Effect of Signal Control on Bimodal Travel Time Distributions

Noah J. Goodall, Ph.D., P.E.
Research Scientist
Virginia Transportation Research Council
Charlottesville, VA
Email: noah.goodall@vdot.virginia.gov

Presented at the 97th Annual Meeting of the Transportation Research Board, January 2017

Abstract

Vehicles traveling along interrupted flow facilities often exhibit travel times with bimodal distributions. The characteristics of these distributions have been studied extensively in the literature, yet the effect of signal control on bimodality have received little attention. Most researchers theorize that the slower group experiences signal delay while the faster group does not. We investigate the effect of signal control on bimodal distributions, specifically the difference between average travel times of two groups of vehicles. Testing is performed in simulation and compared against field results of both sampled data using Bluetooth sensors as well as non-sampled data from the NGSIM vehicle trajectories. In coordinated corridors with platooning, we find that red time is a strong indicator of gap between average travel times.

INTRODUCTION

Travel time is one of the most widely used metrics for road network performance. It has become a standard metric not only in internal assessments, but also for the public in the form of changeable message signs, 511 traveler information systems, and mobile navigation applications.

Travel time is collected through periodic tracking of a sample of vehicles in the traffic stream. Examples include positions of fleet vehicles such as taxis, buses, and trucks using data from Global Positioning Systems (GPS), vehicle license plate recognition, and toll tag transponders. Recently, individual vehicles can be re-identified via media access control (MAC) addresses unique to a mobile phone or other Bluetooth (*1*) or Wi-Fi (*2*) enabled devices.

When measured along signalized intersections and corridors, travel times have been observed to demonstrate bimodality. Whereas freeways typically have a clear average travel time, arterials often exhibit two or more "typical" travel times, sometimes separated by 30 seconds or more. While there has been much work in establishing the bimodality of travel times and fitting the data to various bimodal distributions, there has been little research into determining how the characteristics of the travel time distributions are affected by signal control. Researchers often assumed that the faster distribution was unaffected by signal control, and that the slower distribution was affected, rarely do researchers establish a firm connection to signal control its impact on travel time distributions.

The objective of this research is to investigate the relationship between signal control and the difference in average travel time between two populations in a bimodal distribution. The remainder of this paper will provide an overview of relevant literature, a discussion of the theory behind travel time bimodality, background of data sources and analysis, findings, and recommendations for future research.

LITERATURE REVIEW

While travel times were initially assumed to fit a normal distribution, Wardrop first proposed that they instead followed a skewed distribution (*3*). Later researchers suggested lognormal (*4*), gamma (*5*), Weibull (*6*), and Burr distributions (*7*).

As new technologies allowed the easier collection of travel times across short arterial segments, researchers began to observe that data often failed to fit into unimodal distributions. Davis and Xiong (*8*) used license plate matching along signalized arterials in Minneapolis, Minnesota, and found evidence of bimodal distributions. Susilawati et al. (*9*) found that bimodal travel times were more likely to occur on shorter links with coordinated signal control, but did not investigate the impact of the traffic control scheme on the distribution. Other research has focused on model development (*10*), and the relationship between bimodal travel time distributions and travel time reliability (*11*, *12*) and congestion identification (*13*).

Most efforts have demonstrated statistically that arterial travel times are better described with a bimodal than unimodal distribution. Several suggest that the bimodality is due to signal control – that while the faster population drives through the corridor without stopping, the slower population encounters at least one red signal (*7, 13*). Other studies have observed the impact of signal control directly (*8, 13*). Ji et al. (*10*) studied the signal timing plans along their test corridor and found that their bimodal travel time model had a greater mixture of its two components when the signal was red for a greater percentage of the cycle. The authors suggest that the percentage of red time is directly correlated to the percentage of vehicles stopped at a red light and therefore appearing in the slower of the two populations. No researchers have studied the relationship between signal control and average travel time differences between the modes, although Young et al. (*14*) have proposed that the difference is equivalent to the cycle length of the coordinated signal system.

THEORETICAL FRAMEWORK

Many researchers suggest that the slower population of vehicles represent those that have slowed or stopped due to signal control. This would imply that there is a strong correlation between some aspect of signal control and the difference between the average travel time of each population. Young et al. (*14*) suggests that this difference in travel times is equivalent to the cycle length of the signal. This seems unlikely. For a vehicle to experience a full cycle length of delay at single signal requires that the vehicle wait through not only the red phase but also most of the following green phase.

