THOMAS P. DiNAPOLI
COMPTROLLER

110 STATE STREET
ALBANY, NEW YORK 12236

STATE OF NEW YORK
OFFICE OF THE STATE COMPTROLLER

September 3, 2020

Mr. Patrick J. Foye
Chairman and Chief Executive Officer
Metropolitan Transportation Authority
2 Broadway
New York, NY 10004

Re: Subway Wait Assessment
Report 2019-S-62

Dear Mr. Foye:

Pursuant to the State Comptroller's authority as set forth in Article X, Section 5 of the State Constitution and Section 2803 of the Public Authorities Law, we audited the Metropolitan Transportation Authority (MTA) – New York City Transit (Transit) to determine whether Wait Assessment performance measurements reported to the MTA's Board and customers are accurate and consistently calculated. The audit covered the period between July 1, 2017 and November 30, 2018.

Background

Under the Public Authorities Law, the MTA is required to issue an annual report on its mission statement, measurements, and performance indicators. The MTA's 2017 and 2018 calendar year Mission Statement and Performance Indicator reports list Subway Wait Assessment (WA) as an indicator of the MTA's progress toward its goal of providing reliable service for its customers and federal, State, and local government partners.

Performance measure results are often reviewed by the Board, which provides feedback, guidance, and recommendations to Transit on the success of its operations. On July 25, 2017, in response to the Board's demand for performance improvements, the MTA released the Subway Action Plan to improve service. One of the changes in the Plan was the implementation of new customer-focused performance measures intended to be more relevant and easier to understand. In September 2017, Transit introduced the new metrics[1]: Additional Platform Time, the average time that customers wait at a station beyond their scheduled wait time; Additional Train Time, the average time customers spend on board a train beyond their scheduled travel time; and the sum of these, Additional Journey Time.

With the introduction of these new metrics, WA became a "legacy indicator," but continues to be reported monthly to the MTA Board's Transit and Bus Committee (Committee) and posted on the MTA website's Subway Performance Dashboard (Dashboard).

1 See our prior report, issued January 17, 2020, titled *New Customer-Focused Subway Metrics*.

Results of Audit

Transit's Operations Planning (OP) is responsible for designing effective and efficient subway and bus service. OP's System Data and Research Division calculates WA based on computerized train tracking information and generates reports that are used within OP, reported monthly at Committee meetings, and posted on the Dashboard.

The MTA New York City Transit Service Guidelines Manual (Guidelines), issued in August 2010 and revised in February 2013, is used to develop and maintain comprehensive, cost-efficient transit service that meets the needs of those who live, work, and travel in New York City. The Guidelines provide a structure for evaluating service by determining when, where, and how frequently service should be offered and address factors such as scheduling (minimum frequency) and loading (number of customers on a train).

WA is measured as the percentage of actual intervals between trains that are no more than the scheduled intervals plus 25 percent. Referred to as "headway," for this purpose, it is the time between trains departing a subway station. WA is measured on weekdays and weekends from 6 a.m. to midnight and is calculated only at select subway stations, called time points. Transit used a 2013 case study titled "Measuring Subway Service Performance at New York City Transit: A Case Study Using Automated Train Supervisor (ATS) Track-Occupancy Data" to explain how WA is calculated based on ATS data, which only covers the numbered subway lines (A Division). Since March 2017, Transit developed systems to record and store actual departure times on lettered lines (B Division). OP uses a scheduling software to generate the base and supplement schedules (includes changes such as scheduled track work).

WA calculation involves comparing the actual intervals between trains to the scheduled intervals separately for each route and direction.

To determine whether the reported statistic is accurate, we requested the data used to support WA for calendar year 2018. OP provided us with monthly data sets that compiled each individual scheduled train departure and actual train movement at 82 of the 472 total subway stations.

