2021

Arterial Roadway Travel Time Reliability and the COVID-19 Pandemic

Laurence Rilett

Ernest Tufuor

Sean Murphy

Arterial Roadway Travel Time Reliability and the COVID-19 Pandemic

Laurence R. Rilett, Ph.D., P.E., F.ASCE[1] Ernest Tufuor, Ph.D.[2] and Sean Murphy[3]

1 Distinguished Professor and Director, Nebraska Transportation Center, Keith W. Klaasmeyer Chair of Engineering, 262D Prem S. Paul Research Center at Whittier School, University of Nebraska-Lincoln, Lincoln, NE 68583-0851 (corresponding author). *Email*: lrilett2@unl.edu

2 Postdoctoral Research Associate, Nebraska Transportation Center, 262E Prem S. Paul Research Center at Whittier School, University of Nebraska- Lincoln, Lincoln, NE 68583-0851. *Email*: ernest.tufuor@huskers.unl.edu

3 Graduate Research Assistant, University of Nebraska-Lincoln, Nebraska Transportation Center, 262E Prem S. Paul Research Center at Whittier School, Lincoln, NE 68583-0851; Traffic EIT, HDR Inc., 1917 S. 67th St., Omaha, NE 68106-2973. *Email*: Sean.W.Murphy@hdrinc.com

ORCID

Laurence R. Rilett, https://orcid.org/0000-0002-6925-5306

Ernest Tufuor, https://orcid.org/0000-0002-2816-1055

Abstract

This paper evaluated the effect of the COVID-19 preventive orders on arterial roadway travel time reliability (TTR). A comparative analysis was conducted to examine average travel time distributions (TTD), and their associated TTR metrics, before and during the COVID-19 pandemic. Travel time data for four urban arterial corridors in Nebraska, disaggregated by peak period and direction, were analyzed. It was found that in 2020, the average TTD mean and standard deviation values for all 16 scenarios were reduced by an average of 14.0% and 43.4%, respectively. The travel time index, the planning time index, the level of travel time reliability (LOTTR), and

Published in *Journal of Transportation Engineering, Part A: Systems* 147:7 (2021), #04021034, 13 pp.

DOI: 10.1061/JTEPBS.0000559.

Submitted November 4, 2020; approved March 26, 2021; published May 13, 2021.

the buffer index metrics associated with these TTDs were reduced, on average, by 14.0%, 19.7%, 3.5%, and 35.0%, respectively. In other words, whether the test corridors were more reliable during the pandemic was a function of which TTR metric was used. The paper concludes by arguing for a fundamental change in how arterial TTR is measured and reported to different user groups.

Keywords: Travel time reliability, COVID-19 pandemic, Arterial roadway, INRIX, NPMRDS

Background

In response to the COVID-19 pandemic, caused by the highly infectious and novel SARS-CoV-2 virus, many countries introduced severe intervention measures including stay-at-home orders, self-quarantining, social distancing, and face-coverings. Although these interventions were meant to reduce the transmission of the virus, they also had a direct effect in restricting travel.

The COVID-19 pandemic, and the response to it, provides a natural experiment for studying the performance of transportation systems and how transportation agencies monitor and communicate this information. For example, preliminary data from the Federal Highway Administration (FHWA) show an estimated 3.3% decline in overall fatalities and more than a 16% decrease in total traffic volume on US highways during the first half of 2020 compared with the same period in 2019 (NHTSA 2020). Because the number of fatalities decreased at a slower rate than the decrease in volume, the 2020 traffic fatality rate [e.g., fatalities per 100 million vehicle-miles traveled (VMT)] will be higher than in previous years. In this example, the answer to whether traffic safety improved during the pandemic will be dependent on which metrics are used to analyze the system.

In the last 10 years, the concept of using travel time reliability (TTR) to define how a given roadway is performing has been adopted widely. However, there is no universally accepted definition of TTR and, as a consequence, a number of TTR metrics have been developed for measuring reliability. The COVID-19 pandemic provides a natural experiment to study these TTR metrics. This paper focused on arterial roadways because, although they constitute less than 10% of all roadway mileage, they account for nearly half of all vehicle miles traveled in the US (Reid 2004). Specifically, this paper addressed the following research questions:

1. Did the pandemic result in more reliable arterial roadway travel times?
2. Which of the commonly used TTR metrics best captures the change in TTR caused by the pandemic?
3. Are there better ways of measuring TTR?

This paper compared arterial travel time distributions (TTDs) from Nebraska during the initial enactment of COVID-19 stay-at-home orders (e.g., March–May 2020) with TTDs during the same period in 2018 and 2019. The paper also assessed the sensitivity of both standard descriptive statistics and the commonly used TTR metrics to travel changes brought about by the pandemic. To the authors' knowledge, this is the first paper to (1) examine arterial travel time distributions and their changes in response to the COVID-19 pandemic, and (2) identify which of the commonly used TTR metrics best captures changes in system reliability.

The paper is divided into five sections. The first section introduces the concept of travel time reliability, with a special emphasis on arterial roadways. The second section reviews related research on the effects of the COVID-19 pandemic on travel behavior. The next section describes the 2018, 2019, and 2020 travel time data that were used in the analysis. This is followed by a statistical analysis of the changes in arterial travel times brought about by the pandemic and an analysis of commonly used TTR metrics. The paper concludes with a summary of the research results, and recommendations related to measuring TTR in future studies.

Travel Time Distribution and Travel Time Reliability

Travel time arguably is the most commonly used metric for analyzing how a given system is performing (Yang and Cooke 2018). This is because travel time is understood easily by both roadway users (e.g., the general public and shippers) and managers of the traffic systems (Lomax and Schrank 2002). Because travel time varies across space and time, it is important that the underlying characteristics of the travel time are defined explicitly. Typically, travel times can be categorized in four ways

1. Spatial: Travel time is defined over a given link, corridor, or system. This paper focused on arterial roadway corridors, although the analysis can be generalized to any spatial component of the system.
2. Time of day: Travel time is defined over a given period in the day because travel time can vary because of changes to volume and/or traffic operation conditions. Therefore, users often analyze travel time over specific periods [e.g., morning (AM) peak, evening (PM) peak, off-peak, and so forth]. This paper focused on the AM and PM peak periods because these are the most congested periods of the day.
3. Aggregation: Travel time can be analyzed at an individual vehicle level or aggregated into average values. Often, these periods are disaggregated into smaller periods (e.g., 15 min) to capture changes over a given period. This paper examined travel time at an aggregate or average travel time level for each 15-min period to capture any changes in travel time during a given peak period.
4. Analysis period. Travel time often is analyzed over a set period and for a set number of days. In this paper, the analysis period consisted of weekdays in March, April, and May in 2018, 2019, and 2020.

After the preceding characteristics are defined, travel time may be modeled using a continuous distribution. In this paper, the arterial TTDs were based on empirical travel time measurements. Because travel time is modeled as having a distribution, standard statistical metrics can be used to describe its characteristics, including measures of central tendency (e.g., mean and median), measures of dispersion (e.g., standard deviation and interquartile range), and measures of symmetry (e.g., skewness). Traditionally, transportation officials have tended to use point measures of the travel time distribution to describe how their systems are performing. These were either measures of central tendency, such as the mean, or measures of extreme values, such as the 90th percentile travel time. However, using a single metric to represent system performance has fallen out of favor recently (FHWA 2017). One reason for this is that transportation agencies now have access to much broader and robust travel time information, as is discussed subsequently.

In the last 10 years, transportation officials have begun to utilize the concept of travel time reliability (TTR) to categorize how their systems are performing. For example, the US Federal Highway Administration identified TTR as a key roadway mobility performance indicator (FHWA 2012, 2015). In addition, the most recent version of the *Highway Capacity Manual* includes an arterial TTR estimation methodology for the first time (Transportation Research Board 2016).

Unfortunately, there is no universally accepted definition of travel time reliability. The FHWA noted that travel time reliability measures the extent of unexpected delay to drivers. Specifically, they defined reliability as "the consistency or dependability in travel times, as measured from day-to-day and/or across different times of the day" (FHWA 2017). The Future Strategic Highway Research Program (F-SHRP) defined TTR as the variation in travel times over a period based on hour-to-hour or day-to-day variations (Cambridge Systematics, Texas Transportation Institute, University of Washington, and Dowling Associates 2003). The current sixth edition of the *Highway Capacity Manual* (HCM-6) states that "travel time reliability reflects the distribution of trip travel time over an extended period." This distribution is a function of a number of factors that influence travel time, including weather events, incidents, and work zones (Transportation Research Board 2016). The Strategic Highway Research Program (SHRP) program provided a broad definition:

> TTR aims to quantify the variation of travel time. It is defined using the entire range of travel times for a given trip, for a selected time period (e.g., the P.M. peak hour during weekdays) over a selected horizon (e.g., a year). For the purpose of measuring reliability, a trip can be defined as occurring on a specific segment, facility (combination of multiple consecutive segments), or any subset of the transportation network, or the definition can be broadened to include a traveler's initial origin and final destination. Measuring travel time reliability requires that a sufficient history is described by the travel time distribution for a given trip (Zegeer et al. 2014).

All the previous definitions use qualitative and nonstatistical terms (e.g., consistency, dependability, reflects the distribution, and variation of travel time) to define reliability. Therefore, many metrics can fit a given definition. Equally important, there are some commonalities among the definitions because they all recognize that travel time has a distribution that can vary as a function of a number of factors, and that reliability is related to the variability of this underlying distribution.

