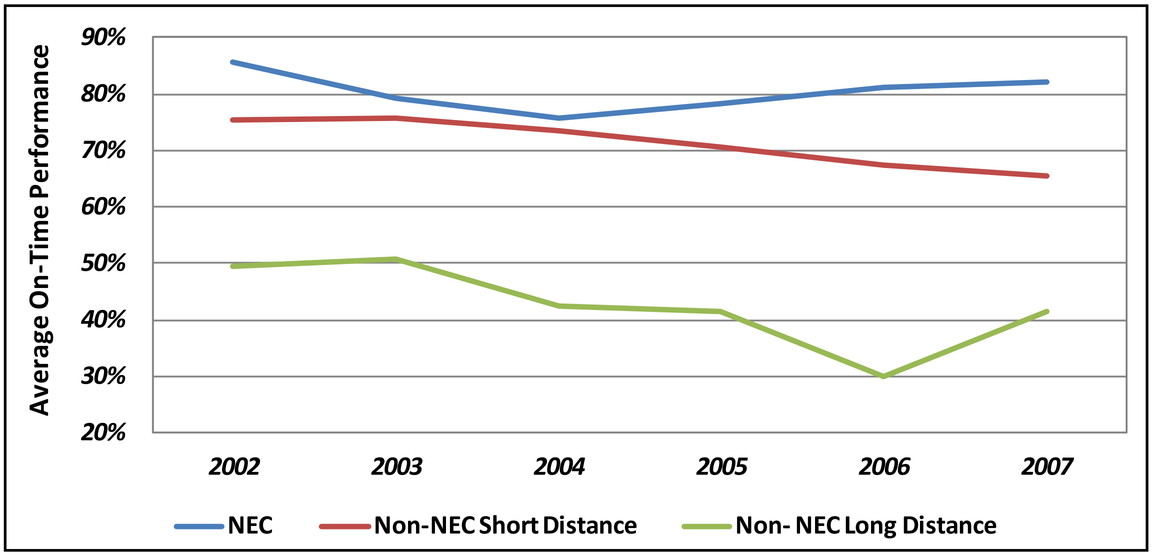
---

[11] 49 U.S.C. § 24308(c)

[12] An Amtrak train is considered on time if it arrives at its endpoint, or final destination, within a fixed number of minutes of its scheduled arrival time. The fixed number of minutes depends upon trip length as follows: arrival at an endpoint 10 minutes or less late of the scheduled arrival time is considered on time on trips of less than 250 miles; 15 minutes or less late is on time on trips between 250 and 350 miles; 20 minutes or less late is on time on routes between 350 and 450 miles; 25 minutes or less late is on time on routes between 450 and 550 miles; 30 minutes or less late is on time on routes of more than 550 miles. Trip length is the total distance traveled by a train from its origin to its final destination. Acela is held to a stricter standard. Even though its trip length exceeds 450 miles, Amtrak considers Acela on time only if it arrives at its endpoint within 10 minutes of its scheduled arrival.

Amtrak's 13 non-NEC short-distance routes varied from 65 to 76 percent. Figure 1 depicts OTP on the NEC, and on non-NEC long- and short-distance services.

Figure 1. Amtrak's Average Annual OTP for Fiscal Year 2002-Fiscal Year 2007



Source: Amtrak data

Delays also differ across routes, which differ in host railroad, congestion, track configuration, and terrain. For example, the average OTP for two long-distance routes, the Sunset Limited and the Empire Builder, is 14 percent and 72 percent, respectively.[13] Two short-distance routes, the Carolinian and the Hiawatha, show a similar range of average OTP at 30 percent and 92 percent, respectively. However, long-distance routes tend to exhibit greater variation in OTP than short-distance routes.

In the current study, we analyzed delays at the station-pair level to capture the causes of delays at their origins. We defined a station-pair as two Amtrak stations on a route—the first station being the origin and the second station being the destination—with no Amtrak stations in between.[14] On average, there were 30 station-pairs on each long-distance route and 20 station-pairs on each short-distance route.

[13] For the locations of Amtrak routes, see Exhibit D.

[14] In our analysis, we also differentiated between station-pairs by direction of travel and by Amtrak service.

RESULTS IN BRIEF

We identified six statistically significant causes of Amtrak delays:

- slow orders, which are speed restrictions imposed by host railroads along track segments in poor condition or undergoing repairs or improvement;[15]
- capacity utilization, which is a measure of congestion;
- host effects, or the effects of operating on track owned and dispatched by a particular host railroad or group of host railroads;[16]
- turn points, where Amtrak changes crews;
- late arrivals, or Amtrak trains operating well outside their scheduled time slot; and
- Amtrak mechanical problems.

Virtually the same causes were significant for both long-distance and short-distance services.[17]

The chief causes of delays system-wide were host effects and slow orders. However, delays caused by host effects differed considerably by host. For example, the difference between the largest and the smallest host railroad contributions to delays along a long-distance route, averaged across the Amtrak system, was 38 minutes and 30 seconds. Only slow orders may have produced sufficiently large delays to exceed the delays caused by the largest host effects.[18] Other factors, such as capacity utilization and activities at turn points, contributed significantly to delays system-wide, but considerably less than either host effects or slow orders. On most individual routes, as opposed to system-wide, slow orders caused the largest delays of any factors other than host effects.[19] Delays caused by different hosts varied even more widely across individual routes than they did on a system-wide basis.

At the locations with the highest average delays, slow orders, capacity utilization, and late arrivals were generally well above their system-wide averages. This suggests that all three causes contributed notably to delays at these locations.[20]

---

[15] Slow orders can also be imposed because of adverse weather conditions. However, we analyzed weather effects separately (see Exhibit B), and our results on slow orders relate only to infrastructure issues.

[16] These effects are broken out by specific host railroad or host railroad group in the body of the report.

[17] As defined by Amtrak, long-distance routes extend 550 miles or more and short-distance routes cover less than 550 miles.

[18] We could determine only minimum values for the delays caused by host effects system-wide. The delays caused by slow orders exceeded the minimum host effect values, but we could not determine how they would compare to the average values of host effects.

[19] On individual routes, we were unable to compare the size of delays caused by host effects versus other factors.

[20] We cannot test this hypothesis because our models, which were estimated on a system-wide basis, cannot generate meaningful delay impact estimates for individual station-pairs.

Host effects may have also played a role, but we cannot measure them at the individual station-pair level. Further analysis found that locations with the highest levels of slow orders also had some of the highest average delays, while locations with the highest levels of either capacity utilization or late arrivals did not have particularly large delays. This suggests that slow orders may have contributed more to delays at these station-pairs than either capacity utilization or late arrivals.

OUR MODELING PROCESS IDENTIFIED SIX STATISTICALLY SIGNIFICANT CAUSES OF DELAYS

In our modeling process, we identified six statistically significant causes of Amtrak train delays under stakeholder control. Four of these causes were under the control of a specific stakeholder: host railroads' effects, slow orders, Amtrak mechanical failures, and activities at Amtrak turn points. The other two—capacity utilization and late arrivals—were not. Capacity utilization is, in part, a function of capital investment, which can be carried out by the freight railroads alone or in partnership with the public sector. Late arrivals—or trains that arrive significantly late at their stations of departure—are out of their scheduled time slots and cause dispatching problems further down their routes. Whether or not an Amtrak train is out of its scheduled time slot arguably depends on the actions of both Amtrak and the host railroad. Virtually the same causes were significant on both long- and short-distance services. Table 1 details these causes.

Table 1. Statistically Significant Causes of Amtrak Train Delays

Cause [a]	**Definition**
Capacity Utilization	The amount of rail traffic divided by infrastructure capacity, which is the maximum amount of cars that can travel on a track segment based on the segment's characteristics. Congestion increases as capacity utilization increases.
Slow Orders	Speed restrictions imposed by the host railroad; we considered only those slow orders due to poor infrastructure conditions, and infrastructure repair and improvement work.
Amtrak Mechanical	Mechanical failures of Amtrak's locomotives and rolling stock.[b]
Late Arrival	The percentage of trains that arrived at least two hours late at the origin station.
Turn Point	A location at which a train's crew changes.
Host Effects	Host effects capture the impacts of such host railroad factors as dispatching practices, management capability, and business model.
BNSF Host Effect	The effect of operating on infrastructure owned and dispatched by Burlington Northern Santa Fe Railroad (BNSF).
CSXT Host Effect	The effect of operating on infrastructure owned and dispatched by CSX Transportation (CSXT).
NS Host Effect	The effect of operating on infrastructure owned and dispatched by Norfolk Southern Railroad (NS).
UP Host Effect	The effect of operating on infrastructure owned and dispatched by Union Pacific Railroad (UP).
CN Host Effect	The effect of operating on infrastructure owned and dispatched by Canadian National Railroad (CN).
Multiple Host Effect	The effect of operating on infrastructure owned and dispatched by multiple railroads between two stations.

[a] All causes were measured on a monthly basis at the station-pair level.

[b] Rolling stock includes all rail cars other than locomotives.

Unlike other causes in our model, we measured host effects relative to a baseline. Measurement of total host effects would require comparison with operations in the absence of any host effect, but this is impossible since trains always operate on some entity's infrastructure. Consequently, we measured delay caused by a host effect relative to a baseline of delay associated with operation on infrastructure owned and dispatched primarily by short-line and regional railroads.[21] Therefore, a host effect is the difference in delays attributed to operating on infrastructure

[21] Generally, Class III carriers are referred to as short lines, and Class II carriers are referred to as regional railroads. The four host railroads whose effects we traced are Class I railroads, which have annual revenues in excess of $250 million adjusted to 1991 dollars.

owned and dispatched by the particular host and delays from operating on infrastructure primarily owned and dispatched by regional and short-line railroads.

HOST EFFECTS AND SLOW ORDERS CAUSED THE LARGEST DELAYS SYSTEM-WIDE AND CONTRIBUTED SUBSTANTIALLY TO ROUTE LEVEL DELAYS

Host effects and slow orders were the chief causes of delays system-wide. However, delays caused by host effects varied considerably from one host railroad to another, for both long- and short-distance services system-wide. Only slow orders may have caused delays greater than those caused by the largest host effects. Capacity utilization and the activities at turn points also contributed notably to delays on both types of services, but capacity utilization had a relatively greater impact on short-distance services than on long-distance services. Amtrak mechanical problems contributed little to delays on either type of service. Delays caused by different hosts varied even more widely across individual routes than they did system-wide. Slow orders dominated other non-host effect causes in their contributions to individual route delays.