To determine whether the WA percentages are accurate, we tested the 696,722 scheduled trains accounted for in the November 2018 WA of 71.6 percent. In addition, we tested the scheduled headways between trains in November 2018 to determine if they were appropriate for each service type (e.g., peak, off-peak). We also reviewed the reasonableness of OP calculating WA based on time points rather than all subway stations.

According to OP, subway WA is calculated at time points, which should consist of between 25 and 50 percent of all stops along each route. However, based on our review of the supporting data, only 82 of the 472 subway stations (17 percent) are considered time points and are included in the WA calculation. These 82 stations are equipped with 239 time points (120 northbound and 119 southbound). For example, 42nd Street-Times Square is a time point on the Number 1, 2, and 3 trains in both directions and counts as 6 of the 239 time points. OP explained that time points are selected because they are typically major transfer subway stations, hubs, and originating terminals, and a majority of the riders enter and exit the subway at these locations.

Although WA depicts the conditions at busier subway stations, it does not reflect the subway system as a whole. Moreover, the MTA does not disclose to the Committee or on the Dashboard that WA is only based on certain stations. Based on our review, only one

train service line, Number 4, calculated WA on more than 25 percent of scheduled stops per direction. Additionally, WA calculations may be affected by trains opting to skip non-time point stops in order to avoid arriving late at time points. For a chart of all 21 subway service lines (excluding three shuttle trains) with the percentage of stops used in the WA formula, see Exhibit.

We found that time points only cover between 11 and 30 percent of all stops on each route per direction. Therefore, the majority of the service provided is not captured in the WA calculation. For example, the Number 1 train, a local service train that runs from 242nd Street in the Bronx to South Ferry in downtown Manhattan, calculates WA based on only 5 of the 38 stops (13.2 percent). See Figure 1 for a map of all stations on the Number 1 line.

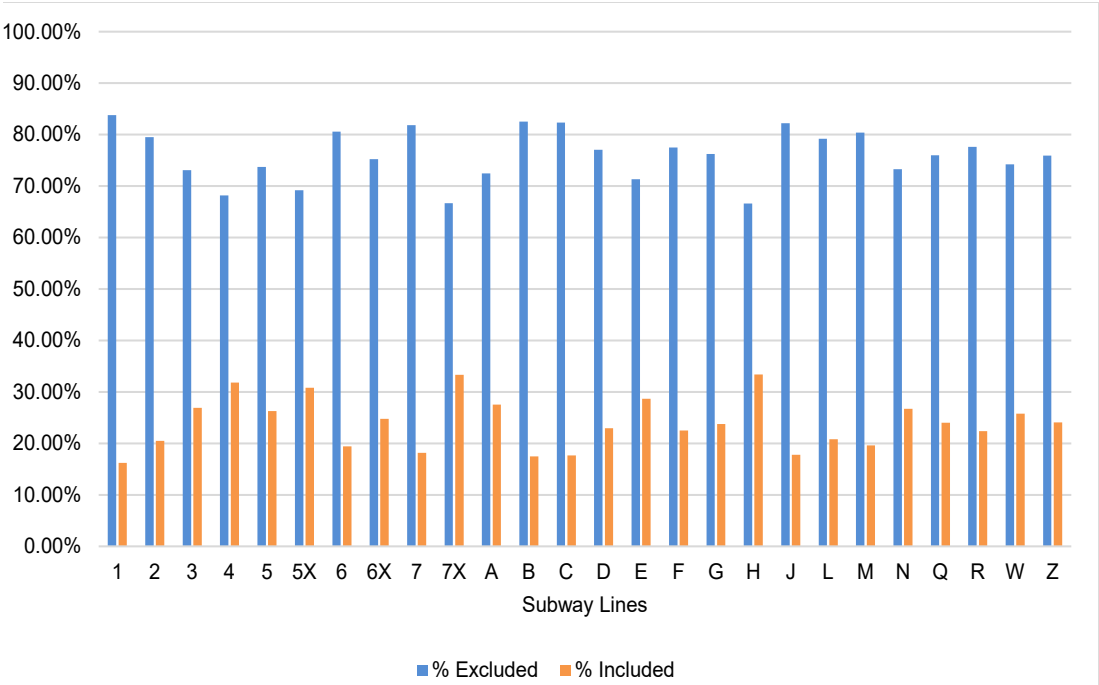
Figure 1 – Time Point Locations for Number 1 Line (Northbound)