Uncoordinated Signalized Intersections

At an isolated or uncoordinated signalized intersection, where arrival rates follow a Poisson distribution, vehicles will arrive throughout the red phase. Figure 1 provides an example of this type of arrival. While vehicles arriving during the green phase experience no delay, vehicles arriving during the red phase experience different amounts of delay depending on how early in the cycle they arrived. The maximum amount of delay a vehicle could experience is the red phase, startup time, and queue clearance time (and assuming no cycle failure). One would expect the average difference between the vehicles experiencing no delay and those experiencing some delay to be equivalent to half the length of the red phase, plus half the average queue clearance time.

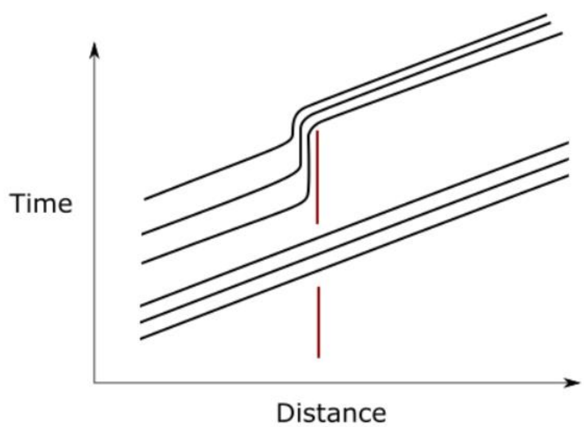

FIGURE 1 Time-space diagram of an isolated intersection

Ji et al. (*10*) studied bus movements through a single isolated signalized intersection at two locations in Columbus, Ohio. Their findings seem to support the idea that vehicles in the slower population experience a random sample of the red time, and therefore the average delay is close to half the length of the red phase. The slower population in Link 1 experienced 13.7 and 17.1 seconds of delay in addition to what was expected from random arrivals, while the slower population in Link 2 experienced 0.3 seconds of additional delay.

TABLE 1 Delay and red phase lengths at an isolated intersection

	Link 1		Link 2	
	AM	PM	AM	PM
Red phase *(s)*[1]	85.0	71.0	63.0	36.0
Expected delay *(s)*[1]	56.2	52.6	31.8	18.3
Estimated lost time *(delay - 0.5 × red phase)*	13.7	17.1	0.3	0.3

[1] From Ji et al. (*10*)

Coordinated Signalized Intersections

Travel times are often collected on corridors with coordinated traffic control. On these corridors, signals communicate directly or are set to a common clock to ensure that vehicles receive green phases as they progress through the network at the design speed. Approximately 60% of traffic signals in North America are coordinated, and are often deployed on high volume roads (*15*). When signals are coordinated, vehicles travel through the corridor in groups known as platoons. As drivers maintain different speeds, platoon disperse as they move through the corridor, with the slowest ones falling below the design speed and reaching a red signal. Vehicles often encounter the red phase as it is beginning, as demonstrated in Figure 2.

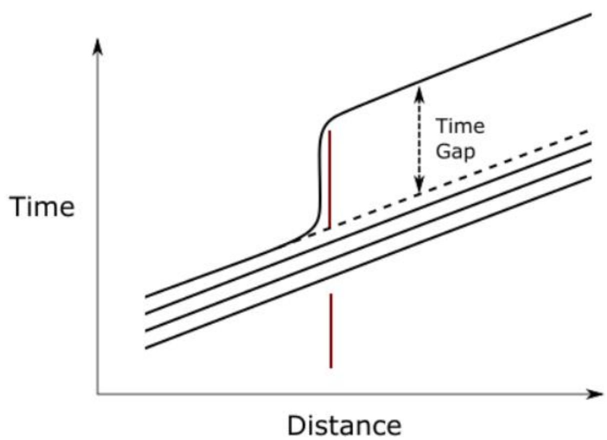

FIGURE 2 Time-space diagram of a coordinated intersection

For these signals, in undersaturated conditions and with good platoon formation, the vehicles in the slower population have stopped at the beginning of the red phase, and have experienced the additional delay of decelerating to a stop, waiting for nearly the entire red phase, and accelerating back up to speed. Vehicles in the slower population, therefore, would be expected to have travel times that are approximately as long or slightly longer than the red phase. When there are several different red phase lengths along a signalized corridor, the slower population will consist of vehicles that have been stopped at different length red phases. It would be expected that the slower population will have a larger variance in travel time, as they experienced different delays.