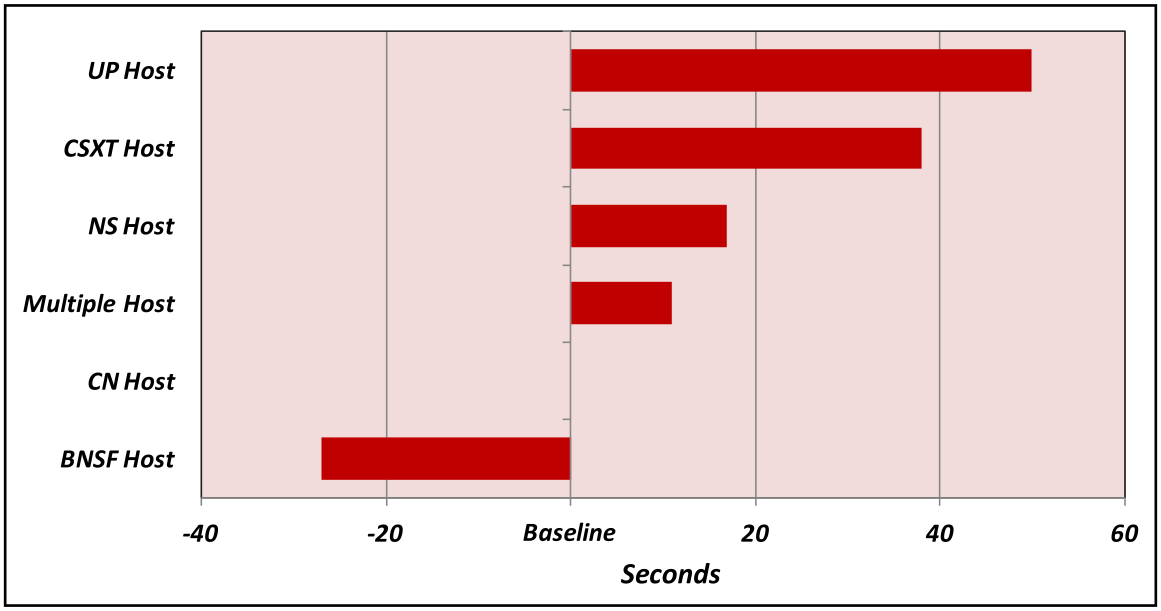
Delays Caused by Host Effects Differed Markedly Across Freight Railroads for Long-Distance Services System-wide

Delays differed substantially depending on which freight railroad owned and dispatched traffic over the infrastructure on which an Amtrak train operated. On average, the UP host effect caused 50 seconds more delay between each station-pair than the baseline railroad group.[22] In contrast, long-distance services on BNSF's rails experienced 27 seconds less delay between each station-pair on average than those operating on the baseline group's infrastructure. Over an entire long-distance route, the UP host effect added an average 25 minutes of delay more than the baseline (50 seconds times 30 stations), while the BNSF host effect reduced delays 13 minutes 30 seconds below the baseline. Figure 2 depicts the contributions of individual host effects to station-pair delays on long-distance services.[23] No bar appears for the CN host effect because it was statistically indistinguishable from the baseline host effect.

[22] We calculated each causal factor's contributions to delays as the product of the average value of the factor and its estimated effect. For example, to calculate slow orders' contribution to system-wide long-distance delays, we multiplied the system-wide average level of slow orders on long-distance routes times slow orders' average effect on delays between a station-pair on a long-distance service.

[23] Exhibit C shows the estimates underlying Figures 2-5, along with their confidence intervals.

Figure 2: Host Effects' Contributions to Station-Pair Delay on Long-Distance Services System-wide



Source: OIG

A minimum value for the baseline host effect can be determined for long-distance services system-wide. Host effect contributions to delays cannot be negative, since trains cannot travel faster than pure runtime. So the baseline host effect must be at least as large and positive as BNSF's host effect in relation to it is negative. At a minimum,[24] the baseline host effect causes 27 seconds of delay at each station-pair. Consequently, UP's host effect caused a minimum of 1 minute 17 seconds of delay (27 seconds minimum baseline value plus 50 seconds UP host delay relative to the baseline) between station-pairs. This translates into an average 38 minutes 30 seconds of delay over an entire long-distance route from traveling on UP controlled infrastructure. Similarly, CSXT, NS, and multi-host effects generated delays equal to the value of their impacts relative to the baseline plus at least 27 seconds, for total minimum contributions to station-pair delays of 65, 44, and 38 seconds, respectively.

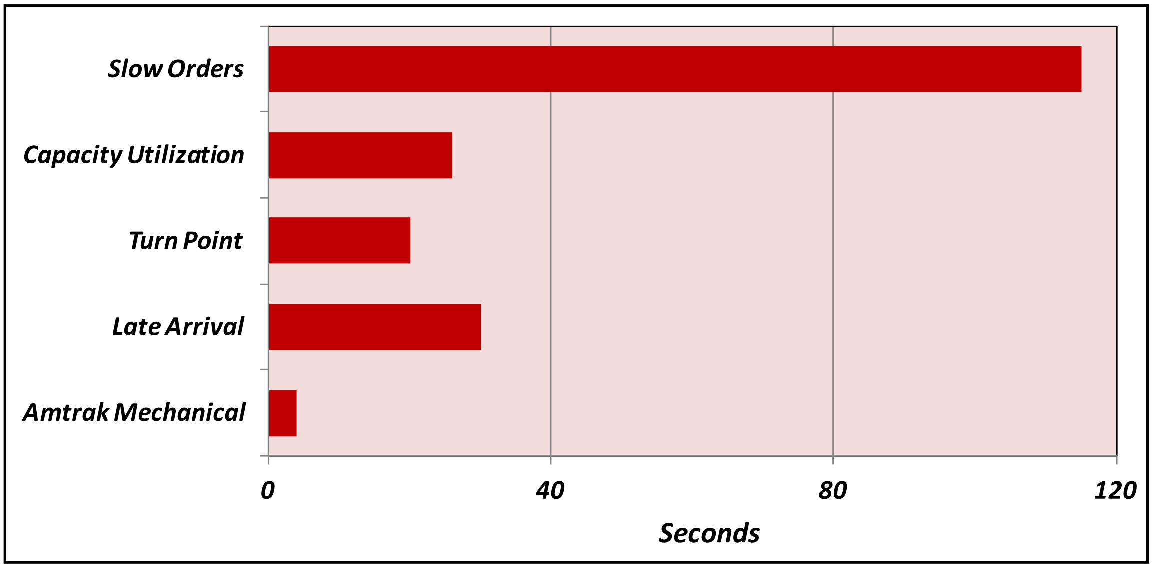
Only Slow Orders May Have Caused Longer Delays Than the Largest Host Effects Caused on Long-Distance Services System-wide

Slow orders caused the greatest delays among factors other than host effects on the long-distance services system-wide—1 minute 55 seconds on average between each station-pair—which represents 23 percent of the average long-distance

[24] This minimum assumes BNSF does not cause any increase in delays.

service station-pair delay of 8 minutes and 25 seconds. Over the entire length of a long-distance route, slow orders contributed 57 minutes 30 seconds of delay on average. No other cause contributed delays greater than the minimum values of most host effects. Figure 3 shows the contributions of all causes other than host effects. After slow orders, late arrivals made the next largest contributions to delays. On average, they added 30 seconds of delay between each station-pair, while capacity utilization and turn points contributed 26 seconds and 20 seconds, respectively. Amtrak mechanical problems contributed little to delays.

Figure 3: Non-Host Effect Causes' Contributions to Station-Pair Delay on Long-Distance Services System-wide

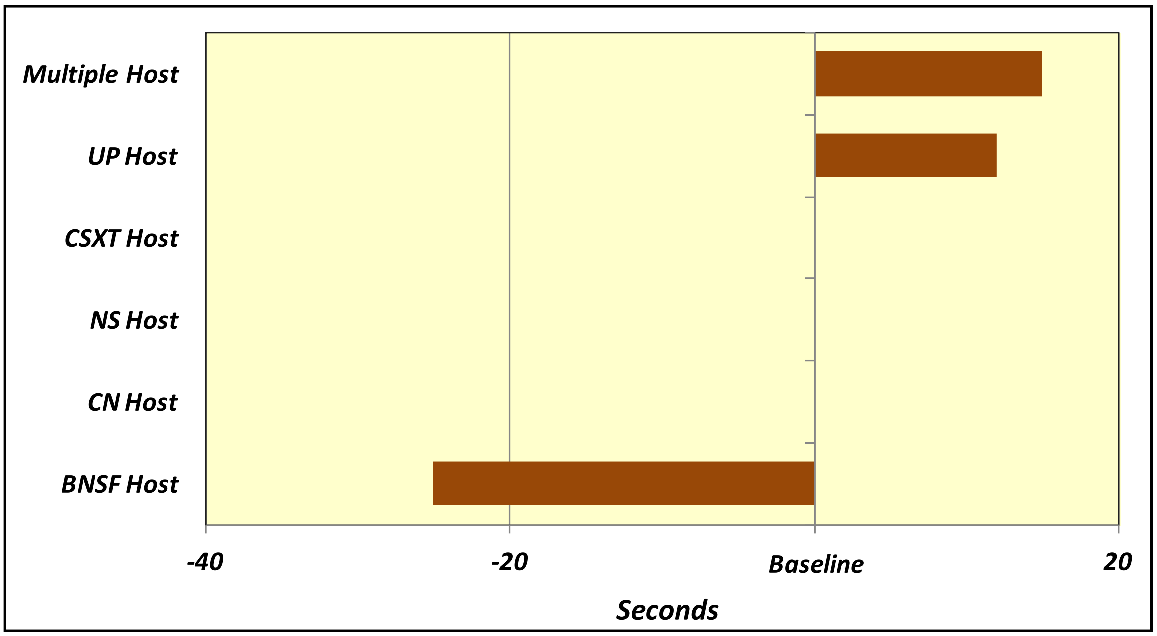


Source: OIG

Two Host Effects and Slow Orders Caused the Largest Delays on Short-Distance Services System-wide

Host effects on short-distance services system-wide also varied widely. As Figure 4 illustrates, the multiple host effect generated the largest delay above the baseline—15 seconds. The UP host effect added 12 seconds, the second largest delay contribution. CN, CSXT, and NS host effects were statistically indistinguishable from the baseline host effect. As with long-distance services, short-distance services operating on BNSF's rails experienced less delay than those operating on the baseline group's infrastructure—24 seconds less on average.