Even though the time points were selected based on customer traffic, OP failed to include key subway stations such as Penn Station, 59th Street-Columbus Circle, and 14th Street. These stations had the second, third, and fourth highest rates of ridership, respectively, in 2017 and on average from 2012–2017. Moreover, these subway stations provide transfer options to nine subway lines, the Long Island Rail Road, Amtrak, New Jersey Transit, and Port Authority Trans-Hudson, as well as transportation options to the airport.

For November 2018, more than 3.1 million actual train records were excluded from the WA calculation because they were recorded at stations that are not time points. As a result,

OP may not be providing its customers and government partners with complete performance results. See Figure 2 for the percentage of trains considered in the WA calculation per service line.

Figure 2 – Percent of Data Excluded/Included in WA November 2018



Scheduled Intervals

The amount of time scheduled between train departures is a key component of WA. We reviewed the scheduled intervals used for the November 2018 WA calculation and found that 5.5 percent of the trains OP reported as having met WA goals were compared to scheduled intervals that were longer than those recommended by the Guidelines.

Using headway that exceeds time frames supported by the Guidelines in WA calculation resulted in inaccurate reporting (see table below).

	Scheduled Trips	WA Met	Maximum Headway Minutes per Guidelines	Headways		
				Did Not Meet Guidelines	Passed WA Without Meeting Guidelines	Avg. Minutes Did Not Meet Guidelines
AM Peak	120,858	83,198	10	6,869	5,689	12
Midday	163,858	123,637	12	2,653	2,436	16
PM Peak	122,539	82,685	10	6,334	5,257	12
Evening	112,111	80,966	12	10,809	8,917	15
Transition Period	177,356	128,391	12*	5,898	5,160	15
Totals	**696,722**	**498,877**		**32,563**	**27,459**	
WA		**71.6%**			**84.3%**	

*Transition period maximum headways are not disclosed in the Guidelines. There is an expectation that service is still being provided at intervals that do not exceed the off-peak time limits.

WA is calculated by dividing the total number of trains that met WA goals by the total number of trains scheduled. In November 2018, OP calculated system-wide WA compliance to be 71.6 percent (498,877/696,722). However, 32,563 of the 696,722 scheduled headways were longer than those recommended by the Guidelines and 27,459 of those trains were still reported as having met WA goals. Eighty-four percent of trains with longer-than-recommended scheduled headways (27,459/32,563) were reported as having met WA goals.

We otherwise found OP's calculation of WA is generally consistent with its public definition, except for in the following respect:

According to OP, a train interval cannot meet WA goals if it does not match the scheduled interval. We found that each month's data had train intervals that were reported as having met WA goals even though they did not match a scheduled interval. In November 2018, system-wide, 1,032 actual train intervals did not match a scheduled interval, but still were reported as having met WA goals.

- 744 of the 1,032 actual train departure records were duplicated, and each was compared against the same scheduled interval. As a result, one train departure at a specific subway station on the same date was counted as two individual trips and passed twice.
- 288 of the 1,032 departures were trains that provided service, but did not match a scheduled interval. However, they were still included as meeting WA goals.

Furthermore, WA is reported for each of the 24 subway service lines and is presented separately in the Committee reports for both weekday and weekend service. However, on the Dashboard, an individual can filter WA data by line, division, and peak/off-peak service for only weekday service – weekend statistics are not provided on the Dashboard. Weekend statistics would be a valuable tool for the public. OP has not closely monitored WA data to ensure that it is accurate and that proper calculations were made.

Recommendations

1. Calculate system-wide WA performance based on all subway stations rather than time point stations. In the interim, disclose that WA is calculated based on only time point locations.