Over longer corridors, it becomes increasingly less likely for a vehicle to progress through without stopping for a red signal.

SIMULATION

To test various signal control strategies in a controlled environment, simulation was used to measure travel times of vehicles across a hypothetical network. The test network consisted of a three-lane, one-way road with four traffic signals. The first traffic signal was used as a meter to create vehicle platoons, as would be expected in a standard coordinated traffic signal system. Travel time measurements began after the first signal and concluded after the fourth. Signals were positioned at 2000-foot (610-meter) intervals with coordinated, fixed time control. Travel times were collected for one hour following a five-minute warmup. Each simulation was repeated ten times using different random seeds. The traffic simulation software VISSIM was used, and vehicle behavior followed the default settings. To isolate the effects of signal control, side street traffic was not modeled.

Travel times were sorted into fast and slow populations using k-means clustering as implemented in the GNU Octave programming language (*16*). In this approach, n observations are divided into k sets S, minimizing the sum of distance functions of each point x in the set to a set's centroid μ. In this analysis, $k = 2$.

$$\sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2$$

Four scenarios were tested, shown in Table 2. The time between cluster centroids, that is, the travel time difference between the two populations, is very close the red time. The time over the red time, about 5.4 seconds in scenarios A-C, represents the time required for a vehicle to accelerate up to the desired speed of 43.5 mi/hr (70 km/hr). From observing the simulation, vehicles exhibited tight platooning, with only the rear-most vehicles unable to stay within the green progression band.

TABLE 2 Timing Plans and Travel Time Differences in Microscopic Simulation

Scenario	Cycle Length *(s)*	Red Time *(s)*	Red Time of Middle Signal *(s)*	Time Between Cluster Centroids *(s)*
A	60	30	30	35.4
B	60	40	40	45.4
C	90	60	60	65.3
D	60	30	40	39.5

Because the simulation ignored traffic from cross streets, pedestrians, parking maneuvers, and other impediments to flow, it represents an ideal environment for coordination. As expected, it exhibits two populations of vehicles, each with very small standard deviations. Figure 3 is a histogram of the simulated travel times. Scenario D is of particular interest. In this scenario, the second of three signals used a red time of 40 seconds, whereas the other signals had a red time of 30 seconds. This has the effect of a larger variance among the slower population, as some vehicles are stopped at a signal for 30 seconds while others are stopped for 40 seconds. It should also be noted that scenario D used lower volumes to prevent cycle failure.