There are a number of TTD statistical measures, such as mean, median, standard deviation, and coefficient of variation (COV), that can be used to quantify how a system is performing. However, it has been argued that these are not easy for a nontechnical audience to understand, and that they may treat early and late arrivals with equal weight (FHWA 2017). As a result, TTR metrics, including the travel time index (TTI), the planning time index (PTI), the level of travel time reliability (LOTTR), and the buffer index (BI), have been developed. Although they have been recommended for use (FHWA 2012, 2017) these metrics tend to be ad hoc in nature and not based on statistical theory. A general argument for avoiding standard statistical measures is that the public does not understand them easily (Pu 2011). These concepts will be explored further in the Proposal for Future TTR Analyses section of the paper.

Review of Other COVID-19–Related Studies

It has been found that the COVID-19 pandemic has resulted in reductions in worldwide travel volumes across all modes of travel. For example, household trips were reduced by 50% across all modes in Australia (Beck and Hensher 2020), and a survey in the Netherlands found a 55% reduction in trips (de Haas et al. 2020). The Nebraska DOT (2020) reported that average volumes on all state highways decreased 29% compared with the previous 3-year average. According to Glanz et al. (2020), the COVID-19 stay-at-home orders resulted in the reduction of the average distance of daily travel from 8.0 to 1.6 km in the US. However, freight movement of essential supplies and food continued to be transported on the US National Highway System (NHS) (Hendrickson and Rilett 2020). Ironically, these shipments were completed more quickly because of the reduction in traffic congestion (Shaver 2020).

The reduction in traffic volume and distance traveled resulted in the improvement of roadway environmental performance measures such as air quality and noise. The COVID-19 pandemic caused cities to be quieter than before, decreased pollution in urban areas, and decreased fatalities, but not necessarily fatality rates. However, preliminary information indicates that the COVID-19 pandemic increased both psychological symptoms such as stress and anxiety (Tull et al. 2020) and alcohol and substance abuse (Vingilis et al. 2020) which may affect road safety. Not surprisingly, the National Highway Traffic and Safety Administration (NHTSA 2020) reported that the fatality rate per 100 million VMT is projected to have increased from 1.06 in the first half of 2019 to 1.25 in the same period during the COVID-19 pandemic. This was because the percentage decline in the number of fatal crashes was lower than the percentage decline in traffic volume.

It is evident that the COVID-19 pandemic directly affected the demand component of the transportation system and not the supply or physical infrastructure (Hendrickson and Rilett 2020). Therefore, the disruptions in travel provide an opportunity to examine TTR performance measures and identify which ones best capture the disruptions due to the COVID-19 pandemic. One of the objectives of this paper was to test the sensitivity of the different TTR metrics to the changes brought about by the COVID-19 pandemic. To the authors' knowledge, this type of study has not been conducted previously.

Data and Methodology

Recent advancements in data collection technology have made new and more-detailed sources of travel time information available to transportation agencies. Although some data sources are public agencies, the majority of the new, high-level travel time data sets are from private-sector sources. Examples of these latter data sources include StreetLight Data (2020), Iteris (2020), and INRIX (2020). The travel data are obtained from a variety of data collection devices, including GPS installed in cell phones, Bluetooth devices, Wi-Fi devices, and probe vehicles. The data may be provided at the individual vehicle or traveler level, but typically the data are provided at various levels of spatial and temporal aggregation.

This paper used INRIX travel time data to analyze the travel time on three arterial corridors in Lincoln, Nebraska, and one arterial corridor in Omaha, Nebraska. INRIX is responsible for developing the US National Performance Management Research Data Set (NPMRDS) for the US Federal Highway Administration. The NPMRDS currently is used by state DOTs and metropolitan planning organizations for research, operational, and performance analyses (Siddiqui and Dennis 2019). In Nebraska, the NPMRDS road network covers major highways and urban arterial roadways. Arterial roadway corridors are disaggregated into different segments of varying lengths. Major cross streets usually delimit the segments endpoints, and each segment has a unique NPMRDS Traffic Message Channel (TMC) location code.

INRIX data from March 1, 2020 through May 31, 2020 were used in this paper. This period was selected because it corresponds to the time when the impact of the COVID-19–related restrictions (e.g., closing of business, stay at home protocols, and so forth) first impacted travel in Nebraska. For comparison purposes, INRIX data were obtained for the same March–May period for 2018 and 2019. The analysis focused on the AM peak (7–10 a.m.) and PM peak (4–7 p.m.) periods because these periods experience the highest traffic volumes and congestion. Each period was analyzed over 15-min subperiods so that any dynamic changes in travel time could be identified and analyzed.

Fig. 1 shows aerial views of the four arterial corridors that were studied. The corridors ranged in length from 3.2 to 4.9 km (2.0 to 3.1 mi), experienced AADT values of 24,500–79,800 vehicles, and had levels of service as defined by the HCM6 that ranged from D to F. Complete details of the corridors were given by Murphy et al. (2020).

The INRIX travel time data that were provided consisted of 15-min average weekday travel times for each of the NPMRDS segments that made up a given corridor. These segment travel times were summed to provide the 15-min average travel time for the corridor. There were 16 unique travel time data sets for each year because there were four corridors, each corridor was bidirectional, and two peak periods were examined for each corridor. Because there were 3 years of data, a total

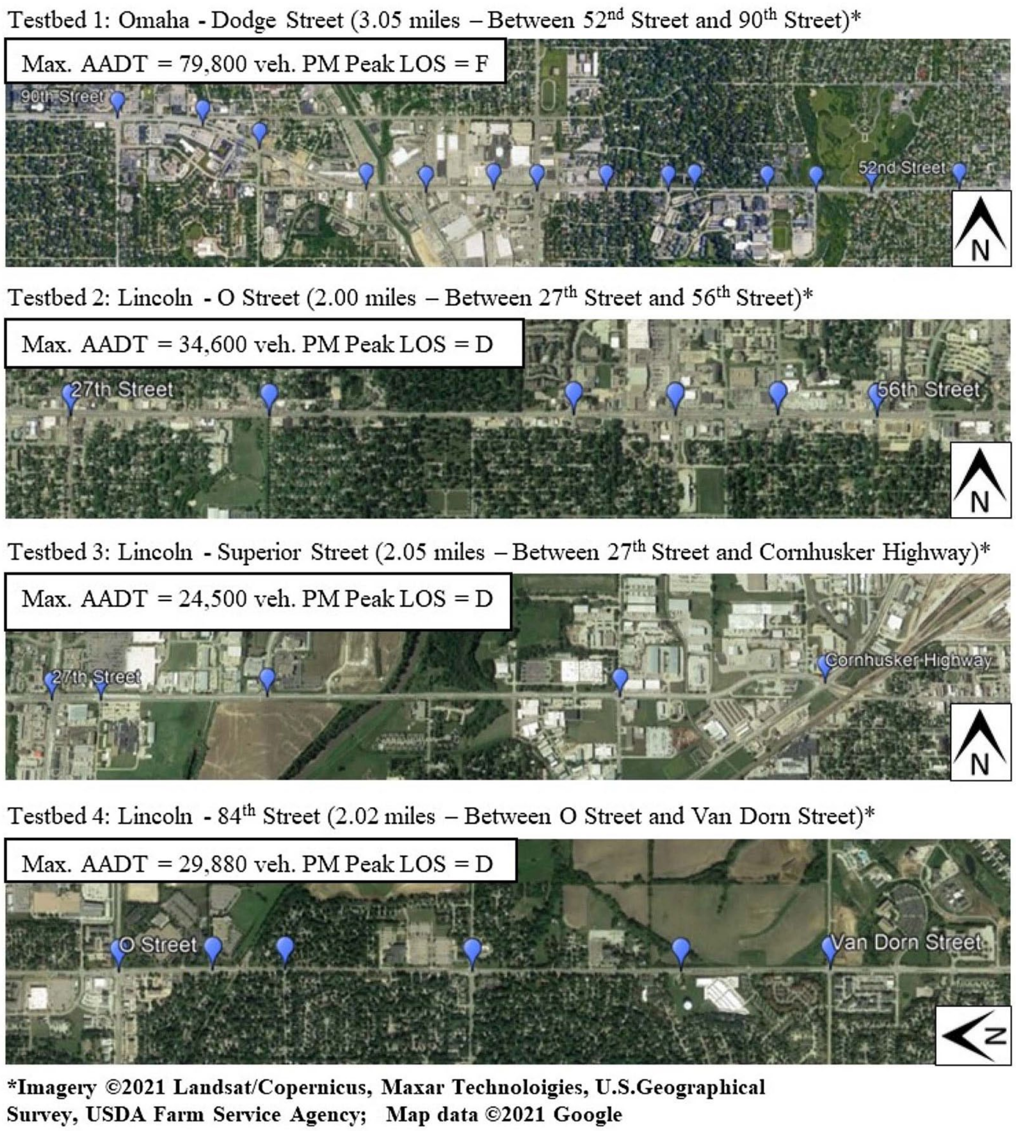

Fig. 1. Aerial maps of testbed corridors. (Map data © 2021 Google, © Landsat/Copernicus, Maxar Technologies, U.S. Geographical Survey, USDA Farm Service Agency.)

of 48 travel time distributions were obtained. In addition, there were 65 weekdays during each analysis period, and each 3-h peak period consisted of 12 15-min periods. Consequently, each of the 48 data sets consisted of a maximum of 780 15-min average travel time observations. Of the 37,440 (e.g., 780 × 48) 15-min periods studied, 1,440 had missing travel time information, and these periods were not considered in the analysis.