2. Provide the public with the supplement schedules that are used to calculate WA. Only include intervals that are in accordance with Guidelines in the WA calculation and, in the interim, disclose that headways may not be in accordance with the Guidelines.

3. Provide weekend WA statistics on the Dashboard.

Audit Scope, Objective, and Methodology

The objective of our audit was to determine whether WA performance measurements reported to the MTA's Board and customers are accurate and consistently calculated. The audit covered the period between July 1, 2017 and November 30, 2018.

To accomplish our objective and evaluate the relevant internal controls, we reviewed MTA – Transit's related policies, procedures, and Guidelines as well as regulations and laws.

We interviewed officials and employees of Transit's OP and Department of Subways to obtain an understanding of the processes used to calculate WA.

We requested the data used to support WA for the 2018 calendar year to determine whether the statistic reported to the Committee and on the Dashboard is complete and accurate. OP provided us with monthly data sets that compiled each individual scheduled train departure and actual train movement at the 82 time points.

To determine whether the WA percentages are accurate, we tested the 696,722 scheduled trains accounted for in November 2018's 71.6 percent WA. In addition, we tested the scheduled headways between trains in November 2018 to determine if they were appropriate for each service type (i.e., peak service, off-peak service). We also reviewed the reasonableness of OP calculating WA based on time points rather than all subway stations.

Statutory Requirements

We conducted our performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objective. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objective.

In addition to being the State Auditor, the Comptroller performs certain other constitutionally and statutorily mandated duties as the chief fiscal officer of New York State. These include operating the State's accounting system; preparing the State's financial statements; and approving State contracts, refunds, and other payments. In addition, the Comptroller appoints members to certain boards, commissions, and public authorities, some of whom have minority voting rights. These duties may be considered management functions for purposes of evaluating organizational independence under generally accepted government auditing standards. In our opinion, these functions do not affect our ability to conduct independent audits of program performance.

Reporting Requirements

We provided a draft copy of this report to MTA – Transit officials for their review and formal comment. Their comments were considered in preparing this final report and are included in their entirety at the end of the report.

In their response, Transit officials disagreed with the recommendations to change the calculation methodology for WA. They did agree there is a need for transparency with respect to the calculation of WA and indicated they have taken steps to disclose that WA is based only on time point locations. While the decision to disclose the basis for WA is a step in the right direction, it does not address the fact that most of the subway stops are not included in the calculation of WA, a condition that needs to be corrected. Our responses to certain comments are included in the report's State Comptroller's Comments.

Within 180 days after the final release of this report, as required by Section 170 of the Executive Law, the Chairman of the Metropolitan Transportation Authority shall report to the Governor, the State Comptroller, and the leaders of the Legislature and fiscal committees, advising what steps were taken to implement the recommendations contained herein, and where the recommendations were not implemented, the reasons why.

Contributors to this report Robert C. Mehrhoff, David DiNatale, Jim Cherian, Celedonia Deaver, Lillian Fernandes, and Susan Gordon.

We wish to thank the management and staff of the Metropolitan Transportation Authority for the courtesy and cooperation extended to our auditors during this audit.