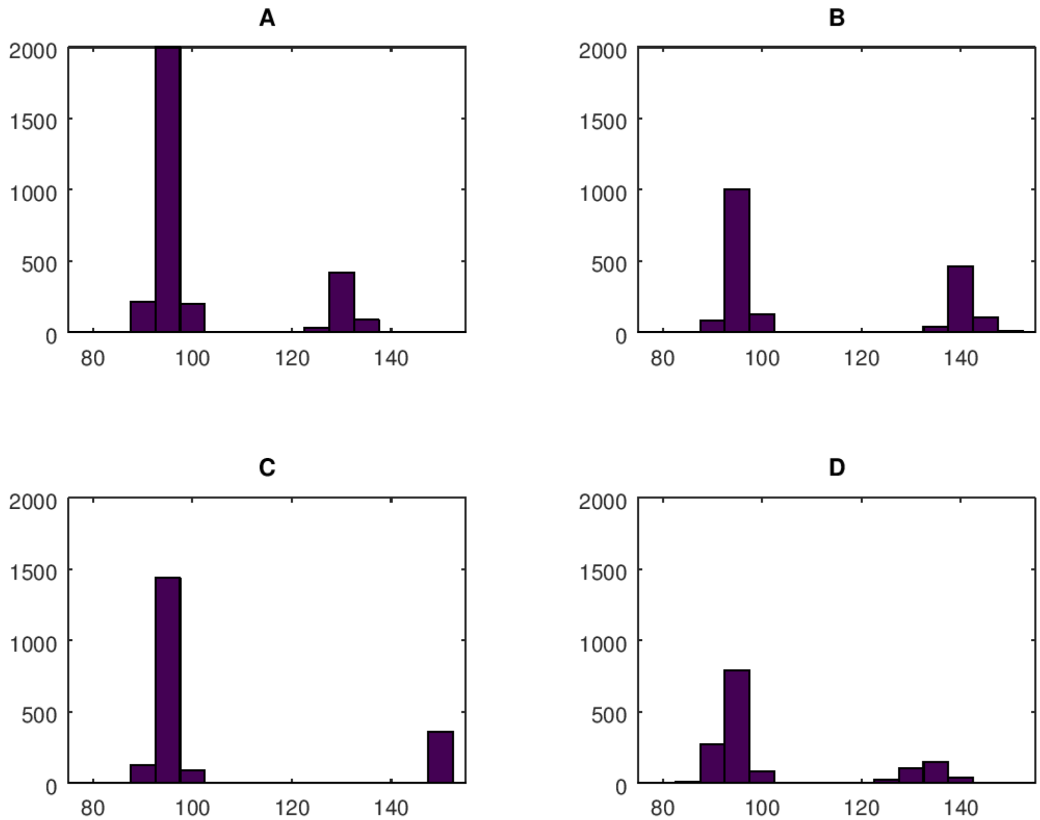

FIGURE 3 Histograms of simulated travel times *(seconds)*

FIELD DATA
To ensure that findings from simulation are relevant to field conditions, data from real-world arterial networks were collected and analyzed. Because signal timings could not be changed merely to test out different theories, analysis focused on existing signal timing strategies. Two sets of data were analyzed: vehicle trajectories and travel times from the Next Generation Simulation dataset on Lankershim Boulevard in Los Angeles, California, and travel times recorded from Bluetooth reidentification sensors along US-29 in Charlottesville, Virginia.

NGSIM Vehicle Trajectories
On June 16, 2005 from 8:30 AM to 9:00 AM, the FHWA collected high-resolution video of a 1,600 foot (500 meter) section of Lankershim Boulevard in the Universal City neighborhood of Los Angeles, California (*17*). As part of the effort, video images were post-processed to determine vehicle positions and speeds ten times per second. Additionally, the status of each signal phase was recorded in real time. Analysis of the data reveals that vehicles traveling southbound through all four signalized intersections exhibited two typical travel times, as indicated in the histogram in Figure 4.

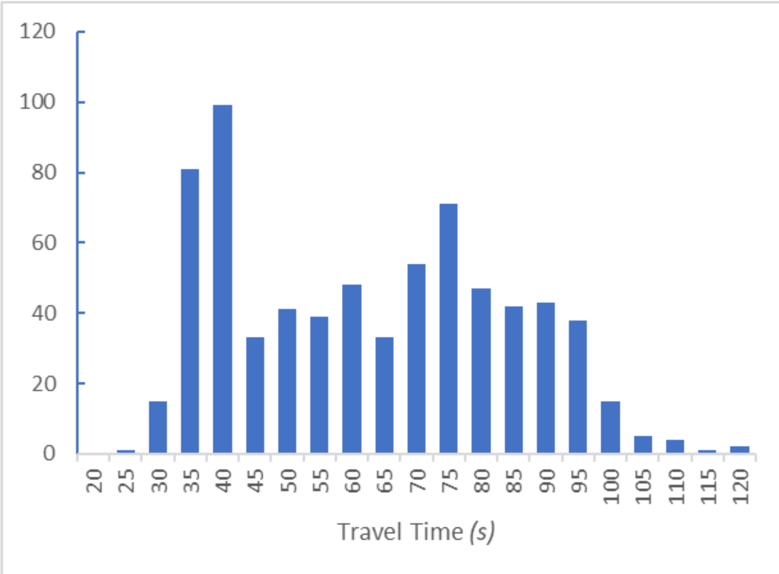

FIGURE 4 Histogram of travel times on Lankershim Blvd southbound

Using k-means clustering with $k = 2$ identified two populations of travel times with centroids of 79 seconds for the slower ($n = 354$) and 42 seconds for the faster ($n = 358$). The cutoff between the two clusters was a travel time of 75 seconds. The difference between the centroids of each cluster was 35 seconds, which matches a visual inspection of the histogram in Figure 4. Vehicles in the slower cluster were far more likely to stop along the corridor, with 93.2% of vehicles in the slow cluster reaching a speed of 0 mi/hr, while only 24.8% of vehicles in the fast cluster came to a stop.