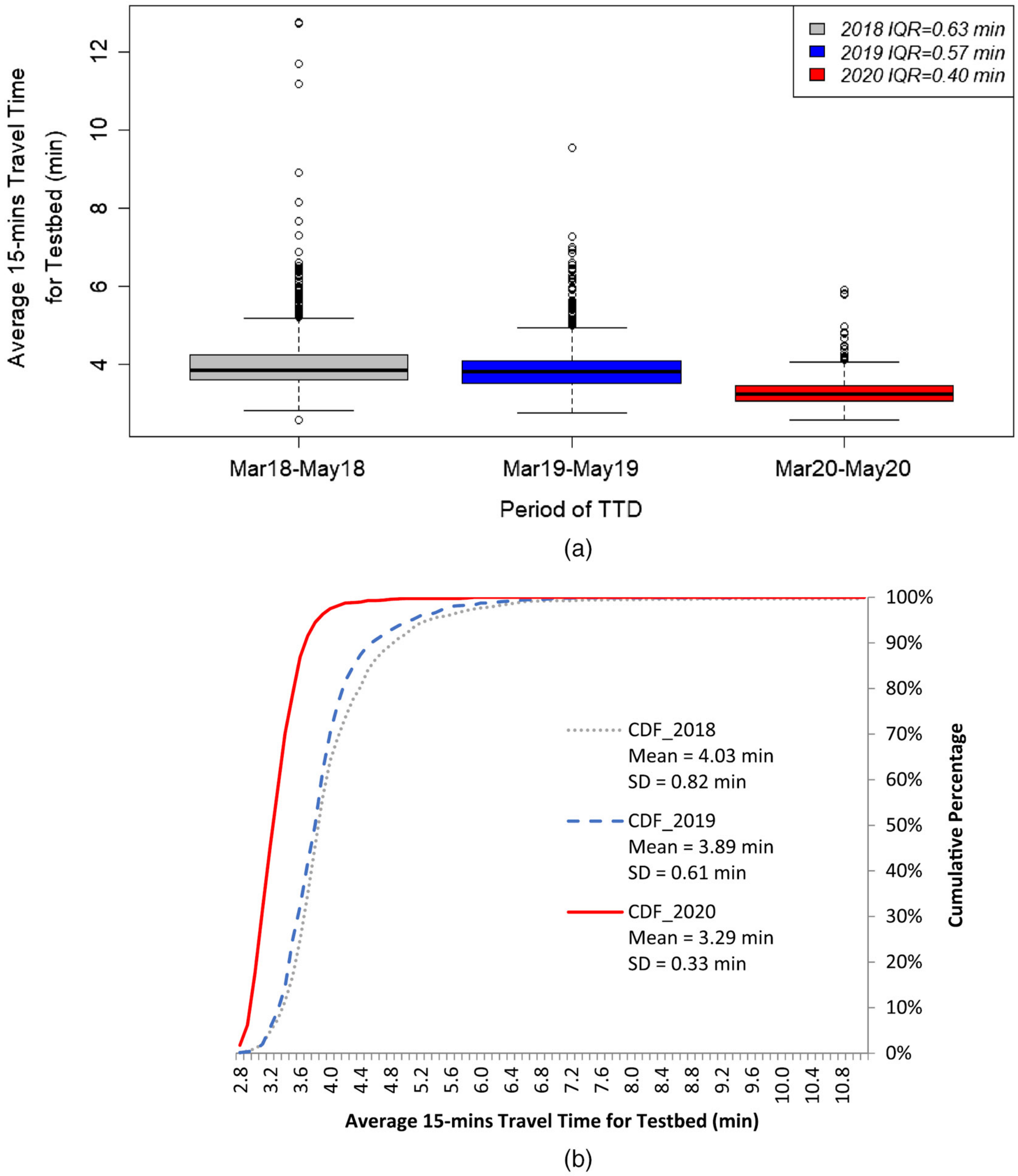

Fig. 2. Travel time distributions on O St. westbound AM peak (7–10 a.m.): (a) standard boxplot; and (b) CDF.

Comparison of Travel Time Distributions

Fig. 2(a) shows a standard boxplot of the average westbound travel times observed on the O Street corridor during the AM peak period. All 3 years in the analysis, 2018, 2019, and 2020, are shown. The top, middle, and bottom of each box plot represent the 75th percentile, the median, and the 25th percentile travel times, respectively. The upper and the lower boundaries are 1.5 times the interquartile range (IQR) (e.g., the difference between the 75th and 25th percentile travel times).

Not surprisingly, the mean and standard deviation values of the average TTD during the COVID-19 pandemic were lower than in prior years. Specifically, the average percentage difference between 2020 and the preceding years with respect to the mean and standard deviation values were −16.9% and −53.8%, respectively. In addition, the IQR in 2020 was 33.3% smaller than the IQR for 2018 and 2019. Because the travel times along the O Street test bed both were lower, on average, and had less variability than for similar periods in 2018 and 2019, it can be concluded that TTR improved in 2020, independent of which reliability definition is used.

Fig. 2(b) shows the cumulative distribution functions (CDF) of the average westbound O Street AM Peak travel time distributions for 2018, 2019, and 2020. The mean value and standard deviation (SD) of each of the CDFs are also presented. The CDF of the average TTD from the INRIX data during the onset of the COVID-19 pandemic in 2020 was considerably different from the average TTD in 2018 and 2019 [Fig. 2(b)]. Similar to Fig. 2(a), Fig. 2(b) shows that in 2020 the travel times along the O Street test bed were both lower (on average) and had less variability than for similar periods in 2018 and 2019.

Fig. 3(a) shows a standard boxplot of the average westbound travel times observed on the Dodge Street corridor during the AM peak. Not surprisingly, the mean and standard deviation values of the average TTD during the COVID-19 pandemic were, on average, 16.7% and 36.7% lower than in 2018 and 2019, respectively. In addition, the interquartile range of the average TTD, which is a measure of dispersion, in 2020 was 39.7% smaller than the interquartile range for 2018 and 2019. Similar to the O Street corridor, the travel times along the Dodge Street corridor both were lower, on average, and had less variability than for similar periods in 2018 and 2019.

Fig. 3(b) shows the cumulative distribution function (CDF) of the average westbound travel time distributions for 2018, 2019, and 2020 on the Dodge Street corridor. The CDF of the average TTD during the onset of the COVID-19 pandemic in 2020 was considerably different