Very truly yours,

Carmen Maldonado
Audit Director

cc: M. Woods, MTA, OAG
D. Jurgens, MTA, OAG
NYS Division of the Budget

Exhibit

All 21 Subway Service Lines and Percentage of Stops Used in WA Formula

Line*	Stops	Northbound (NB) Time Points	Southbound (SB) Time Points	Percent of Time Points per NB Route	Percent of Time Points per SB Route	Total Percent of Time Points in Both Directions
1	38	5	5	13.16%	13.16%	26.32%
2	49	8	8	16.33%	16.33%	32.65%
3	34	7	7	20.59%	20.59%	41.18%
4	27	8	8	29.63%	29.63%	59.26%
5	44	8	9	18.18%	20.45%	38.64%
6	38	5	5	13.16%	13.16%	26.32%
7	22	3	3	13.64%	13.64%	27.27%
A	44	8	6	18.18%	13.64%	31.82%
B	45	5	5	11.11%	11.11%	22.22%
C	40	5	5	12.50%	12.50%	25.00%
D	36	5	5	13.89%	13.89%	27.78%
E	22	4	4	18.18%	18.18%	36.36%
F	45	7	7	15.56%	15.56%	31.11%
G	20	4	4	20.00%	20.00%	40.00%
J/Z	30	4	4	13.33%	13.33%	26.67%
L	24	4	4	16.67%	16.67%	33.33%
M	34	6	6	17.65%	17.65%	35.29%
N	26	6	6	23.08%	23.08%	46.15%
Q	29	4	4	13.79%	13.79%	27.59%
R	45	7	7	15.56%	15.56%	31.11%
W	21	4	4	19.05%	19.05%	38.10%

*Excludes shuttles.

Agency Comments

2 Broadway
New York, NY 10004
212 878-7000 Tel

Patrick J. Foye
Chairman and Chief Executive Officer

Metropolitan Transportation Authority
State of New York

July 2, 2020

Ms. Carmen Maldonado
Audit Director
The Office of the State Comptroller
Division of State Government Accountability
59 Maiden Lane, 21st Floor
New York, NY 10038

Re: Draft Report #2019-S-62 (Subway Wait Assessment)

Dear Ms. Maldonado:

This is in reply to your letter requesting a response to the above-referenced draft report.

I have attached for your information the comments of Sarah E. Feinberg, Interim President, MTA New York City Transit, which address this report.

Additionally, I will be working with staff to ensure that management is following up on and enforcing the audit's recommendations, where appropriate, and requesting regular, interim reports to that effect.

Sincerely,

Patrick J. Foye
Chairman and Chief Executive Officer

c: Anni Zhu, Acting Chief of Staff to the MTA Chairman & Chief Executive Officer
Michele Woods, Acting Auditor General, MTA Audit Services

Memorandum

MTA New York City Transit

Date July 1, 2020

To Patrick J. Foye, Chairman, Metropolitan Transportation Authority

From Sarah E. Feinberg, Interim President, MTA New York City Transit

Re **New York State Comptroller's Request for Response to the Audit Report "Subway Wait Assessment" 2019-S-62**

This is in reply to the Office of the State Comptroller's audit of New York City Transit's Wait Assessment Performance Measure, dated June 3, 2020. New York City Transit has reviewed the Office of the State Comptroller's recommendations and responds as follows:

This audit examines NYCT's Wait Assessment (WA) performance metric. As the preliminary finding notes, WA is a "legacy indicator," developed long before electronic train movement data was available. Recognizing the limitations of this metric, NYCT introduced in September 2017 a new series of customer-focused metrics—Additional Platform Time, Additional Train Time, and Customer Journey Time Performance—all of which were intended to better reflect service quality at all stations and utilize the newly available electronic train movement data. When the new customer-focused metrics were introduced, WA was maintained as a legacy indicator primarily to provide some basis for comparing historical performance with current performance. In other words, the primary purpose of WA at present is to allow historical year-on-year comparisons. NYCT believes that any attempt to calculate WA differently negates the metric's principal present-day purpose, while doing nothing to improve reporting on service quality at all stations, because the current customer-focused metrics better serve that objective.

Comment 1

Accordingly, NYCT disagrees with the audit's recommendations to change the calculation methodology. NYCT does agree, however, with the need for transparency with respect to the calculation of WA and has already taken steps to disclose details of the WA methodology.