Figure 4: Host Effects' Contributions to Station-Pair Delay on Short-Distance Services System-wide

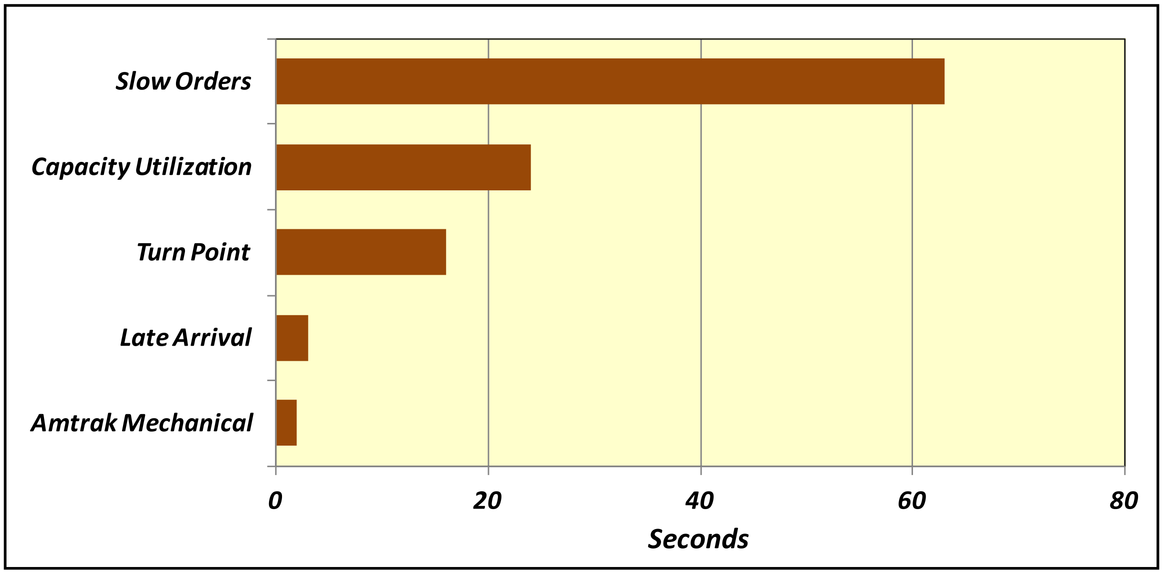


Source: OIG

Again, the baseline delay must be at least as large and positive as the BNSF host effect is large and negative. At a minimum then, the multiple host effect between a station-pair added 39 seconds to delay (24 seconds minimum baseline value plus 15 seconds multiple host delay relative to the baseline), and the UP host effect added 36 seconds. The minimum values of these two host effects exceeded the contributions of all non-host effect causes except slow orders.

Slow orders contributed 1 minute 3 seconds to the average total delay on short-distance services of 5 minutes 10 seconds, or approximately 20 percent. Figure 5 shows the contributions of non-host effect causes to delays on short-distance services system-wide. Capacity utilization played a relatively more prominent role in delays on short-distance services than on long-distance services, contributing 24 seconds between each station-pair, or 8 percent of the average total. Turn points increased delays between station-pairs by 16 seconds. Few trains arrived at their departure stations over two hours late on short-distance services, so late arrivals did not have a major impact on short-distance delays. As on long-distance services, Amtrak mechanical problems contributed little to delays on short-distances services.

Figure 5: Non-Host Effect Causes' Contributions to Station-Pair Delay on Short-Distance Services System-wide

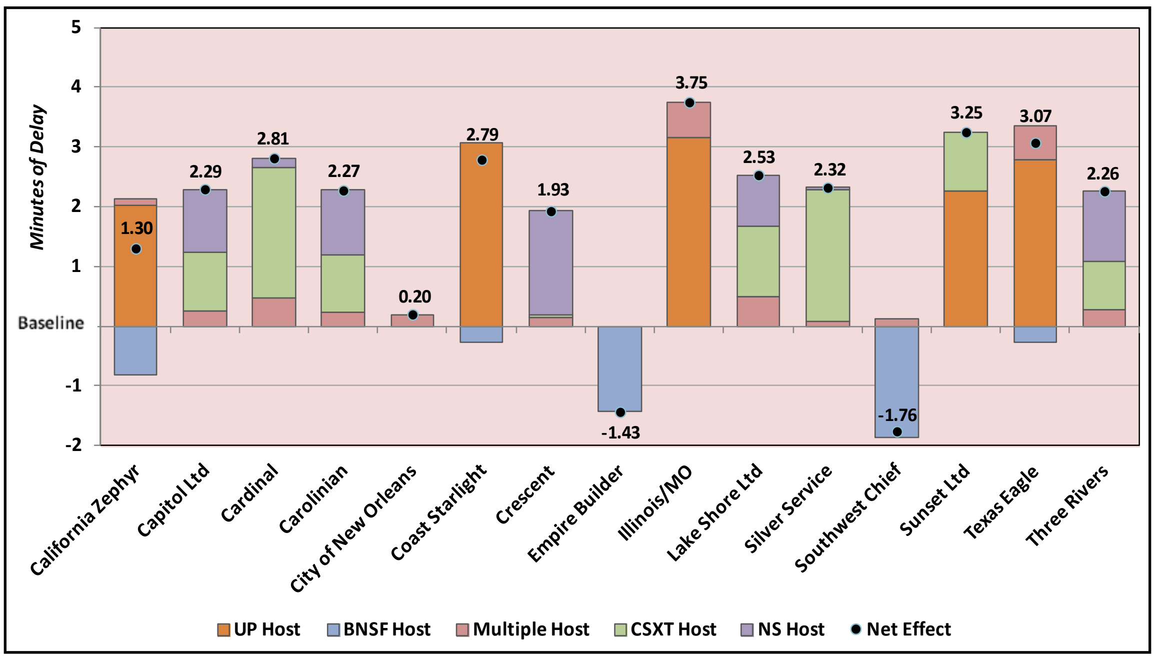


Source: OIG

Host Effects Had Widely Varying Impacts and Slow Orders Caused Substantial Delays Across Individual Long-Distance Routes

Host effects on individual long-distance routes varied greatly from one route to another, as Figure 6 shows.[25] The impact of host effects on individual routes depends on the proportion of the infrastructure owned by the host on that route. For example, the Illinois/Missouri Service travels on UP infrastructure for the majority of its length, and the UP host effect contributes 3 minutes 11 seconds to station-pair delays on that route. In contrast, the UP host effect's contribution to long-distance service delays system-wide captures UP's impact only on UP's portion of the entire long-distance system's infrastructure. Because we generated route level results using the system-wide model, we cannot determine a minimum value for the baseline for individual routes.

Figure 6: Host Effects' Contributions to Station-Pair Delay on Individual Long-Distance Routes

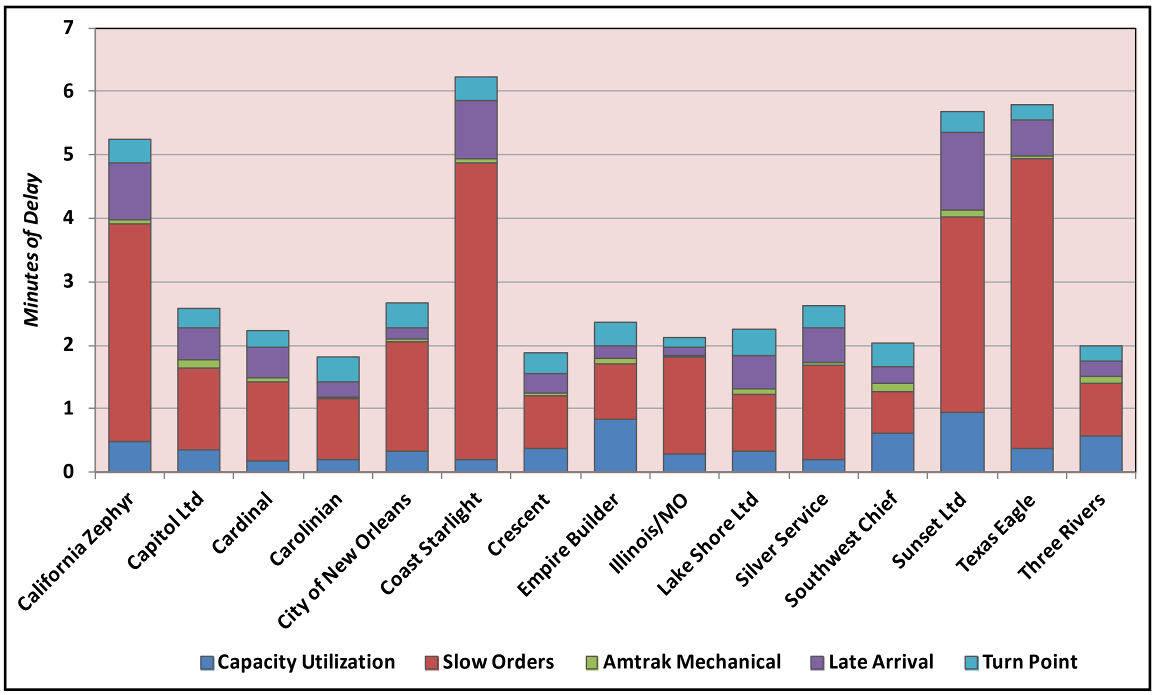


Source: OIG

[25] Exhibit E shows the estimates underlying Figures 6-7.