The difference between average travel times of the two populations suggests that the average delay experienced by a stopped vehicle was 35 seconds. As this is a coordinated signal system, vehicles will generally stop towards the beginning of a red phase as the vehicle drops back from the platoon. Table 3 shows the average duration and standard deviation of red phases along the corridor. These red times seem to support the histogram in Figure 4. Many drivers in the slow population experienced delays of 35 seconds more than those in the fast population, due in large part to the 175 stops at the Lankershim Blvd. Off-Ramp intersection's 31 second red time. Simultaneously, many vehicles experienced an additional 40 to 55 seconds of delay, which corresponds to the high number of stops at Campo De Cahuenga Way / Universal Hollywood Dr. intersection and its 50 seconds of red time. A time space diagram of vehicle movements along the Lankershim Blvd corridor is shown in Figure 5.

TABLE 3 Observed Red Times During NGSIM Data Collection

Cross Street	Mean Red Time *(s)*	Standard Deviation *(s)*	Number of Stops
Valley Heart Dr. / James Stewart Ave.	43	13.2	58
MTA Dwy. / Main St.	27	8.8	32
Campo De Cahuenga Wy. / Universal Hollywood Dr.	50	8.8	194
Lankershim Boulevard Off-Ramp	31	7.5	175

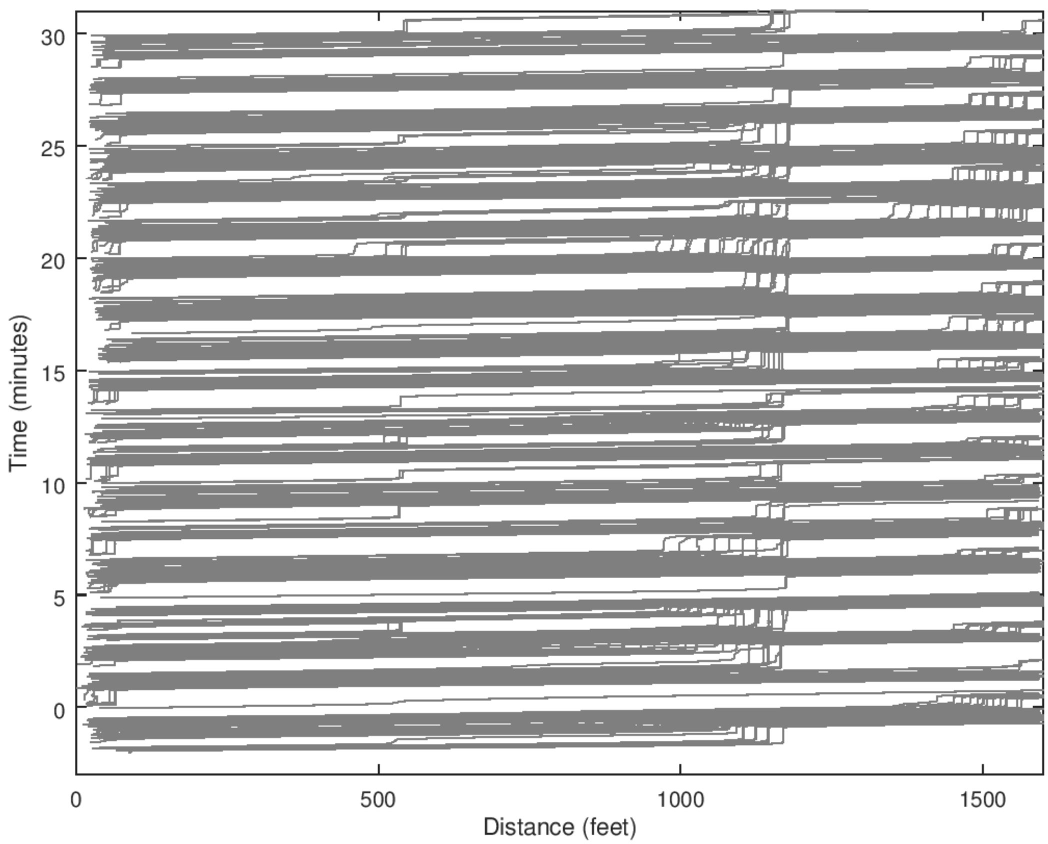

FIGURE 5 Time-space diagram of Lankershim Blvd. NGSIM data

Bluetooth Reidentification Data

Additional data collection was performed using Bluetooth reidentification technology. Bluetooth sensors can record the unique MAC addresses of certain mobile phones and other Bluetooth-equipped devices. When two units are used at different locations and record the same vehicle, the difference between the timestamps can be used to calculated the vehicle's travel time.