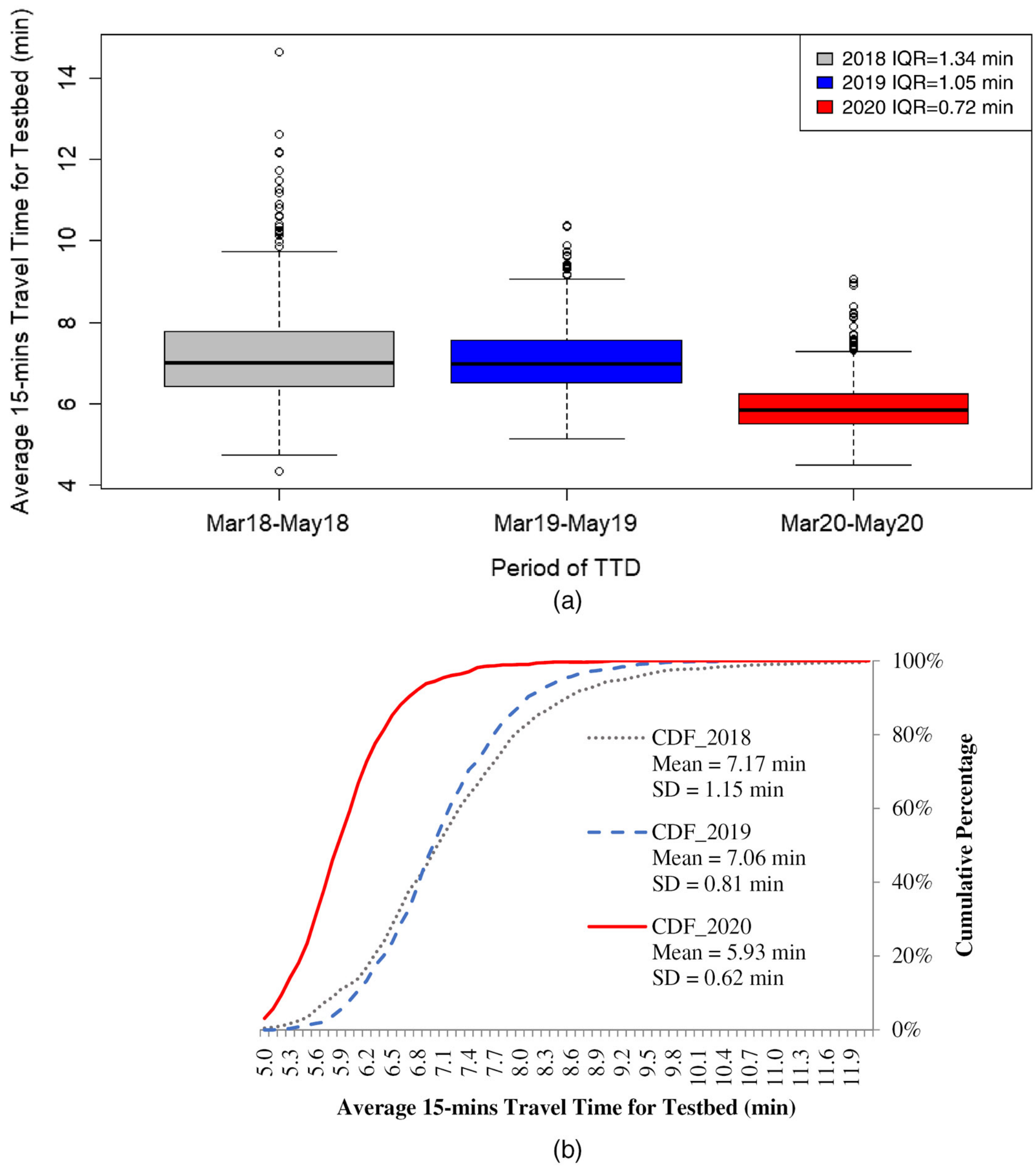

Fig. 3. Travel time distributions on Dodge St. westbound AM peak (7–10 a.m.): (a) standard boxplot; and (b) CDF.

from the average TTD from the 2018 and 2019 values. Similar to Fig. 3(a), Fig. 3(b) shows that in 2020 the travel times along the Dodge Street corridor both were lower, on average, and had less variability than for similar periods in 2018 and 2019.

Space constraints prevent publishing the boxplots and CDF distributions from the other 14 (e.g., 42 TTDs) scenarios that were analyzed. However, similar results were obtained for all 14 scenarios examined. In all cases, the average travel time CDF for 2020 was shifted to the left of the average travel time CDF for 2018 and 2019. In addition, for each scenario, the variance and range of the average travel times in 2020 were considerably smaller than those in 2018 and 2019. It is hypothesized that this result was a direct result of the lower traffic volumes that occurred because of the COVID 19 pandemic, which resulted in faster speeds and lower variability in speed along each of the four corridors.

A natural question is whether a given 2020 TTD was statistically different from the corresponding TTD in 2018 and 2019. To answer this question a Kolmogorov–Smirnov (KS) test was used to test the differences between cumulative distribution functions of the average TTDs for all 16 scenarios (e.g., 48 TTDs). This test was conducted at the 95% significance level. It was found that the differences in TTD across all 16 scenarios between 2020 and either of the preceding 2 years (e.g., 2018 or 2019) were statistically significant. In other words, the travel times on all four test corridors for both peak periods and both directions had lower means and smaller standard deviations during the COVID-19 pandemic than in prior years. These differences in TTDs were significant at the 95% level. Based on the generic reliability definitions listed earlier, it can be concluded that all 16 scenarios studied had more-reliable travel times in 2020 than in 2018 or 2019.

Comparison of Travel Time Distribution Metrics

Although it was not possible to show all 48 travel time distributions because of paper size limitations, it is possible to illustrate their statistical metrics. **Table 1** lists the mean and median travel times values, which are standard measures of central tendency, of all 48 TTDs that were analyzed. The columns correspond to the 16 scenarios, whereas the rows indicate the particular metric (e.g., mean and median) for a given year. As would be expected based on the preceding analysis, all 16 mean travel times observed in 2020 decreased compared with those

Table 1. Measures of central tendency

	Dodge Street				O Street				Superior Street				84th Street			
	EB		WB		EB		WB		EB		WB		NB		SB	
Year	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM
Mean (min)																
2018	6.36	8.35	7.17	8.45	3.99	4.40	4.03	4.77	3.74	4.00	3.92	3.92	3.32	3.40	3.25	3.25
2019	6.49	7.87	7.06	8.43	3.80	4.27	3.89	4.43	3.71	3.87	3.95	4.02	3.44	3.52	3.25	3.34
2020	4.95	6.80	5.93	6.48	3.36	3.63	3.29	3.60	3.46	3.48	3.54	3.60	3.02	3.07	3.00	3.01
Percentage difference (negative %)																
%Δ[a]	23.0	16.2	16.7	23.2	13.7	16.3	16.9	21.7	7.1	11.6	10.0	9.3	10.7	11.3	7.7	8.6
Median (min)																
2018	6.13	8.27	7.02	8.20	3.88	4.19	3.84	4.61	3.68	3.81	3.81	3.86	3.28	3.33	3.19	3.22
2019	6.31	7.67	6.98	7.80	3.75	4.13	3.80	4.32	3.68	3.81	3.89	3.91	3.38	3.39	3.21	3.28
2020	4.89	6.74	5.85	6.26	3.32	3.60	3.24	3.50	3.43	3.44	3.49	3.57	3.00	3.04	2.97	2.99
Percentage difference (negative %)																
%Δ[a]	21.4	15.4	16.4	21.8	13.0	13.5	15.2	21.6	6.8	9.7	9.4	8.1	9.9	9.5	7.2	8.0

EB = eastbound; WB = westbound; NB = northbound; and SB = southbound.

a. Estimated as quotient of (1) difference between 2020 value and average of 2018 and 2019 values, and (2) average of 2018 and 2019 values.

in 2018 and 2019. The decrease ranged from 7.1% to 23.2%, and the average decrease was 14.0%.

Similarly, all 16 median travel times observed in 2020 decreased compared with those in 2018 and 2019. The decrease ranged from 6.8% to 21.8%, and the average decrease was 12.9%. The slight difference between the mean and median decreases (e.g., 14.0% and 12.9%) was attributable to the fact that the 2020 distributions had fewer outliers (e.g., travel times considerably higher than the mean) than those for 2018 and 2019, and the median is more robust to outliers than the mean (Rousseeuw and Hubert 2011).

For all 48 TTDs examined, the mean travel time was greater than the median travel time, indicating that the underlying travel time distributions were not symmetric but rather were skewed to the right. In addition, the average differences between the mean and median travel times were 2.6% and 2.4% in 2018 and 2019, respectively. This difference was reduced to 1.3%, or by approximately 50%, in 2020. It may be inferred that the travel time distributions became more symmetric in 2020. This phenomenon is explored subsequently.

Table 2 lists two standard measures of dispersion of the travel time distribution: the standard deviation, and the IQR. The skewness, which is a measure of symmetry, for all 48 TTDs also is shown.

Interestingly, the change in standard deviations from 2018 and 2019 compared with that in 2020 was much more severe than the changes in the measures of central tendency (e.g., mean and median). In particular, the decrease in standard deviation between 2020 and the years 2018 and 2019 ranged from 16.5% to 69.4%, with an average decrease of 43.4%. The IQR sometimes is preferred as a measure of dispersion because it is robust to outliers (Rousseeuw and Hubert 2011). The differences in IQR ranged from −67.5% to 10.3%, with an average IQR decrease of 37.5%. Westbound Superior Street experienced a modest increase in IQR during the pandemic, which is in contrast to the other 14 scenarios in which the IQR decreased. The overall average decrease in IQR (37.5%) and standard deviation (43.4%) indicates that in 2020 there were considerably fewer periods with higher-than-normal travel times (e.g., fewer outliers). This was confirmed visually through an examination of all 48 boxplots.

Because the mean and standard deviation for all 16 scenarios (e.g., all four corridors, both directions, and both peak periods) decreased

Table 2. Standard measures of dispersion and symmetry

	Dodge Street				O Street				Superior Street				84th Street			
	EB		WB		EB		WB		EB		WB		NB		SB	
Year	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM
Standard deviation (min)																
2018	1.12	1.48	1.15	1.84	0.56	0.73	0.82	0.82	0.41	0.68	0.49	0.35	0.36	0.34	0.29	0.31
2019	1.43	1.30	0.81	2.33	0.38	0.56	0.61	0.60	0.33	0.43	0.41	0.44	0.35	0.53	0.25	0.30
2020	0.39	0.64	0.62	1.12	0.27	0.30	0.33	0.44	0.27	0.30	0.36	0.33	0.17	0.29	0.15	0.12
Percentage difference (negative %)																
%Δ[a]	69.4	54.0	36.7	46.3	42.6	53.5	53.8	38.0	27.0	45.9	20.0	16.5	52.1	33.3	44.4	60.7
Interquartile range (min)																
2018	1.32	1.89	1.34	2.34	0.54	0.88	0.63	0.92	0.47	0.60	0.44	0.36	0.31	0.34	0.25	0.24
2019	0.98	1.53	1.05	3.26	0.47	0.61	0.57	0.74	0.34	0.38	0.41	0.42	0.37	0.41	0.24	0.31
2020	0.50	0.89	0.72	0.91	0.33	0.33	0.40	0.37	0.37	0.35	0.46	0.43	0.18	0.18	0.16	0.12
Percentage difference (negative %)																
%Δ[a]	56.5	48.0	39.7	67.5	34.7	55.7	33.3	55.4	8.6	28.6	8.2	10.3	47.1	52.0	34.7	56.4
Skewness																
2018	1.81	0.46	1.24	1.16	1.65	1.31	4.31	1.29	1.66	2.87	2.02	1.21	4.18	1.71	3.63	8.09
2019	5.99	2.32	0.65	1.44	0.99	1.52	2.36	1.34	2.50	2.50	2.01	2.24	2.75	4.94	1.65	2.64
2020	0.65	0.35	1.08	3.66	1.08	1.11	2.01	2.60	0.59	1.08	0.89	0.44	2.00	6.00	1.31	1.48
Percentage difference (negative %)																
%Δ[a]	83.3	74.8	14.3	181.5	18.2	21.6	39.7	97.7	71.6	59.8	55.8	74.5	42.3	80.5	50.4	72.4

EB = eastbound; WB = westbound; NB = northbound; and SB = southbound; italics indicate positive change.

a. Estimated as the quotient of (1) the difference between the 2020 value and the average of the 2018 and 2019 values, and (2) the average of the 2018 and 2019 values. Note that numbers in italics indicate a positive change.

in 2020, it can be argued that the corridors became more reliable with respect to travel time because of the pandemic. A driver would experience lower average travel times and these average travel times would have much less variability in 2020 than in 2018 or 2019, all else being equal. This observation is analyzed further in the "Comparison of Travel Time Reliability Metrics" section.