Among non-host effect causes, slow orders produced the largest delays on most individual long-distance routes. However, the amount of delay due to slow orders varied considerably across routes. For example, slow orders added 4 minutes 34 seconds to station-pair delays on the Texas Eagle, but only 39 seconds on the Southwest Chief. Figure 7 depicts the causes' relative contributions. Late arrivals increased delays on many of the long-distance routes, but most notably on the Sunset Limited, Coast Starlight, and California Zephyr.

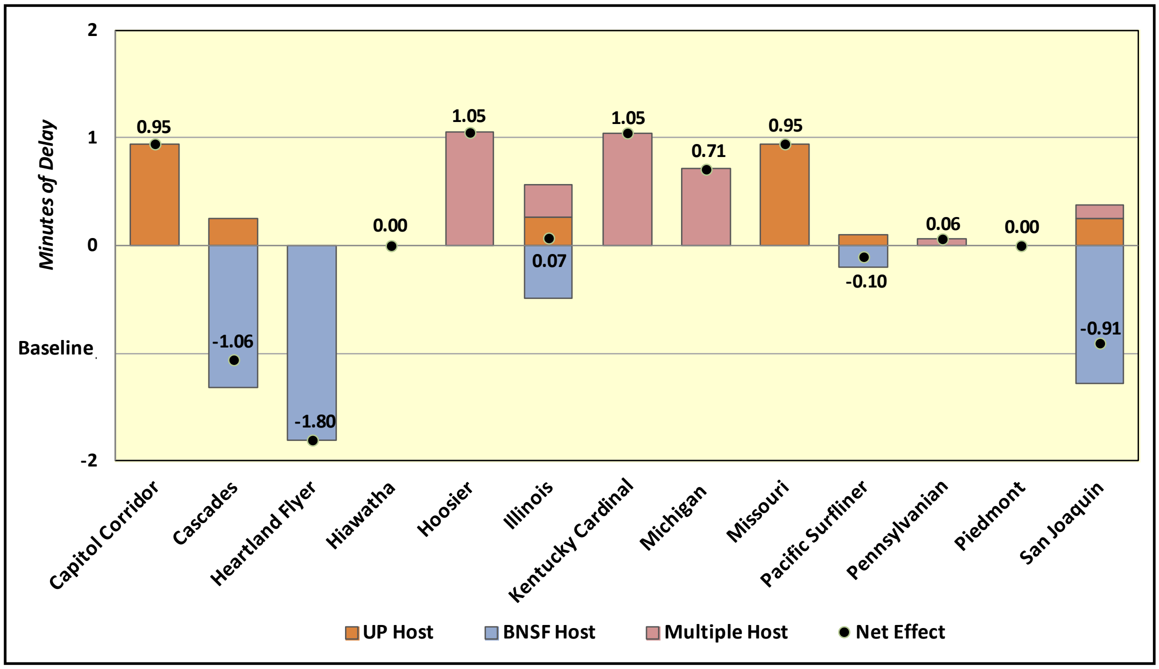
Figure 7: Non-Host Effect Causes' Contributions to Station-Pair Delay on Individual Long-Distance Routes



Source: OIG

As depicted in Figure 8, the multiple host effect caused sizable delays on several individual short-distance routes, while the UP host effect caused substantial delays on the Capitol Corridor and Missouri routes.[26] Short-distance routes hosted at least in part by BNSF experienced reductions in delays relative to the baseline system-wide host effect. Among these routes, the Cascades, Heartland Flyer, and San Joaquin experienced particularly substantial reductions in relation to the baseline.

Figure 8: Host Effects' Contributions to Station-Pair Delay on Individual Short-Distance Routes

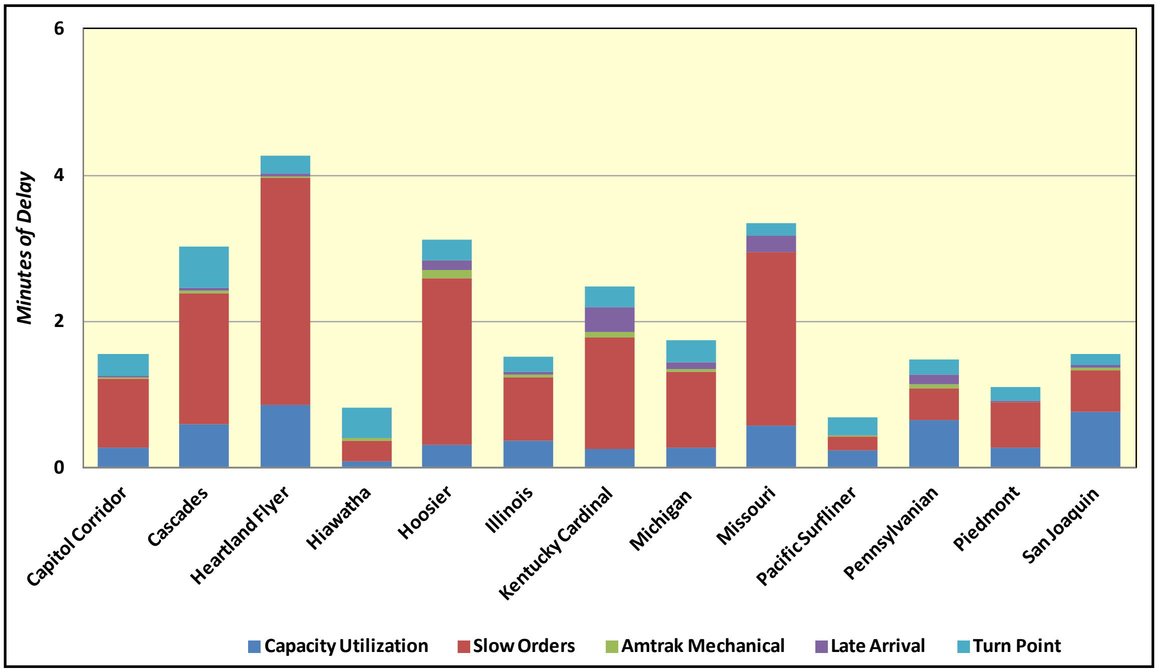


Source: OIG

[26] Exhibit F shows the estimates underlying Figures 8-9.

Slow orders caused large delays, but not always the largest of non-host effect causes on individual short-distance routes. For example, capacity utilization contributed more to delays on the Pacific Surfliner, Pennsylvanian and San Joaquin routes than slow orders, as Figure 9 indicates. Late arrivals had less impact on short-distance routes than on long-distance routes because far fewer trains ran so far outside of their scheduled time slots over the shorter distances.

Figure 9: Non-Host Effect Causes' Contributions to Station-Pair Delay on Individual Short-Distance Routes



Source: OIG

SEVERAL DELAY CAUSES HAD ELEVATED VALUES AT STATION-PAIRS WITH THE HIGHEST AVERAGE DELAYS, BUT SLOW ORDERS MAY HAVE BEEN THE MOST IMPORTANT

Slow orders, capacity utilization, and late arrivals at the locations with the highest average delays were all generally above their system-wide averages—suggesting that all three causes contributed notably to these delays. We cannot test this hypothesis, however, because our system-wide models cannot generate delay impact estimates at individual station-pairs.[27] Furthermore, without using our models, we cannot measure host effects. However, further analysis of the raw data reveals that locations with the highest levels of slow orders also had some of the highest average delays, while locations with the highest levels of either capacity

---

[27] Use of a system-wide model at the station-pair level would involve errors beyond acceptable levels.

utilization or late arrivals did not have particularly large delays. These data suggest that slow orders may have primarily driven the largest station-pair delays.

Slow orders, capacity utilization, and late arrivals were all generally above their system-wide averages at the station-pairs with the highest average delays, as illustrated in Table 2 for a selection of those station-pairs.[28] This fact suggests that these causes all contributed notably to delays at these locations.

Table 2. Delay Causes' Average Values at Selected Station-Pairs with High Average Delay

Station-Pair	Route	Average Delay (minutes)	Slow Orders (minutes)	Capacity Utilization (%)	Late Arrival (%)
Long-Distance Services					
San Antonio – Houston	Sunset Ltd.	48.4	15.3	52.5	42.6
Alpine-El Paso	Sunset Ltd.	46.0	14.0	49.3	44.6
St. Louis-Poplar Bluff	Texas Eagle	43.2	37.9	2.8	3.6
Winnemucca-Sparks	California Zephyr	41.1	28.0	41.4	53.7
Salt Lake City-Elko	California Zephyr	40.0	19.6	7.9	22.0
System-wide		***8.4***	***2.1***	***15.2***	***17.5***
Short-Distance Services					
Vancouver-Bellingham	Cascades	29.4	5.5	6.8	0.0
Springfield-Alton	Illinois	25.6	6.0	6.4	1.7
Sedalia-Jefferson City	Missouri	24.4	1.9	14.7	3.0
Lee's Summit-Warrensburg	Missouri	21.1	1.5	12.7	1.1
Pittsburgh-Alliance	Pennsylvanian	19.1	3.3	12.2	2.2
System-wide		***5.2***	***0.9***	***10.1***	***1.5***

Source: OIG

[28] We chose to display a selection from among the station-pairs with the highest average delays to avoid primarily showing data for station-pairs that differed only in direction, and to provide a sampling from different routes. Similar concerns determined our selections of station-pairs for Tables 3-4.

We examined average delays at the station-pairs with the highest levels of slow orders. Delays at these station-pairs matched the magnitudes of the delays at the station-pairs with the highest average delays. Table 3 illustrates this for selected station-pairs.