1. *Recommendation 1: Calculate system-wide WA performance based on all subway stations rather than timepoint stations. In the interim, disclose that WA is calculated based on only timepoint locations.*

NYCT disagrees with the first part of this recommendation. The primary purpose of WA at present is to allow historical year-on-year comparisons, not to best represent service quality at all stations. WA is a legacy indicator that originated before electronic train movement data was widely available. As a result of this data limitation, NYCT calculated WA based on a sample of observations taken at timepoints, because those are the only locations (other than terminals) where there are scheduled arrival times. Even though, over the past decade, electronic train location data has become available systemwide, we do not agree that it would be appropriate to alter the WA algorithm. Calculating WA using the same methodology is important for

Comment 2

continuity and evaluation of historical trends. Our new, customer-focused performance metrics, which have been introduced since systemwide train location data has become available and measure performance at all stations, better capture customer experience using the system than legacy metrics like WA. Comment 3

NYCT agrees with second part of the recommendation and has already implemented it. The description of WA on the public dashboard and in Board materials has been updated to specify that the metric is measured at timepoints.

2. *Recommendation 2: Provide the public with the Supplemental Schedule that is used to calculate WA. Only include intervals that are in accordance with NYCT Service Guidelines in the WA calculation and in the interim, disclose that headways may not be in accordance with the Guidelines.*

NYCT agrees that the public should be notified of service changes, but disagrees that any additional action is needed with respect to the first part of this recommendation. Planned changes to service are already available to the public in advance, through the "Planned Service Changes" section of www.mta.info. Customers are also advised of planned work as it is happening in the "Service Status" box on the website's home page. In addition, NYCT provides and updates daily a supplemented GTFS feed for the subway (available in machine-readable form for use in third-party applications) which accounts for scheduled service diversions for the following seven days. This information allows customers to plan ahead in response to subway service changes, and it is why NYCT evaluates performance relative to supplemental schedules, rather than the base schedule only. Comment 4

NYCT disagrees with the second part of the recommendation. In addition to the concerns already explained with altering WA given its role as a historical comparison metric, NYCT also disagrees with any suggestion that scheduled headways greater than the policy maximum headways are not in accordance with the service guidelines. The service guidelines call for maximum headways ***where operationally feasible***. For example, the service guidelines allow that work on the right-of-way may necessitate headways greater than what are called for by the policy maximum headways. Similarly, guidelines allow for variations in headways that may exceed the maximum headways for other operational reasons, such as allowing for other routes (sometimes on different scheduled headways) to share the same tracks. Such headways are still in accordance with the guidelines because it is not operationally feasible to maintain the policy maximum headways. As such, it is fully appropriate for WA to be based on headways greater than the policy maximum headways where operational considerations so dictate. In addition, NYCT believes transparency is undercut if the agency reports WA for only a subset of intervals. Comment 5

3. *Recommendation 3: Provide weekend WA statistics on the dashboard.*

NYCT disagrees with this recommendation. Operations Planning (OP) has considered adding Weekend Wait Assessment to the dashboard, but decided against it. Because WA is a legacy Comment 6

metric, we do not believe that it is an efficient use of limited staff resources to develop and implement this metric on the dashboard. It should be noted that Weekend Wait Assessment results are publicly available online in the meeting materials of the Transit Committee of the MTA Board and can be found at http://web.mta.info/mta/news/books/.

* * *

Please also note the following clarifications to the other sections of the preliminary finding:

- Page 2, under "Background": Please note that OP's System Data and Research unit no longer exists. These functions now reside within the Data, Research and Development and the Operations Improvement and Analysis units within OP.

- Page 2, WA Calculation at Timepoint Stations: The audit leaves the incorrect impression that there is something problematic about the fact that 17 percent of subway stations are timepoints included in the WA calculation. While it is true that 25 to 50 percent of all stops on each route could be timepoints, the research report that documents the WA methodology (and that is referenced in the audit) makes clear that only "a subset of" timepoints "are used for WA reporting." It is therefore not discrepant for 17 percent of stations to be timepoints used for WA reporting. Thus, any suggestion that the number of timepoints selected to calculate WA is inconsistent with the WA calculation methodology is misplaced.