All 48 TTDs were skewed positively, and this finding is typical of most travel time distributions (Tufuor et al. 2020; Tufuor and Rilett 2019; Mahmassani et al. 2014). On average, the COVID-19 pandemic caused an 18.2% reduction in the skewness values. In other words, the 2020 TTDs in general were more symmetrical than those in 2018 and 2019. This confirms the observation that the differences between the mean and median travel time values were smaller in 2020 than in 2018 or 2019. Four of the 16 scenarios became less symmetrical. It was hypothesized that the overall move towards greater symmetry in the TTDs occurred because the number of extremely high travel times was reduced during the pandemic. This was true for the O Street and Dodge Street examples — the number of high travel times were reduced in 2020 [Figs. 2(a) and 3(a)]. This general pattern was confirmed in the other scenarios.

In addition, the testbeds with the highest annual average daily traffic values (e.g., Dodge Street and O Street) had the greatest reduction in the mean, median, standard deviation, IQR, and skewness values. This is not surprising, because these corridors are used extensively by commuters in the peak periods, and would be the most affected by the reduction in work travel caused by the pandemic.

Two tests were used to identify whether the differences that were identified previously were statistically significant. Welch's t-test was used to test the differences between the mean values. This test was selected because, in contrast to Student's t-test, Welch's t-test controls the Type I error when comparing unequal variance (Derrick et al. 2016). The difference between the median values was tested by using the Mann–Whitney–Wilcoxon rank-sum test (U test), which is a powerful nonparametric test (Landers 1981).

The results showed that for all 48 cases analyzed, the differences in mean travel time between 2020 and both 2019 and 2018 were statistically significant at the 95% significance level. Comparable results were found for the median values. This is not surprising, because similar conclusions were found previously when comparing the empirical TTDs using a KS test.

In summary, the results indicated that travel times on all four corridors decreased because of the pandemic, and these decreases were statistically significant. In addition, the measures of dispersion values were reduced, and these reductions were greater than the reductions in the measures of central tendency. It was hypothesized that this occurred because there were fewer outliers (e.g., periods of higher travel times) in 2020 than in 2018 and 2019. In other words, the lower traffic volumes associated with the COVID-19 pandemic had a greater impact on travel time measures of dispersion than on measures of central tendency. Lastly, the travel time reliability on all four test beds for both peak periods and both directions improved in 2020, regardless of which qualitative reliability definition was used.

Comparison of Travel Time Reliability Metrics

The characteristics of distributions and their various statistical-based metrics are well known and have received considerable study in the field of statistics (Spiegelman et al. 2011). Interestingly, although transportation agencies have access to comprehensive TTDs, they have gravitated to the use of nonstatistical metrics to described reliability. It is the authors' belief that this has occurred because of the qualitative terminology used to define reliability. Without specific quantitative descriptions, transportation professionals can interpret reliability terms such as consistency, dependability, and variations in a variety of ways—and have, as is described subsequently.

The common reliability metrics described in this section all are based on the underlying TTD and/or key statistical metrics related to the TTD. In fact, the HCM6 arterial travel time reliability methodology explicitly estimates the TTD first, and subsequently uses this TTD to identify TTR metrics. Although many TTR metrics can be used for arterial roads, there is no clear answer to the question of which metric is best for describing reliability. One of the few benefits of the 2020 pandemic is that it has created a natural experiment in which these types of research questions can be explored in detail. As described in the preceding section, the pandemic had a profound and statistically significant effect on travel times for the 16 scenarios examined on the test corridor. This section explores which, if any, of the TTR metrics best captures this change.

One of the most widely used TTR metrics is the travel time index [Eq. (1)] (Schrank et al. 2019; FHWA 2017). The TTI is a ratio of two different TTD metrics. The numerator is the mean travel time for the period, corridor, and direction under consideration [e.g., weekday westbound (WB) PM peak period on Dodge Street]. The denominator is the free-flow travel time, typically measured in the off-peak, on the same corridor for the same conditions. This paper estimated the free-flow travel time using the methodology presented in the current sixth edition of the *Highway Capacity Manual* (Transportation Research Board 2016). This approach is deterministic, and results in a static value for each scenario. The higher the TTI value, the greater is the difference in the free-flow travel time (e.g., a TTI of 4 means that that the average travel time will be 4 times the free-flow travel time). This metric often is used to analyze potential operational improvements (Schrank et al. 2019). Because the TTI measures how the mean travel time relates to free flow conditions, it implicitly assumes that a more reliable travel time is one that is closer to the free-flow travel time, all else being equal.

$$\mathrm{TTI}_{tcd} = \frac{\mu_{Ttcd}}{\mu_{Ftcd}} \qquad \forall\ t = 1, 2;\ \ \forall\ c = 1, 4;\ \ \forall\ d = 1, 2 \tag{1}$$

where TTI_{tcd} = travel time index for period t, corridor c, and direction d; t = period (1 = AM peak, 2 = PM peak); c = corridor (1 = O Street, 2 = Superior Street, 3 = 27th Street, and 4 = Dodge Street); d = direction [1 = eastbound (EB) or northbound (NB), 2 = WB or southbound (SB)]; μ_{Ttcd} = mean TTD for period t, corridor c, and direction d (s); and μ_{Ftcd} = mean free flow TTD for period t, corridor c, and direction d (s). Free-flow travel time usually is measured under uncongested conditions, and it often is assumed to be deterministic and static.

The planning time index [Eq. (2)] is of the same format as the TTI in that both are ratios. The difference is that in the PTI, the 95th percentile travel time replaces the mean travel time in the numerator. This metric compares the near-worst-case travel time to the free-flow travel time conditions and, as the name implies, attempts to capture how much extra time a traveler should plan to add to their trip, compared with free flow conditions, to ensure an on-time arrival (FHWA 2017).

$$\mathrm{PTI}_{tcd} = \frac{P_{95tcd}}{\mu_{Ftcd}} \qquad \forall\, t = 1, 2;\ \ \forall\, c = 1, 4;\ \ \forall\, d = 1, 2 \tag{2}$$

where PTI_{tcd} = planning time index for period t, corridor c, and direction d; and P_{95tcd} = 95th percentile travel time for period t, corridor c, and direction d (s).

Table 3 lists the TTR metrics calculated for the 48 periods analyzed in this paper. In 2020, the TTI values decreased by 7.2%–23.4%, with an average reduction of 14.0%. The PTI values decreased by 9.4%–33.7%, with an average reduction of 19.7%. These results are not surprising, because although both the travel time mean and travel time standard deviation decreased, the latter had the greatest change. Therefore, the PTI, which is related to the spread or dispersion of the distribution, had a greater percentage drop than the TTI. The key to understanding why the PTI had a greater change than the TTI lies in the underlying travel time distributions and their associated statistics, as shown previously.

In addition, the percentage difference in means across all 3 years were exactly the same as the percentage differences in TTI, all else being equal. This can be seen by comparing the percentage change in mean travel time values in Table 1 and the percentage change in TTI values in Table 2. This occurs because the TTI is simply the mean divided by a constant. For similar reasons, the percentage difference in the 95th percentile values across all 3 years were the same as the percentage difference in PTI values. Both the TTI and PTI define reliability with respect to a single metric— in the case of the TTI it is the mean travel time, and in the case of the PTI it is the 95th percentile travel time. In summary, using the TTI and PTI metrics to compare changes in reliability over time (e.g., 2020 versus 2019) gives the exact same result as using the mean and 95th percentile travel time values, respectively.

The LOTTR [Eq. (3)] is the ratio of the 80th percentile travel time to the 50th percentile travel time (e.g., median travel time) for a given period, corridor, and direction under consideration.