Table 3. Delay Causes' Average Values at Selected Station-Pairs with High Average Slow Orders

Station-Pair	Route	Average Delay (minutes)	Slow Orders (minutes)	Capacity Utilization (%)	Late Arrival (%)
Long-Distance Services					
St. Louis- Poplar Bluff	Texas Eagle	43.2	37.9	2.8	3.6
Elko-Winnemucca	California Zephyr	41.0	33.9	34.5	31.0
Salt Lake City-Elko	California Zephyr	40.0	19.6	7.9	22.0
Chico-Sacramento	Coast Starlight	27.9	19.4	6.5	44.3
Klamath Falls-Dunsmuir	Coast Starlight	31.9	17.8	10.9	25.9
System-wide		***8.4***	***2.1***	***15.2***	***17.5***
Short-Distance Services					
Hermann-Jefferson City	Missouri	16.1	6.1	24.2	4.2
Springfield-Alton	Illinois	25.6	6.0	6.4	1.7
Gainesville-Fort Worth	Heartland Flyer	18.8	5.7	40.6	0.8
Vancouver-Bellingham	Cascades	29.4	5.5	6.8	0.0
Rensselaer-Dyer	Hoosier	10.4	5.1	2.9	3.7
System-wide		***5.2***	***0.9***	***10.1***	***1.5***

Source: OIG

By comparison, we did not find an association between the highest values of capacity utilization and late arrival and the highest station-pair delays. Table 4 illustrates this with the average delays at station-pairs selected from among those with the highest levels of capacity utilization. Taken together, these results suggest that slow orders alone could have driven the bulk of the delays at the station-pairs with the highest average delays

Table 4. Delay Causes' Average Values at Selected Station-Pairs with High Average Capacity Utilization

Station-Pair	Route	Average Delay (minutes)	Slow Orders (minutes)	Capacity Utilization (%)	Late Arrival (%)
Long-Distance Services					
Pomona-Ontario	Sunset Ltd.	4.7	0.2	95.0	5.5
Deming-Lordsburg	Sunset Ltd.	12.0	0.6	87.7	51.1
Tucson-Maricopa	Sunset Ltd.	18.1	0.3	81.0	49.5
Yuma-North Palm Springs	Sunset Ltd.	28.4	2.9	77.0	59.7
Glasgow-Wolf Point	Empire Builder	2.2	0.6	62.7	8.4
System-wide		***8.4***	***2.1***	***15.2***	***17.5***
Short-Distance Services					
Martinez-Antioch	San Joaquin	1.3	1.2	68.5	0.2
Gainesville-Fort Worth	Heartland Flyer	18.8	5.7	40.6	0.8
Wasco-Corcoran	San Joaquin	3.5	0.5	33.7	0.2
Fresno-Madera	San Joaquin	1.9	0.6	29.8	0.5
Seattle-Edmonds	Cascades	1.2	0.4	28.3	0.3
System-wide		***5.2***	***0.9***	***10.1***	***1.5***

Source: OIG

CONCLUDING OBSERVATIONS

Train delays substantially increase Amtrak's need for subsidies and reduce the value of intercity passenger rail as an option for travelers. Under the HSIPR program, substantial investment funds have been made available to improve passenger rail's speed and reliability. Stakeholders making and overseeing investments in rail infrastructure involving public funds need to be aware of causes of the delays that they are intending to reduce as part of the decision making process. Of key concern is the need to differentiate those factors in passenger train delays that may not require capital investment to address, such as host effects and slow orders, from those factors that require investment. Making this distinction is critical to ensure publically funded rail investments are targeted to locations where capacity constraints are the primary cause of delays, and not expended on locations where other delay causes predominate.

#

cc: Audit Liaison, OST, M-1
Audit Liaison, FRA, RAD-43
Amtrak Liaison

EXHIBIT A. SCOPE AND METHODOLOGY

We conducted this analysis to identify the primary causes of delays and determine their relative impacts. In our analysis, we: (1) identified statistically significant causes of delay—those whose impacts were greater than would be expected to occur by chance—outside the NEC and under the control of a stakeholder, either Amtrak or a freight railroad; (2) assessed the degree of influence each cause had on Amtrak's delays system-wide and by individual route; and (3) examined delay determinants at locations of consistent, substantial delay.

We built our sample from Amtrak's 28 non-NEC passenger routes, 15 of which were long-distance routes—or longer than 550 miles—and 13 of which were short-distance. The sample period included fiscal years 2002 through 2007. We compiled 1,117 unique, non-overlapping Amtrak origin and destination station-pairs. The station-pairs differed by direction and route because delay times can differ according to these parameters. The unit of observation in our data was a unique combination of origin station, destination station, route, and month.

We designed and estimated delay models to explain the variations in delay of Amtrak's trains. The dependent variable—the one we wished to explain—was the average minutes of on-track delay at the station-pair level. The independent variables—those used to explain the variations in delays—included precipitation; temperature; capacity utilization; slow orders; mechanical delays; late arrivals; turn points—locations at which crews changed; distance between stations; host effects; meetings between Amtrak and various freights to address delay issues; and fiscal year dummy variables.

We divided our sample into long-distance and short-distance routes when estimating our delay models. The models for these two types of routes were the same except that we excluded the meeting variables from the short-distance model because these meetings were unrelated to short-distance routes. The decision to estimate separate models was based upon the different characteristics associated with these services as well as test results. For example, outside of the NEC, short-distance routes are served by shorter, lighter trains that can travel at higher speeds than those on long-distance routes.

We generated our measure of delay with respect to pure runtime (PRT). PRT is the time required to make a trip in the absence of any interference. Our delay variable was the difference between the time a train departed an origin station and the time it arrived at its destination station minus PRT. PRT delays differ from delays with respect to scheduled travel time, which incorporates schedule padding introduced to accommodate expected delays and as a result of negotiations between Amtrak and the host railroads.

Capacity utilization is traffic volume divided by railroad infrastructure capacity. We used the maximum value of the calculated capacity utilization rate across all track segments between a station-pair. Oak Ridge National Lab (ORNL) developed the capacity values for each infrastructure segment based on four different factors: number of tracks, type of signal system, frequency of passing sidings, and terrain grade.

Traffic volume is defined as the total number of rail cars—locomotives, Amtrak rail cars, commuter rail cars, and freight rail cars—that pass over a segment of track in a given month. We obtained raw data on daily freight traffic from the Surface Transportation Board's Carload Waybill Sample.[1] ORNL used the Transportation Routing Analysis Geographical Information System (TRAGIS), a national railroad routing simulation program, to assign routes to all freight traffic identified in the Waybill Sample. We determined Amtrak train frequency using Amtrak timetables and the number of cars on each train from Amtrak's consist books.[2] We obtained information on commuter train frequency and number of cars per train from 10 commuter agencies.

We captured host effects using dummy variables for the host railroads that own and dispatch the associated infrastructure. The host railroads for each of our 1,117 station-pairs remained constant in our sample.[3] Amtrak provided the data on freight railroad infrastructure ownership. We created seven host railroad categories for the station-pairs in our sample: one category for each of five Class I railroads—BSNF, UP, CSXT, NS, and CN; one category for multiple hosts between a single station-pair; and one category for other, largely short-line and regional, railroads.

We chose the Hausman-Taylor estimator for our sample after a series of tests. The Hausman-Taylor estimator allowed us to address potential measurement error associated with the capacity utilization variable and to obtain coefficients for our time-invariant variables, such as host effects. The measurement error with capacity utilization arises because of the assumptions we had to make during the construction of the variable. These assumptions include no change in capacity along Amtrak routes during our study period and the use of the maximum capacity utilization value for each station-pair.

[1] The Waybill Sample provided origin, destination, and, in some cases, intermediate points for each trip sampled.

[2] Amtrak's consist books detail the numbers and types of rail cars making up each train on each Amtrak route.

[3] Station-pairs which changed host railroad accounted for less than three percent of our total observations, and we excluded them from our analysis.

EXHIBIT B. DETAILED SCOPE AND METHODOLOGY

This exhibit provides a detailed discussion of our scope and methodology broken into the following sections: data, model, estimation, and sensitivity analyses.

DATA

Our sample includes nearly all of Amtrak's non-NEC corridor passenger routes from Amtrak fiscal year 2002 through 2007.[1] The data covers 28 Amtrak long-distance—defined as longer than 550 miles—and short-distance routes.[2] We compiled 1,117 unique, non-overlapping Amtrak origin and destination station-pairs from these routes. The unit of observation was a unique combination of origin station, destination station, route, and month.[3] Altogether, we had 64,684 unique observations.

Station-pairs serviced by different routes were treated as separate observations. Many instances in which multiple routes served the same station-pair involved both short- and long-distance routes. Our treating these instances separately allowed for the possibility that short-distance and long-distance trains were affected differently by factors such as host railroad dispatching practices. Ninety-three percent of the 1,117 station-pairs were served by a single route. Approximately six percent were shared by two routes. The remaining station-pairs were served by three to six routes.

The station-pairs in our model also differed by direction. We used directional markets because delay times can differ according to travel direction. For example, on the Texas Eagle, which runs between Chicago, Illinois and San Antonio, Texas, northbound trains consistently take 100 minutes longer, on average, than southbound trains to reach their final destination.

---

[1] We excluded the following NEC routes from our sample: Acela Express, Metroliner, Regional, Empire Service, Keystone, Vermonter and Downeaster. We also excluded one non-NEC service, the Auto Train, which travels from Lorton, Virginia to Sanford, Florida with no intermediate station stops.

[2] Long distance routes include the California Zephyr, City of New Orleans, Coast Starlight, Empire Builder, Cardinal, Lake Shore Limited, Silver Service, Southwest Chief, Sunset Limited, Texas Eagle, Capitol Limited, Crescent, Three Rivers and Illinois/Missouri. Short-distance routes include the Heartland Flyer, Hoosier State, Illinois, Pacific Surfliner, San Joaquin, Capitol Corridor, Carolinian, Cascades, Hiawatha, Michigan, Missouri, Pennsylvanian, Piedmont, and Kentucky Cardinal.

[3] We have unbalanced panel data in that we did not have the full 72 months worth of data for all 1,117 station-pairs. The minimum number of observations for any station-pair included in the sample was 12 months. At least 68 months of data was available for half of the station-pairs.

Measure of Delay

We generated our measure of delay with respect to pure runtime (PRT).[4] PRT is the time required to make a trip in the absence of any interference. Our delay measure is the difference between the time a train departs an origin station and the time it arrives at its destination station minus PRT. Note that PRT delays differ from delays with respect to scheduled travel time. Scheduled travel time incorporates schedule padding, which may be introduced to accommodate expected delays or as a consequence of negotiations with the host railroad. We initially analyzed delays with respect to scheduled travel time but found that padding hid important determinants of delay. In addition, current year padding appeared to be related to previous year delays, and consequently introduced endogeneity issues. The remainder of this report solely considers delays with respect to PRT.