Comment 7

- Pages 2-3: While the audit correctly identifies examples of how timepoints are selected based on the referenced research paper documenting the WA methodology, the audit also gives the incorrect impression that each timepoint must meet all of the conditions given in these examples. Note that, historically, the selection of timepoints was based not on ridership volume, but rather on the locations of towers or control points where staff in the field supervised train operations, by routing trains onto different tracks, holding trains in stations using holding lights, and recording the times of trains as they passed. That is why these points were selected in the timetables to be timepoints. When OP developed WA as a measure, it took advantage of these pre-existing timepoints, adapting them for analytical purposes, because those were the locations where reliable data was available. Even as control and supervision of some lines has been centralized and automated, the control points/timepoints have remained unchanged. This is one of the reasons WA is considered a legacy metric, and labeled as such in our public reporting. As explained earlier, to make the metric comparable across prior time periods, NYCT has not changed the calculation methodology or the locations of the timepoints on which the metric is based. NYCT's customer-focused metrics, introduced in 2017, take advantage of the greater availability of electronic train location data. These metrics include measurements at every station and on every line, and even further, they are weighted by passengers. These metrics are a better reflection of passenger trip experience.

Comment 8

Comment 9

- Page 5: NYCT considers these differences to be due to slight differences in calculation methodology, rather than an error. The differences are so small that they have no effect on the calculation. The 1,032 records represent only 0.14% of the total records for November 2018. The inclusion or exclusion of these records is immaterial – the State Comptroller calculated the

Comment 10

same figure for November 2018 WA when excluding these records as NYCT did when including them. Furthermore, NYCT disagrees with the statement that "OP has not closely monitored WA data to ensure that it is accurate and that proper calculations were made." NYCT has a robust data validation process that is used to verify the metric on a daily basis.

State Comptroller Comments

1. A customer-friendly statistic based on more complete data provides the riding public with transparency. The new metrics, unlike WA, rely on significant assumptions. While compatibility is important, so are accuracy and transparency.

2. Transit has the historic data to update WA for the A Division. For the B Division, it can use ATS and the new electronic systems installed over the past several years. Our primary concern is that these performance metrics are based on data, while the new metrics rely on significant assumptions. The riding public and decision makers require accurate and complete information to assess agency performance.

3. The new statistics make several assumptions, including where a rider detrains and which train a rider will take when more than one train is available, and add additional platform and train time for each part of a trip to determine the total additional trip time. The numerous assumptions create doubt regarding these metrics' reliability.

4. This response does not address the recommendation, which is to provide the supplemental schedule actually used to calculate WA.

5. Scheduled headways that exceed the Guidelines should be reported as such and not as "meeting" the Guidelines. Inconsistent application of the Guidelines undermines their value because any headway can be declared acceptable. The planned work should be covered in the supplemental schedules.

6. The MTA's 3.1 million Saturday riders and 2.4 million Sunday riders would benefit from being able to access weekend WA information on the Dashboard. Transit points out that the weekend WA is available in the Committee materials, but that requires the public to navigate through the MTA's website and usually over 100 pages to locate the information versus simply visiting the Dashboard for weekend WA.

7. OP officials stated that time points should consist of between 25 and 50 percent of all stops along the route. As stated in the report, the actual number ranged from 11 to 30 percent. WA accuracy is impacted by omitting the majority of stops in any route. Additionally, a train that is running late can skip a stop that is not a time point to improve WA performance, but riders at the missed stop have an increased wait time. Using more time points could increase WA accuracy as well as accountability to the public.

8. There appears to be a miscommunication between the workers and management as to what WA represents, because, on the A Division, the information is obtained from ATS, as stated in the 2013 study, and the B Division information is from three different sources (e.g., Beacon and Programmable Logic Control). Prior to March 2017, WA was calculated based on information from Traffic Checkers.

9. See Comment 3.

10. OP officials claimed that a train cannot pass WA if it did not match with a scheduled interval. In response to the draft report, Transit officials claim they have a "robust validation process," but did not provide any support for the statement. Moreover, a truly robust validation process would have identified the exceptions.