Table 3. Travel time reliability metrics (unitless)

	Dodge Street				O Street				Superior Street				84th Street			
	EB		WB		EB		WB		EB		WB		NB		SB	
Year	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM
Travel time index (TTI)																
2018	1.47	1.93	1.63	1.92	1.37	1.51	1.38	1.63	1.33	1.42	1.40	1.40	1.16	1.19	1.14	1.14
2019	1.50	1.82	1.60	1.92	1.31	1.47	1.33	1.51	1.32	1.38	1.41	1.44	1.21	1.24	1.14	1.17
2020	1.15	1.57	1.35	1.47	1.15	1.25	1.12	1.23	1.23	1.24	1.26	1.29	1.06	1.08	1.05	1.06
Percentage change (negative %)																
%Δ[a]	23.0	16.2	16.7	23.2	13.7	16.3	16.9	21.7	7.1	11.6	10.0	9.3	10.7	11.3	7.7	8.6
Planning time index (PTI)																
2018	1.94	2.56	2.10	2.66	1.71	2.05	1.82	2.15	1.59	1.88	1.73	1.64	1.34	1.41	1.30	1.27
2019	1.81	2.40	1.93	2.89	1.56	1.85	1.73	1.89	1.50	1.63	1.64	1.69	1.40	1.49	1.30	1.36
2020	1.32	1.83	1.60	1.84	1.33	1.44	1.31	1.51	1.40	1.43	1.50	1.50	1.17	1.17	1.15	1.13
Percentage change (negative %)																
%Δ[a]	29.7	26.3	20.4	33.7	18.5	26.2	26.2	25.2	9.2	18.5	10.9	10.2	14.5	19.5	11.6	13.9
Level of travel time reliability (LOTTR)																
2018	1.16	1.15	1.13	1.20	1.12	1.17	1.15	1.16	1.10	1.13	1.08	1.06	1.06	1.08	1.06	1.05
2019	1.10	1.14	1.10	1.33	1.09	1.12	1.09	1.13	1.06	1.06	1.08	1.09	1.08	1.10	1.06	1.07
2020	1.08	1.09	1.09	1.11	1.07	1.06	1.09	1.08	1.07	1.07	1.09	1.08	1.04	1.04	1.04	1.03
Percentage change (negative %)																
%Δ[a]	5.0	5.0	2.4	12.6	3.1	7.4	2.7	5.2	0.4	2.4	0.5	0.5	2.5	4.7	1.5	2.8
Buffer index (BI)																
2018	0.32	0.33	0.29	0.38	0.25	0.36	0.32	0.32	0.20	0.32	0.23	0.17	0.15	0.18	0.14	0.11
2019	0.20	0.32	0.20	0.51	0.19	0.26	0.30	0.25	0.13	0.18	0.16	0.18	0.16	0.21	0.14	0.16
2020	0.15	0.16	0.19	0.25	0.15	0.15	0.16	0.23	0.14	0.15	0.18	0.16	0.10	0.08	0.09	0.07
Percentage change (negative %)																
%Δ[a]	42.6	49.6	22.4	44.4	30.1	50.0	47.1	19.3	15.9	38.9	6.2	6.4	32.6	56.5	34.1	46.9
Coefficient of variation (COV)																
2018	0.18	0.18	0.16	0.22	0.14	0.17	0.20	0.17	0.11	0.17	0.13	0.09	0.11	0.10	0.09	0.10
2019	0.22	0.17	0.11	0.28	0.10	0.13	0.16	0.14	0.09	0.11	0.10	0.11	0.10	0.15	0.08	0.09
2020	0.08	0.09	0.10	0.17	0.08	0.08	0.10	0.12	0.08	0.09	0.10	0.09	0.06	0.09	0.05	0.04
Percentage change (negative %)																
%Δ[a]	60.3	45.0	24.0	30.0	33.1	44.4	44.3	20.5	21.4	38.7	11.1	7.8	46.4	24.6	39.8	56.9

EB = eastbound; WB = westbound; NB = northbound; and SB = southbound; italics indicate positive change.

a. Estimated as the quotient of 1) the difference between the 2020 value and the average of the 2018 and 2019 values, and 2) the average of the 2018 and 2019 values. Note that numbers in italics indicate a positive change.

$$\mathrm{LOTTR}_{tcd} = \frac{P_{80tcd}}{P_{50tcd}} \quad \forall\, t = 1, 2; \quad \forall\, c = 1, 4; \quad \forall\, d = 1, 2 \qquad (3)$$

where LOTTR_{tcd} = level of travel time reliability for period t, corridor c, and direction d; P_{80tcd} = 80th percentile travel time for period t, corridor c, and direction d (s); and P_{50tcd} = 50th percentile travel time (or median travel time) for period t, corridor c, and direction d (s).

One might be tempted to argue that the LOTTR is a measure of dispersion, because it is a direct function of a travel time range defined by the 50th to the 80th percentile travel times. It was found that for all 16 scenarios the correlation coefficient values between the LOTTR and the IQR and the LOTTR and standard deviation were 0.90 and 0.87, respectively. This demonstrates a high level of linear correlation between LOTTR and two of the most commonly used measures of dispersion (standard deviation and IQR). In essence, as the measure of dispersion increases for a given median travel time, so too does the LOTTR value.

The differences in LOTTR values ranged from −12.3% to 0.5%, with an average reduction of 3.5% (Table 3). The 2020 LOTTR values for the westbound a.m. and PM peak periods on Superior Street had marginal increases of approximately 0.9% and 0.5%, respectively. These results indicate that this corridor actually became less reliable in the AM and PM peaks during the pandemic in spite of the fact that both the mean and standard deviation of travel time for these scenarios decreased. The other 14 LOTTR values all experienced a decrease. Regardless, the LOTTR values had the least change of all the TTR metrics between 2020 and the nonpandemic years of 2018 and 2019. A closer examination of the data across all 16 scenarios found that the reductions in the 80th percentile and the 50th percentile travel times were approximately proportional. This is why the LOTTR values were relatively inelastic to the large reductions in volume associated with the pandemic.

Focusing solely on the LOTTR values might lead an analyst to conclude that the COVID-19 pandemic had only a marginal effect on travel time reliability on the test corridors. The correct interpretation is that the pandemic had a significant effect on travel times (e.g., both the mean and standard deviation decreased significantly), but its effects on the 50th percentile travel time and the 80th percentile travel time were approximately equal. The fact that the LOTTR values were relatively inelastic to the changes brought about by the pandemic is not a small issue, because this metric is used as the performance measure of the National Highway System. Both O Street in Lincoln and Dodge Street in Omaha are part of the NHS.

Interestingly, both the standard deviation and IQR had considerable differences between 2020 and the preceding years (e.g., 2018 and 2019). This was because the standard deviation and IQR are strict measures of dispersion, whereas the LOTTR is a ratio of two TTD metrics. Therefore, although the standard deviation and IQR are correlated to the LOTTR, they measure different attributes of the underlying distribution, and this becomes clearer when a comparison across years is conducted.

The buffer index [Eq. (4)] is also a ratio. The numerator is the difference between the 95th percentile travel time and the mean travel time, and the denominator is the mean travel time. The BI attempts to capture the extra time that an average traveler would need to add to the average travel time to ensure on-time arrival for 95% of all trips (FHWA 2017).

$$\mathrm{BI}_{tcd} = \frac{P_{95tcd} - \mu_{Ttcd}}{\mu_{Ttcd}} \quad \forall\, t = 1, 2; \quad \forall\, c = 1, 4; \quad \forall\, d = 1, 2 \qquad (4)$$

where BI_{tcd} = buffer index for period t, corridor c, and direction d.

The BI is aligned closely with the coefficient of variation [Eq. (5)]. The COV is the ratio of the TTD standard deviation to the TTD mean. Both the COV and the BI attempt to capture the same phenomena—the ratio of a metric representing the spread of the TTD to a metric representing the central tendency of the TTD. The COV of travel time has been used to quantify TTR (Pu 2011), although the USDOT has discouraged its use for reasons noted previously (FHWA 2017).

$$\mathrm{COV}_{tcd} = \frac{\sigma_{Ttcd}}{\mu_{Ttcd}} \quad \forall\, t = 1, 2; \forall\, c = 1, 4; \forall\, d = 1, 2 \qquad (5)$$

where COV_{tcd} = COV of travel time for time period t, corridor c, and direction d (s); and σ_{Ttcd} = standard deviation of travel time for time period t, corridor c, and direction d (s).

The BI values decreased by 7.7%–59.0%, with an average reduction of 35.0% (Table 3). In addition, of all four TTR metrics examined, the BI metric had the greatest percentage change in values because of the pandemic year. As discussed previously, the BI is a ratio of a measure of dispersion to a measure of central tendency. Because the decrease in the measures of dispersion was greater than the decrease in the measures of central tendency, it is not surprising that the BI had the greatest change in all TTR metrics. Similar results were found for the COV, the average reduction of which was −35.1%, with a range of −10.0% to −60.0%. The fact that the BI and COV gave similar results is not surprising, because both metrics measure essentially the same components of the underlying TTD.

In summary, the TTR metrics during the COVID-19 pandemic were, on average, lower than comparable TTR metrics for the 2018 and 2019 period (Table 3). In other words, travel times on the test corridors were more reliable during the pandemic. A natural question is, How much did travel time reliability improve during the pandemic? As shown previously, it is difficult to answer this question using the common TTR metrics. Specifically, the LOTTR metric indicated that the change in reliability was 3.5%, whereas the buffer index metric showed that the change in reliability was 35.0%. The TTI and PTI metrics indicated that the increase in reliability was between these two extremes, at 14.0% and 19.7%, respectively. This wide range of answers to a relatively straightforward question clearly is problematic from a user perspective.