Though we do not present the results here, we also examined delays occurring in stations. We found that on average, on-track delays accounted for 77 percent of total delay minutes, while station delays accounted for 23 percent. However, while there are scheduled terminal wait times, there is no analogue for PRT that would correspond to time spent in a terminal. Consequently, we could not construct a model for station delays analogous to our model of on-track delays with respect to PRT. All delays discussed in this report are on-track delays.

Daily delays were aggregated into monthly averages by station-pair. To avoid unrepresentative monthly averages, we eliminated station-pairs that were missing more than two-thirds of their expected observations.[5] In addition, we eliminated outliers that were obviously the result of data entry error. Some delays were long enough to have compelled the cancellation of the associated trains, but did not, and others were sufficiently negative so as to be physically impossible. Both types of extreme delays fell within the top and bottom one percent of daily delay data, which we removed from our sample. Figure B1 displays a histogram for average minutes of delay for the data used in the analysis. The distribution has a longer tail on the right, but the tail thins out dramatically as delays increase.[6] We calculated the skewness and kurtosis to compare the degree of deviation from a normal

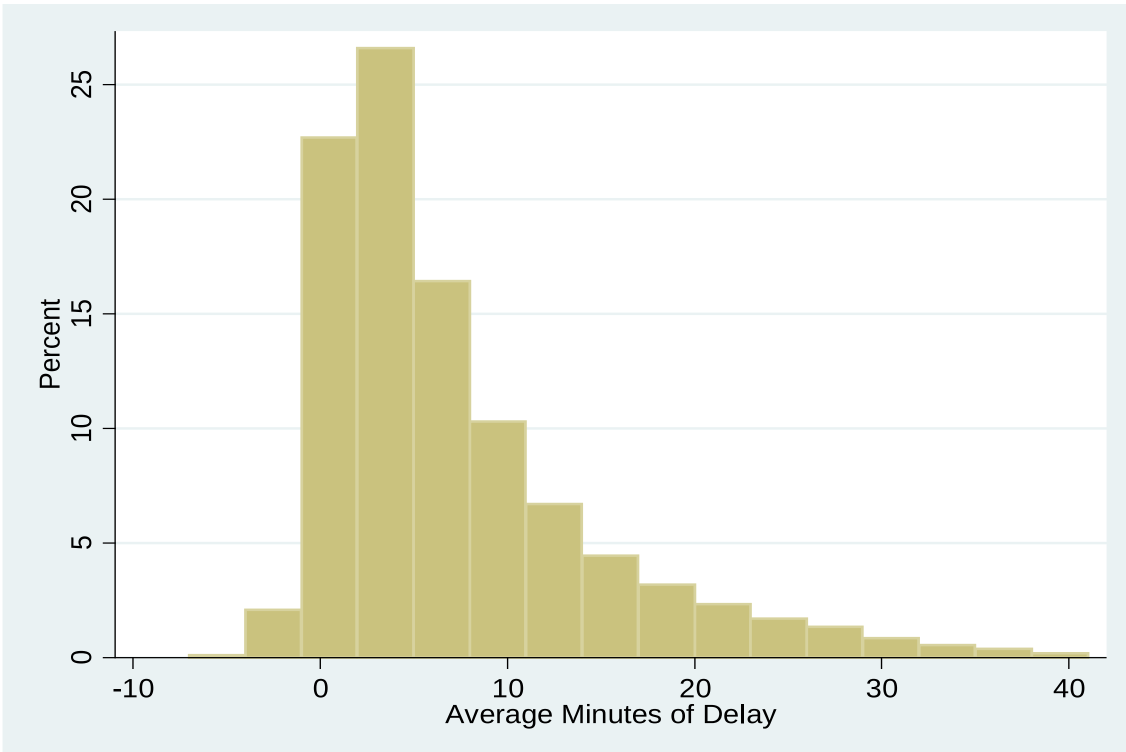
[4] Amtrak provided PRT data and data on arrival and departure times for each station at a daily frequency at the individual train level. We removed data on temporary trains—those operating less than once a week. Nearly all routes are serviced at least once a day.

[5] According to Amtrak, missing observations most likely occur as the result of station staff's failure to record trains' entry and exit times. Amtrak is currently automating the recording process for all stations.

[6] The method used to measure PRT creates considerable opportunity for noise around zero minutes of delay. Specifically, PRT represents a time agreed upon by representatives of both Amtrak and the relevant host railroad as they ride a route and gauge the minimum possible travel time. In addition, PRT changes infrequently; short-term travel time adjustments are usually made only in scheduled recovery time and miscellaneous adjustments rather than in PRT itself.

distribution. The skewness was 2 and kurtosis 10, as compared with 0 and 3 for the normal distribution.

Figure B1: Minutes of Delay Histogram



Note: Figure B1 captures 99 percent of observations; the remainder have delays in excess of 40 minutes.
Source: OIG

Host Effects

We captured host railroad effects—host effects—using dummy variables for the individual host railroads who own and dispatch the associated infrastructure. Amtrak provided the data on railroad infrastructure ownership. We identified seven host railroad categories associated with the station-pairs in our sample: one category for each of five Class I railroads—Burlington Northern and Santa Fe (BSNF), Union Pacific (UP), CSX Transportation (CSXT), Norfolk Southern (NS) and the Canadian National Railway Company (CN); one category for multiple hosts between a single station-pair; and one category for other, primarily short-line and regional, railroads.

The host railroads for our 1,117 station-pairs remained constant throughout the sample period. Three host railroads divided ownership of roughly three-fourths of the track miles fairly evenly among them. Track miles owned by multiple hosts accounted for 2 percent of total track miles, and those owned by the "other" railroad group accounted for 4 percent. We eliminated the station-pairs between which Amtrak hosts the tracks because there were so few of them.

Capacity Utilization Rate

Our capacity utilization measure has four components: the capacity of the railroad infrastructure, freight traffic, Amtrak traffic, and commuter rail traffic. Traffic is defined as the total number of cars—locomotives, passenger cars, and freight cars—that pass over a segment of track in a given month. The Transportation Routing Analysis Geographical Information System (TRAGIS), a national railroad system model and routing simulation program developed by Oak Ridge National Laboratories (ORNL), provides information on track characteristics (number of tracks, type of signal system, frequency of passing sidings) for each segment of intermediate track between each station-pair. ORNL used this information in combination with data on terrain grade to develop capacity values for each segment based on a method designed by Clark (1995). ORNL determined it was reasonable to assume that capacity did not change along Amtrak routes over our sample period.

We obtained raw data on daily freight traffic from the Surface Transportation Board's Carload Waybill Sample.[7] ORNL used TRAGIS to assign routes to all freight traffic identified in the Waybill Sample, and translated daily traffic levels into monthly averages for each track segment. We determined Amtrak train frequency using Amtrak timetables, and the number of cars on each train from Amtrak consist books. We obtained information on commuter train frequency and number of cars per train from the 10 relevant commuter agencies.

Monthly freight, Amtrak, and commuter traffic estimates were divided by the number of days in each month to get average daily traffic to match daily capacity. We calculated the capacity utilization rate for each segment of track by totaling the daily averages of freight, Amtrak, and commuter rail traffic and dividing the sum by segment capacity. The maximum value of the calculated capacity utilization rate across all track segments between a station-pair was used to represent segment capacity in the analysis.

MODEL

We constructed the following model to measure delays on long-distance routes. The dependent variable is ON-TRACK DELAY$_{ijt}$. It is the average on-track

[7] The Waybill Sample provided origin, destination, and, in some cases, some intermediate points for each trip sampled.

minutes of delay with respect to PRT for all trains that operate at station-pair *i*, route *j*, in month *t*. The rest of the model is as follows:

$$
\begin{aligned}
\text{ON-TRACK DELAY}_{ijt} = \; & \alpha_0 + \beta_1 \text{ PRECIP}_{ijt} + \beta_2 \text{ TEMP}_{ijt} + \gamma_1 \text{ CAP UTIL}_{ijt} \\
& + \gamma_2 \text{ SLOW ORDER}_{ijt} + \gamma_3 \text{ MECH}_{ijt} + \gamma_4 \text{ LATE ARR}_{ijt} \\
& + \delta_1 \text{ DISTANCE}_i + \delta_2 \text{ TURN}_i + \eta_1 \text{ MEET UP}_{\text{jt}} \\
& + \eta_2 \text{ MEET NS}_{\text{jt}} + \theta_1 \text{ BNSF}_i + \theta_2 \text{CN}_i + \theta_3 \text{ CSXT}_i \\
& + \theta_4 \text{ NS}_i + \theta_5 \text{ UP}_i + \theta_6 \text{ MULTI HOST}_i \\
& + \sum_{l=2}^{7} \lambda_l \text{ FISCAL YEAR}_{lt}
\end{aligned}
$$

where *i* indexes the station-pair, *j* the route, and *t* the month. The Greek letters represent the coefficients which quantify the effect of each variable on Amtrak delays. β's control for weather related delays. γ's represent non-weather related delay determinants with both cross-sectional and time series variation. δ's are associated with time invariant variables. η's represent the effects of meetings between Amtrak and host railroads to address delays, θ's are dummy variables for individual host railroads, and λ's capture fiscal year-specific effects.

This model was also used for short-distance routes with MEET UP_{jt} and MEET NS_{jt} excluded, because the meetings were unrelated to short-distance routes. MEET UP_{jt} equals one if route *j* is California Zephyr and if *t* is later than April 2007, the date UP and Amtrak met to discuss addressing delays on that route. MEET NS_{jt} equals one if route *j* is either Capitol Limited or Lake Shore Limited and *t* is later than June of 2007, marking a meeting between NS and Amtrak to address delays on those routes.

We generated seven host railroad dummy variables to capture the effects of individual railroads and groups of railroads: UP_i, BNSF_i, CSXT_i, NS_i, CN_i, MULTI HOST_i, and Other RR_i. UP_i takes the value 1 if Union Pacific is the host railroad for the O&D station-pair *i* and 0 otherwise. The rest of the individual host dummies are defined in the same fashion. The definitions for MULTI HOST_i and Other RR_i are slightly different. MULTI HOST_i takes the value 1 if there is more than a single host railroad for a given station-pair. Other RR_i represents the group of other, primarily short line and regional, freight railroads.