All four metrics answered a slightly different variation of the same question. Specifically

- TTI: What is the change in the relationship between the mean travel time and the baseline free flow travel time, across the different years? This is the same question as: What is the difference in mean travel time across the different years? The answers to both questions are the same.
- PTI: What is the change in the relationship between 95th percentile travel time and the baseline free flow travel time, across the different years? This is the same question as: What is the difference in the 95th percentile travel time across the different years? The answers to both questions are the same.

- LOTTR: What is the change in the ratio of the 80th percentile travel time to the 50th percentile travel time across the different years? This essentially is asking if the rate of change of the 80th percentile change is less than, equal to, or greater to the change in the 50th percentile travel time. If the change in both metrics is approximately the same, the LOTTR does not change—as was found in this analysis.
- BI: What is the change in the ratio of the difference between the 95th percentile travel time and the mean travel time? This is a similar question to: What is the change in the ratio of a measure of dispersion (IQR or standard deviation) to a measure of central tendency (mean or median)? As shown, the BI and COV are highly correlated, and the answers to both questions are similar.

The answer of which is the best metric to use to measure reliability is up to the analyst. It is hypothesized that the preceding analysis also shows why no TTR metric has received widespread acceptance in the transportation community. All four metrics describe an answer to a different question related to reliability.

As shown previously, the 2020 travel time distributions were different from the comparable travel time distributions in 2018 and 2019, and these differences were statistically significant. Specifically, the 2020 TTD in general had reduced measures of central tendency (e.g., mean and median), reduced measures of dispersion (e.g., standard deviation and IQR), and reduced measures of symmetry (e.g., skewness). The authors contend that standard statistical measures describe changes in reliability over time as well, and arguably better, than do the standard TTR metrics.

Proposal for Future TTR Analyses

Because of the wide range of definitions of travel time reliability, a number of diverse TTR metrics have been developed. It is clear from the preceding analyses that none of these TTR metrics were able to capture, on their own, the change in reliability on the test corridors as a result of the COVID-19 pandemic. This was because the metrics answered different versions of the same question: Did travel time reliability change? If, similar to the MAP-21 legislation, one believed that the LOTTR was the best reliability metric then the answer is that travel time reliability experienced only a marginal change of 3.5%. If one believes that reliability is based on a ratio of a measure of dispersion to a measure of central tendency, then the BI and COV analyses indicate that TTR increased substantially— on the order of 35%. Alternatively, if one believes that reliability is related to the mean travel time or the 95th percentile travel time, then the TTI and PTI analyses indicate an answer that lies between these two extremes.

The truth is that if one wants to analyze a difference between two TTDs, then the most efficient way is to compare the distributions directly. As was shown in this paper, this can be accomplished easily. The benefit to this approach is that there are accepted techniques for measuring whether differences in distributions are statistically significant. The second-best approach is to analyze changes in various metrics associated with the distributions. As shown in this paper, the typical process is to examine the changes in (1) central tendency (mean, median, and so forth), (2) dispersion (standard deviation and IQR), and (3) symmetry. It can be argued that the change in TTD kurtosis also could be analyzed. Critically, all of the aforementioned metrics have accepted techniques for inferring whether any differences are statistically significant.

As this paper has demonstrated, measuring travel time reliability is actually a multiattribute decision-making problem. Based on the current definitions and associated TTR metrics, reliability clearly is a function of changes in the measures of central tendency (mean or median), changes in dispersion (standard deviation or IQR), changes in the relationship between measures of central tendency and measures of dispersion, and changes in symmetry (skewness). The key question is how much weight a given analyst places on each of these components.

Consequently, to identify the best TTR metric, it is necessary to know (1) what components of the TTD the end user considers important for identifying reliability, and (2) how much weight the end user puts on each of these components (e.g., improvement in a given measure of central tendency versus a given improvement in a measure of dispersion). Clearly, the developers of the TTI assigned zero weight to changes in the measure of dispersion. The developers of the LOTTR clearly felt that if both the 50th and 80th travel times improved (or became worse) at the same rate, then the corridor would not be identified as becoming more (or less) reliable.

Because of the aforementioned issues, the authors recommend that statistically based metrics be used for measuring travel time reliability in the future. The benefit of using commonly accepted statistical metrics is twofold. First, there is a rich literature on the characteristics of these measures, including how to test if differences are statistically significant and how to identify confidence intervals about the predicted values. Secondly, there are a large number of robust statistical measures that would be appropriate for travel time analysis which have not yet been used by the transportation community. For example, it is well known that travel times often are susceptible to outliers. There is a rich literature on metrics that are robust to outliers, and these would be natural candidates for new TTR metrics (Arachchige et al. 2020; Rousseeuw and Hubert 2011; Spiegelman et al. 2011).

The argument for developing and using nonstatistical TTR metrics is that the users have trouble understanding standard statistical metrics. However, it was shown that the TTR metrics are all highly correlated with existing statistical metrics. The authors contend that if users are able to understand the nonstatistical metrics, they also are capable of understanding the statistically based metrics.

The authors want to stress that adopting standard statistical metrics and techniques does not obviate the need for communicating appropriate information to end users. With respect to system users such as Transportation Systems Management and Operations (TSMO) operators, the authors argue that it is not too much to ask that these users be familiar with introductory statistical concepts as described in this paper. The ability to estimate confidence intervals for travel time forecasts, and to understand what they mean, would be invaluable to the engineers and analysts responsible for designing, planning, and operating the transportation system. With respect to the public, the authors agree that using statistical terms for communication is problematic. However, the authors also believe that the current TTR metrics are not much better. For example, if the transportation agencies in Lincoln and Omaha were to tell the traveling public that travel time reliability essentially was unchanged during the pandemic, as would be the conclusion based on the LOTTR analysis, there clearly would be a great deal of confusion. Most travelers would feel that reliability had improved because both average travel time and the variability of travel time was reduced significantly. The authors argue that the overarching goal should be to provide answers to users related to their specific needs. Specific questions, such as: When should I leave to be on time 95% of the time?, What will be the travel time corridor if I leave in 15 minutes?, and so forth, all can be answered using the travel time statistics and techniques advocated in this paper. The added benefit is that the transportation agencies can put confidence bounds on the answers. This might entail telling users that the agency cannot reasonably provide an answer because the confidence bounds are so large.

Concluding Remarks

This paper examined the effect of COVID-19 on travel time and travel time reliability on arterial roadways in Nebraska. Specifically, the travel time distribution of previous years (e.g., 2018 and 2019) within the same period (i.e., March–May) was compared with the travel time distribution during the COVID-19 pandemic in 2020. The paper also assessed the sensitivity of the TTR metrics to changes in the travel time distribution caused by the pandemic.

Four arterial roadways were used as the test corridors. The travel time data from INRIX on these corridors within the AM peak (7–10 a.m.) and the PM peak (4–7 p.m.) were used to analyze the travel time distributions and reliability metrics. A total of 16 scenarios, each with 3 TTDs for the years 2018, 2019, and 2020, were examined. It was found that during the COVID-19 pandemic

1. The 2020 TTDs were different from the equivalent 2018 and 2019 distributions, and these differences were statistically different at the 95% level of significance according to the KS test. In all 16 cases, the box plots showed that the measures of central tendency, the measures of dispersion, and the measure of symmetry all were reduced compared with those of the previous years, all else being equal.

2. The average TTD mean and standard deviation values for all 16 scenarios were reduced by an average of 14.0% and 43.4%, respectively. In general, the greatest changes occurred on the test corridors that had the highest prepandemic volumes. Not surprisingly, the test corridors could be considered more reliable during the pandemic regardless of which TTR definition was used.
3. Four standard travel time metrics were compared for all 16 scenarios. It was found that the travel time index, the planning time index, the level of travel time reliability, and the buffer index were reduced, on average, by 14.0%, 19.7%, 3.5%, and 35.0%, respectively. In other words, the question of whether the test corridors were more reliable during the pandemic was a function of which TTR metric was chosen. This was not surprising, because each TTR metric provides information on different components of reliability.
4. Interestingly, the LOTTR metric had the lowest percentage change. This was attributed to the fact that the LOTTR is a ratio and both the numerator (i.e., 80th percentile travel time) and the denominator (i.e., 50th percentile travel time) were reduced at approximately the same rate across all 16 scenarios. The USDOT has chosen the LOTTR metric for evaluation of TTR on the US National Highway System. This metric may be problematic for monitoring purposes, because it was found to be inelastic to relatively major changes in traffic volume.

This paper illustrated the importance of selecting appropriate metrics—and having a deep understanding of these metrics—when evaluating transportation systems. The authors argue in this paper that travel time reliability, as commonly defined in the US, has three components: changes in measures of central tendency, which are measured by metrics such as the mean, median, and TTI; changes in measures of dispersion, which are measured by metrics such as the standard deviation and PTI; and relational changes in the measure of dispersion to measures of central tendency, which are measured by metrics such as the BI and COV. It also can be argued that changes in symmetry (e.g., skewness) also are a component of reliability.

The authors argue in this paper that a better approach to measuring TTR and evaluating transportation systems is to use statistically based theory and practice. It was demonstrated that the most common TTR metrics used are highly correlated with standard statistical measurements. This fact is not surprising, because both the TTR metrics and the statistical measures often attempt to measure the same phenomena. For example, the TTI, PTI, and BI metrics are highly correlated with the mean, P_{95}, and coefficient of variation (COV), respectively. Because (1) the statistical measures are well-understood and documented, and (2) there are standard statistical tests that can be used to tell whether changes in travel time are statistically significant, the authors argue that these would be more appropriate metrics for measuring travel time reliability.

The rationale for using nonstatistical metrics is that the traveling public does not understand statistical metrics. The authors agree with this statement. However, the authors do not believe that the statistical measures advocated in this paper are too complicated for transportation engineers tasked with operating the transportation system, because the metrics are covered in most introductory statistical courses and textbooks (Spiegelman et al. 2011). More importantly, it is easy to translate these statistical measures into information that an average user of the system can understand (e.g., When should I leave?, How long will my trip take?, and so forth). The authors argue that using the approach advocated in this paper will lead to more-accurate information being provided to users—or even providing no information if the statistics indicate that the forecast bounds are such that the information would be of little value to the user.

Lastly, using the TTD to measure TTR is not a new concept. The HCM6 TTR methodology first estimates/forecasts a TTD, and then estimates TTR metrics from the TTD. The authors propose that when system operators evaluate reliability, they should use accepted, statistically based metrics rather than the ad hoc TTR metrics currently used. This does not preclude using other metrics for users (e.g., roadway users and decision makers) who may have difficulty understanding the statistical metrics. However, as stated previously, these metrics should be tailored to the user's specific needs, and would be based on sound, statistical theory.

Data Availability Some or all data, models, or code used during the study were provided by a third party, including the INRIX travel time data. Direct request for these materials may be made to the provider as indicated in the Acknowledgments.

Acknowledgments The authors would like to thank the Omaha Office of HDR for making INRIX travel time data for Lincoln, NE and Omaha, NE readily available. This research was made possible through the funding of the U.S. Department of Transportation and the Nebraska Department of Transportation [DPS-STWD-118 and DPS-STWD-118(a)], as well as the Nebraska Transportation Center. This support is gratefully acknowledged. The contents of this paper reflect the views of the authors, who are responsible for the facts and accuracy of the information presented herein and are not necessarily representative of the U.S. Department of Transportation, the Nebraska Department of Transportation, the City of Lincoln, or the City of Omaha.

References

Arachchige, C. N. P. G., L. A. Prendergast, and R. G. Staudte. 2020. “Robust analogs to the coefficient of variation.” *J. Appl. Stat.* 1–23. https://doi.org/10.1080/02664763.2020.1808599

Beck, M. J., and D. A. Hensher. 2020. “Insights into the impact of COVID-19 on household travel and activities in Australia—The early days of easing restrictions.” *Transp. Policy* 99 (Dec): 95–119. https://doi.org/10.1016/j.tranpol.2020.08.004

Cambridge Systematics, Texas Transportation Institute, University of Washington, and Dowling Associates. 2003. *Providing a highway system with reliable travel times: Study 3—Reliability*. Final Report, NCHRP Project 20-58(3). Washington, DC: Transportation Research Board of the National Academies.

de Haas, M., R. Faber, and M. Hamersma. 2020. “How COVID-19 and the Dutch ‘intelligent lockdown’ change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands.” *Transp. Res. Interdiscip. Perspect.* 6 (Jul): 100150. https://doi.org/10.1016/j.trip.2020.100150

Derrick, B., D. Toher, and P. White. 2016. “Why Welch’s test is Type I error robust.” *Quant. Methods Psychol.* 12 (1): 30–38. https://doi.org/10.20982/tqmp.12.1.p030

FHWA (Federal Highway Administration). 2012. “Moving ahead for progress in the 21st century.” Accessed August 23, 2020. http://www.fhwa.dot.gov/map21/

FHWA (Federal Highway Administration). 2015. “Fixing America’s surface transportation act.” Accessed August 23, 2020. https://www.fhwa.dot.gov/fastact/

FHWA (Federal Highway Administration). 2017. “Travel time reliability: Making it there on time, all the time.” Accessed September 23, 2020. https://ops.fhwa.dot.gov/publications/tt_reliability/TTR_Report.htm

Glanz, J., B. Carey, J. Holder, D. Watkins, J. Valentino-DeVries, R. Rojas, and L. Leatherby. 2020. “Where America didn’t stay home even as the virus spread.” Accessed July 30, 2020. https://www.nytimes.com/interactive/2020/04/02/us/coronavirus-social-distancing.html

Hendrickson, C., and L. R. Rilett. 2020. “The COVID-19 pandemic and transportation engineering.” *J. Transp. Eng. Part A Syst.* 146 (7): 01820001. https://doi.org/10.1061/JTEPBS.0000418

INRIX. 2020. “Intelligence that moves the world.” Accessed February 12, 2020. https://inrix.com/about/

Iteris. 2020. “ClearGuide: Real-time contextual mobility intelligence.” Accessed July 15, 2020. https://www.iteris.com/clearguide

Landers, J. 1981. *Quantification in history, topic 4: Hypothesis testing II-differing central tendency*. Oxford: All Souls College.

Lomax, T. J., and D. L. Schrank. 2002. *Using travel time measures to estimate mobility and reliability in urban areas*. No. FHWA/TX-02/1511-3. College Station, TX: Texas A&M Transportation Institute.

Mahmassani, H. S., J. Kim, Y. Chen, Y. Stogios, A. Brijmohan, and P. Vovsha. 2014. *Incorporating reliability performance measures into operations and planning modeling tools*. Washington, DC: Transportation Research Board.

Murphy, S., E. Tufuor, and L. Rilett. 2020. *A comparison of empirical data to the highway capacity manual estimated travel time distribution on urban streets*. Washington, DC: Transportation Research Board.

Nebraska DOT. 2020. “Media release on the impact of COVID-19.” Accessed July 27, 2020. https://dot.nebraska.gov/media/113416/3-31-2020-covid19-impactspdf.pdf

NHTSA (National Highway Traffic and Safety Administration). 2020. “Early estimates of motor vehicle traffic fatalities.” Accessed October 9, 2020. https://crashstats.nhtsa.dot.gov/?_ga=1.195381676.1482461981.1490551667#!/PublicationList/51

Pu, W. 2011. “Analytic relationships between travel time reliability measures.” *Transp. Res. Rec.* 2254 (1): 122–130. https://doi.org/10.3141/2254-13

Reid, J. 2004. *Unconventional arterial intersection design, management and operations strategies*. Charlotte, NC: Parsons Brinckerhoff Quade & Douglas.

Rousseeuw, P. J., and M. Hubert. 2011. “Robust statistics for outlier detection.” *Wiley Interdiscip. Rev.: Data Min. Knowl. Discovery* 1 (1): 73–79. https://doi.org/10.1002/widm.2

Schrank, D., T. Lomax, and B. Eisele. 2019. “2019 urban mobility report—Appendix A.” Accessed December 20, 2020. http://mobility.tamu.edu/ums/report

Shaver, K. 2020. “As coronavirus precautions take hold, large US cities see rush hour traffic jams vanish.” Accessed April 19, 2020. https://www.washingtonpost.com/transportation/2020/03/18/coronavirus-precautions-take-hold-many-cities-see-heavy-rush-hour-traffic-congestion-vanish/

Siddiqui, C., and M. Dennis. 2019. "Developing a guideline for setting targets for national performance management measures to assess performance of the national highway system in South Carolina." *Transp. Res. Rec.* 2673 (9): 266–276. https://doi.org/10.1177/0361198119843863

Spiegelman, C., E. S. Park, and L. R. Rilett. 2011. *Transportation statistics and microsimulation*. Boca Raton, FL: CRC Press.

StreetLight Data. 2020. "Understand the impact of COVID-19 on traffic, travel patterns, toll revenues and more." Accessed July 27, 2020. https://www.streetlightdata.com/covid-transportation-metrics/

Transportation Research Board. 2016. *Highway capacity manual: A guide for multimodal mobility analysis*. 6th ed. Washington, DC: Transportation Research Board.

Tufuor, E., L. R. Rilett, and L. Zhao. 2020. "Calibrating the Highway Capacity Manual arterial travel time reliability model." *J. Transp. Eng. Part A Syst.* 146 (12): 04020131. https://doi.org/10.1061/JTEPBS.0000451

Tufuor, E. O. A., and L. R. Rilett. 2019. "Validation of the *Highway Capacity Manual* urban street travel time reliability methodology using empirical data." *Transp. Res. Rec.* 2673 (4): 415–426. https://doi.org/10.1177/0361198119838854

Tull, M. T., K. A. Edmonds, K. M. Scamaldo, J. R. Richmond, J. P. Rose, and K. L. Gratz. 2020. "Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life." *Psychiatry Res.* 289 (1): 1–6. https://doi.org/10.1016

Vingilis, E., et al. 2020. "Coronavirus disease 2019: What could be the effects on Road safety?" *Accid. Anal. Prev.* 144 (Sep): 105687. https://doi.org/10.1016/j.aap.2020.105687

Yang, S., and P. Cooke. 2018. "How accurate is your travel time reliability?—Measuring accuracy using bootstrapping and lognormal mixture models." *J. Intell. Transp. Syst.* 22 (6): 463–477. https://doi.org/10.1080/15472450.2017.1421075

Zegeer, J., et al. 2014. *Incorporating travel time reliability into the Highway Capacity Manual*. SHRP 2 Report S2-L08-RW-1. Washington, DC: National Research Council, Transportation Research Board.