Other RR_i is the reference group and its dummy is deleted from the estimation. The coefficients on the remaining host dummies measure the difference in delays relative to Other RR_i. These dummy variables capture the time invariant characteristics specific to each host railroad. They can include, but are not limited to, host business model, managerial ability, and dispatching practices.

CAP $UTIL_{ijt}$ is the maximum monthly level of the capacity utilization rate between station-pair *i*, on route *j*, in month *t*. It is the maximum ratio of traffic volume to capacity, both measured in carloads, observed between the station-pair. An increase in the level of capacity utilization is expected to result in an increase in train delays. Therefore, we expect a positive coefficient on CAP $UTIL_{ijt}$.

To capture variation in minutes of delay due to weather, we include precipitation ($PRECIP_{ijt}$) and temperature ($TEMP_{ijt}$).[8] $PRECIP_{ijt}$ is the percent of days during which at least one inch of rain fell at station-pair *i* in month *t*. $TEMP_{ijt}$ is the percentage of days with temperatures exceeding 90 degrees for station-pair *i* in month *t*. The weather variable for each station-pair was measured at the weather station closest to the mid-point between the station-pair. Trains operate at reduced speed during heavy rains. Also, most freight railroads have heat order policies that reduce train speeds when the temperature exceeds certain thresholds.[9]

SLOW $ORDER_{ijt}$ is the average minutes of slow order delays from all the trains in our sample operating at station-pair *i*, on route *j*, in month *t*. A slow order is a local speed restriction issued by a host railroad that requires trains to travel at less than a track segment's normal speed limit. Several triggers can result in slow orders: poor infrastructure conditions, infrastructure maintenance or improvement, and weather-related issues.[10] Our data includes all except weather-related slow orders, and came from Amtrak's conductor delay reports.

$MECH_{ijt}$ is the average minutes of delays resulting from mechanical problems with all the Amtrak trains in our sample operating between station-pair *i*, on route *j*, in month *t*. The data for this measure was also taken from Amtrak conductor delay reports.

LATE ARR_{ijt} was used to capture delays induced by trains running well outside their scheduled time slot. It is defined as the percent of trains that arrive more than two hours late at the origin station of station-pair *i*, on route *j*, in month *t*. LATE ARR_{ijt} could be endogenous because factors not captured in our model could be expected to cause late arrivals. We use the percentage of trains that arrive late rather than the actual minutes of delay to minimize this potentiality.

---

[8] We obtained our weather data from the National Climatic Data Center.

[9] Amtrak provided data on the freights' heat-order policies. Different freights have different thresholds for imposing heat-related speed restrictions. We use 90 degrees since it is the minimum threshold across the various freight railroads' heat-order policies.

[10] Rail infrastructure includes, but is not limited to, track, ballast, signal systems, and sidings.

$DISTANCE_i$ is the weighted average distance between station-pair *i*, with the weights being the percentages of train and date combinations.[11] An increase in distance increases the probability of a delay event occurring, including as a result of factors not otherwise captured. Among these factors are delay-inducing track configurations, such as crossovers.

$TURN_i$ takes the value 1 if the origin station of the station-pair *i* on route *j* is an Amtrak crew base. A crew base is a location at which Amtrak changes crews.

Finally, we include fiscal year dummy variables to control for delay changes due to time trends that affect the whole system. These could include changes in Amtrak or host railroad practices, Amtrak ridership, and overall network congestion. FISCAL $YEAR_{2t}$ takes the value 1 if month *t* belongs to fiscal year 2002 and 0 otherwise. FISCAL $YEAR_{3t}$ through FISCAL $YEAR_{7t}$ are generated in a similar fashion.

ESTIMATION

We chose the Hausman-Taylor estimator for our panel data after a series of tests.[12] The tests showed fixed effects estimators to be preferable and the variable CAP $UTIL_{ijt}$ to be endogenous. The Hausman-Taylor estimator allowed us to address the endogeneity issue and to obtain coefficients for our time-invariant variables, such as host effects. We also tested the equality of coefficients across the two samples for long- and short-distance routes. The difference in coefficients confirmed our expectation that we need to estimate separate models for the two groups. Table B1 summarizes the estimation results. The reported bootstrapped standard errors are cluster-robust.[13]

There are several reasons to expect that CAP $UTIL_{ijt}$ has some degree of measurement error. It was constructed using data from multiple sources and several assumptions. Most notably, translating the Carload Waybill data into freight traffic by station-pair required the use of a routing algorithm, in our case the one underlying TRAGIS. Use of the Waybill data also required assumptions about the percentage of actual freight traffic the Waybill sample represented. It

[11] We obtained the distance information from Amtrak. It is weighted because of the slight changes in distance over time in approximately 40 percent of the station-pairs. We suspect that this difference in reported mileage occurs because of small changes in the methodology used in recording it.

[12] We performed an F-test on the significance of station-pair effects. The rejection of the null hypothesis of no station-pair effects indicated that the Fixed Effects (FE) estimator is preferable to the OLS estimator. To compare Random Effects (RE) and OLS estimators, we ran the Breusch-Pagan LM test and rejected the null hypothesis of no station-pair effects. This led to the RE estimator being preferred to the OLS estimator. We ran a bootstrapping Hausman test on each time-varying variable to compare FE and RE estimators. The results suggest that the variable CAP $UTIL_{ijt}$ is endogenous in our long-distance model.

[13] We compared OLS and cluster-robust standard errors. The significant differences between them on most of the variables in our model indicate the existence of clusters in the independent variables.

was further assumed that capacity did not change during the sample period along Amtrak routes. Lastly, we summarized the information on capacity utilization levels for all the track segments between each station-pair using a single measure, the maximum. All these considerations could be expected to have greater impact on the long-distance model as long-distance routes have many more track segments because of having almost twice as much distance between station-pairs on average as the short-distance routes. Consequently, we were not surprised that tests showed CAP $UTIL_{ijt}$ to be endogenous.

SENSITIVITY ANALYSES

We performed a series of sensitivity analyses to check the robustness of our findings. To test for the possibility of a nonlinear conditional expected value for the dependent variable resulting from the truncation of delays near zero, we re-estimated our model using a Tobit procedure.[14] The results deviated very little, if at all, from those reported above. To investigate the possibility of a nonlinear relationship between minutes of delay and the level of capacity utilization, we tested the significance of various thresholds for capacity utilization--the 50th, 60th, 70th, 80th, and 90th percentiles-in our models. [15] The results did not suggest a nonlinear relationship in our data.

We also experimented with augmenting capacity utilization by adding a traffic heterogeneity variable. An increase in traffic heterogeneity, or traffic mix, has been shown to increase train delays (Dingler *et al*, 2009). This occurs because different types of traffic travel at different speeds. The more diverse the speeds of the trains on a track segment, the more complicated the task facing a dispatcher trying to coordinate their movements.

We generated two variables to capture traffic heterogeneity: $HET1_{ijt}$ and $HET2_{ijt}$. $HET1_{ijt}$ is similar to the Herfindahl-Hirschman index. It is the sum of squares of service frequencies of different types of traffic, measured by the number of cars, for all the traffic at station-pair *i*, on route *j*, in month *t*.[16] $HET2_{ijt}$ is defined in the same fashion except that we weighted the components of the sum by the average speed of each type of traffic. We chose the minimum value of the index across all track segments between a station-pair for $HET1_{ijt}$ and $HET2_{ijt}$, because the segment with the minimum heterogeneity index has the most diverse traffic.[17] In both the long- and short-distance models, neither $HET1_{ijt}$ nor $HET2_{ijt}$ was

[14] Only 0.26 percent of delays were recorded as zero.

[15] Krueger (1999) found that the relationship between train delay and traffic volume can be expressed as an exponential equation. Since traffic volume is the numerator for our congestion variable, this led us to consider the possibility of a nonlinear relationship between the level of congestion and the average on-track delays.

[16] The types of traffic considered were passenger, commuter, and four types of freight trains.

[17] We also experimented with defining $HET1_{ijt}$ and $HET2_{ijt}$ as the average value of all the links in any given station-pair. The results did not change.

statistically significant, nor did they notably change the estimated coefficients on CAP $UTIL_{ijt}$.

We tried redefining LATE ARR_{ijt} to indicate that a train arrived either three or four hours later than scheduled. On long-distance routes, the estimated coefficient on LATE ARR_{ijt} increased by 9 percent using the three-hour definition and by 12 percent using the four-hour definition. On short-distance routes, the estimated coefficient on LATE ARR_{ijt} indicating a three-hour late arrival was not significant and a four-hour late arrival was marginally significant at the 10 percent level. The remaining coefficients change very little with the various definitions of LATE ARR_{ijt} in both the long- and short-distance route models. We expected the impact of late arrivals on delays to increase with the degree of lateness, and such results were observed in long-distance routes. However, trains rarely arrived more than two hours late in our short-distance route sample.

Alternative thresholds were also tested for the weather variables. $PRECIP_{ijt}$ is defined to be the percentage of days having greater than or equal to an accumulation of one inch of precipitation, which represents the 97.5th percentile in our sample. Cut-offs at the 85th, 90th, 95th percentiles, and average precipitation, were tested. The estimated coefficients in the long-distance model were 26 percent to 56 percent smaller than those on the precipitation variable as first defined, but were still significant at the one percent level. Similarly, in the short-distance model the estimated coefficients on the alternative measures were 48 percent to 64 percent smaller than with the initial definition.

For $TEMP_{ijt}$, which is currently defined as the percentage of days on which temperatures greater than or equal to 90 degrees are registered, we tested thresholds of 95, 100, and 110 degrees. The results showed that the higher the threshold, the greater the effect on delays on long-distance routes. However, the effect was insignificant using a 110 degree threshold, as temperatures above 110 degrees rarely occur. On short-distance routes, the coefficient on the temperature variable increases by 28 percent using a 95-degree threshold, and is significant at the 5 percent level. It becomes insignificant with a 100-degree definition. The coefficient becomes negative with an increase in absolute size of 116 percent when using a threshold of 110 degrees, and is significant at the one percent level.

REFERENCES

Clarke, D. B. (1995): "An Examination of Railroad Capacity and its Implications for Rail-Highway Intermodal Transportation", Ph.D. Dissertation. The University of Tennessee, Knoxville.

Dingler, M. H., Y-C. (R). Lai, C. P. L. Barkan (2009): "Impact of Train Type Heterogeneity on Single-Track Railway Capacity", *Transportation Research Record: Journal of the Transportation Research Board*, 2117, 41-49.

Krueger, H. (1999): "Parametric Modeling in Rail Capacity Planning", *Proceedings of the 1999 Winter Simulation Conference*, 2, 1194-1200.

Table B1: Hausman-Taylor Estimators for Long-Distance and Short-Distance Samples

Variable	Long-Distance Routes	Short-Distance Routes
Precipitation	2.836***	3.108***
	(0.614)	(0.774)
Temperature	0.804***	0.457***
	(0.161)	(0.159)
Capacity Utilization	1.869***	2.544**
	(0.692)	(1.041)
Slow Order	0.896***	1.128***
	(0.0322)	(0.0438)
Amtrak Mechanical Failure	0.156***	0.200***
	(0.0570)	(0.0694)
Late Arrival	2.843***	3.648**
	(0.304)	(1.557)
Meet UP	0.682	
	(0.649)	
Meet NS	0.158	
	(2.185)	
Distance	0.781***	1.895***
	(0.0575)	(0.131)
Turn Point	1.447***	1.457**
	(0.451)	(0.586)
UP	3.669***	0.945*
	(0.583)	(0.527)
BNSF	-1.902***	-1.805***
	(0.583)	(0.598)
CSXT	2.575***	-0.572
	(0.507)	(0.924)
NS	1.812***	-0.485
	(0.555)	(0.602)
CN	0.707	1.010
	(0.548)	(0.708)
Multiple Hosts	7.078***	5.233***
	(1.466)	(1.236)
FY 2002	-0.0956	-0.804***
	(0.203)	(0.173)
FY 2003	-0.294	-0.681***
	(0.187)	(0.176)
FY 2004	-0.136	-0.600***
	(0.162)	(0.150)
FY 2005	-0.0365	-0.655***
	(0.137)	(0.124)
FY 2006	0.568***	-0.228*
	(0.110)	(0.121)
Observations	48649	16035
Number of Station-Pairs	810	307

Notes: *Significant at the 10 percent level. **Significant at the 5 percent level.
***Significant at the 1 percent level.

EXHIBIT C. CONFIDENCE INTERVALS ON SYSTEM-WIDE DELAY ESTIMATES, IN SECONDS, BY CAUSE

Long-Distance Services

Variable	Estimate	95% CI Lower Bound	95% CI Upper Bound
Slow Orders	115	107	123
UP Host Effect	50	34	66
CSXT Host Effect	38	23	53
Late Arrival	30	24	36
Capacity Utilization	26	7	45
Turn Point	20	8	33
NS Host Effect	17	7	28
Multiple Host Effect	11	7	16
Amtrak Mechanical	4	1	7
BNSF Host Effect	-27	-44	-11

Short-Distance Services

Variable	Estimate	95% CI Lower Bound	95% CI Upper Bound
Slow Orders	63	58	67
Capacity Utilization	24	5	44
Turn Point	16	3	29
Multiple Host Effect	15	8	22
UP Host Effect	12	-1	25
Late Arrival	3	1	6
Amtrak Mechanical	2	1	4
BNSF Host Effect	-24	-39	-8

Note: The confidence interval for each estimate provides a range that the true value of the variable would be expected to fall within 95 percent of the time in the sample period.

EXHIBIT D. AMTRAK ROUTES

ROUTE	ENDPOINTS	
California Zephyr	Chicago, IL	Emeryville, CA
Capitol Corridor	Auburn, CA	San Jose, CA
Capitol Ltd	Washington, DC	Chicago, IL
Cardinal	Chicago, IL	New York, NY
Carolinian	New York, NY	Charlotte, NC
Cascades	Eugene, OR	Vancouver, BC
City of New Orleans	Chicago, IL	New Orleans, LA
Coast Starlight	Seattle, WA	Los Angeles, CA
Crescent	New York, NY	New Orleans, LA
Empire Builder	Chicago, IL	Seattle, WA
Heartland Flyer	Oklahoma City, OK	Fort Worth, TX
Hiawatha	Chicago, IL	Milwaukee, WI
Hoosier	Chicago, IL	Indianapolis, IN
Illinois (a)	Chicago, IL	Carbondale, IL
(b)	Chicago, IL	Quincy, IL
(c)	Chicago, IL	St. Louis, MO
Illinois/MO	Chicago, IL	Kansas City, MO
Kentucky Cardinal	Chicago, IL	Indianapolis, IN
Lake Shore Ltd	Chicago, IL	New York, NY
Michigan	Chicago, IL	Pontiac, MI
Missouri	Kansas City, MO	St. Louis, MO
Pacific Surfliner	San Luis Obispo, CA	San Diego, CA
Pennsylvanian	New York, NY	Pittsburgh, PA
Piedmont	Charlotte, NC	Raleigh, NC
San Joaquin	Sacramento, CA	Bakersfield, CA
Silver Service (a)	New York, NY	Savannah, GA
(b)	New York, NY	Miami, FL
Southwest Chief	Chicago, IL	Los Angeles, CA
Sunset Ltd (a)[a]	Orlando, FL	Los Angeles, CA
(b)[b]	New Orleans, LA	Los Angeles, CA
Texas Eagle	Chicago, IL	Los Angeles, CA
Three Rivers	Chicago, IL	New York, NY

[a] October 2001 to October 2005
[b] October 2005 to Present

EXHIBIT E. CONTRIBUTIONS TO AVERAGE STATION-PAIR DELAYS, IN MINUTES, ON INDIVIDUAL LONG-DISTANCE ROUTES BY CAUSE

Services	Capacity Utilization	Slow Orders	Amtrak Mechanical	Late Arrival	Turn Point	UP Host	BNSF Host	Multiple Host	CSXT Host	NS Host	Net Effect
California Zephyr	0.48	3.43	0.07	0.88	0.38	2.02	-0.83	0.11	0	0	6.55
Capitol Ltd	0.36	1.29	0.11	0.51	0.31	0	0	0.25	0.98	1.06	4.87
Cardinal	0.17	1.26	0.07	0.46	0.26	0	0	0.48	2.16	0.17	5.03
Carolinian	0.20	0.97	0.02	0.23	0.39	0	0	0.23	0.96	1.07	4.09
City of New Orleans	0.33	1.73	0.05	0.16	0.40	0	0	0.20	0	0	2.86
Coast Starlight	0.21	4.67	0.06	0.91	0.37	3.07	-0.28	0	0	0	9.01
Crescent	0.37	0.84	0.04	0.31	0.32	0	0	0.13	0.05	1.74	3.81
Empire Builder	0.83	0.89	0.07	0.20	0.38	0	-1.43	0	0	0	0.93
Illinois/MO	0.28	1.53	0.02	0.13	0.15	3.17	0	0.58	0	0	5.86
Lake Shore Ltd	0.33	0.90	0.07	0.54	0.41	0	0	0.50	1.16	0.87	4.78
Silver Service	0.20	1.48	0.04	0.55	0.35	0	0	0.08	2.21	0.03	4.93
Southwest Chief	0.62	0.65	0.13	0.26	0.36	0	-1.87	0.11	0	0	0.27
Sunset Ltd	0.94	3.08	0.12	1.23	0.33	2.25	0	0	0.99	0	8.94
Texas Eagle	0.37	4.57	0.05	0.57	0.22	2.78	-0.27	0.56	0	0	8.85
Three Rivers	0.56	0.83	0.12	0.25	0.23	0	0	0.28	0.80	1.17	4.24

Exhibit E. Contributions to Average Station-Pair Delays, in Minutes, on Individual Long-Distance Routes by Cause

EXHIBIT F. CONTRIBUTIONS TO AVERAGE STATION-PAIR DELAYS, IN MINUTES, ON INDIVIDUAL SHORT-DISTANCE ROUTES BY CAUSE

Services	Capacity Utilization	Slow Orders	Amtrak Mechanical	Late Arrival	Turn Point	UP Host	BNSF Host	Multiple Host	Net Effect
Capitol Corridor	0.27	0.95	0.02	0.01	0.31	0.95	0	0	2.50
Cascades	0.60	1.78	0.05	0.03	0.56	0.26	-1.32	0	1.96
Heartland Flyer	0.85	3.11	0.01	0.04	0.24	0	-1.80	0	2.45
Hiawatha	0.08	0.29	0.03	0	0.42	0	0	0	0.82
Hoosier	0.30	2.28	0.12	0.13	0.29	0	0	1.05	4.17
Illinois	0.37	0.87	0.04	0.03	0.20	0.26	-0.49	0.30	1.58
Kentucky Cardinal	0.26	1.51	0.08	0.33	0.29	0	0	1.05	3.52
Michigan	0.27	1.04	0.03	0.09	0.31	0	0	0.71	2.45
Missouri	0.57	2.36	0.01	0.22	0.17	0.95	0	0	4.29
Pacific Surfliner	0.24	0.17	0.03	0.01	0.24	0.10	-0.21	0	0.58
Pennsylvanian	0.65	0.42	0.07	0.13	0.21	0	0	0.06	1.55
Piedmont	0.27	0.63	0.01	0.01	0.19	0	0	0	1.10
San Joaquin	0.75	0.58	0.04	0.04	0.15	0.25	-1.28	0.13	0.65

Exhibit F. Contributions to Average Station-Pair Delays, in Minutes, on Individual Short-Distance Routes by Cause

EXHIBIT G. MAJOR CONTRIBUTORS TO THIS REPORT

Name	Title
Betty Krier	Program Director/ Supervisory Economist
Chia-Mei Liu	Senior Economist
Brian McNamara	Senior Economist
Jerrod Sharpe	Senior Economist
Kang Hua Cao	Economist
Susan Neill	Writer/Editor