Bluetooth sensors were placed along a four-signal, 1 mile (1.6 km) segment of US-29, a high-volume arterial in Charlottesville, Virginia. Units were placed mid-block to avoid measuring side street vehicles. Data was collected continuously from October 19-27, 2016. Analysis focuses on the southbound morning peak period, on weekdays between 6 AM and 9 AM. There are several reasons for this. First, the roadway is near capacity at this time of day, which ensures that most traffic movements are gapping out and therefore the maximum red phase lengths from the timing plans are likely the actual red phase lengths experienced in the field. Second, the southbound morning peak period exhibited an obvious bimodal distribution from a visual inspection of the histogram.

Travel times were classified into two groups using k-means clustering. The slow group had average travel times of 147 seconds, while the faster group averaged 90 seconds for an average speed of 40 mi/hr (64 km/hr), very close to the 45 mi/hr (72 km/hr) speed limit. The histogram of travel times is shown in Figure 6.

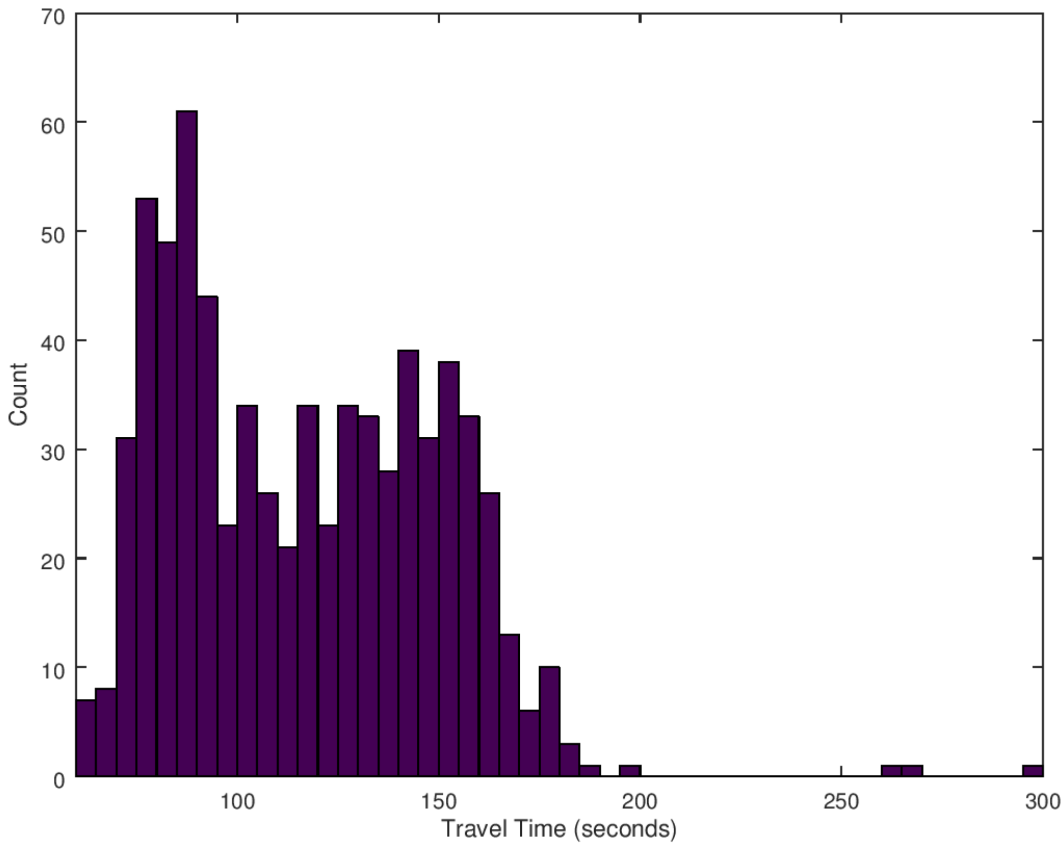

FIGURE 6 Histogram of travel times for US-29 in Charlottesville northbound during AM peak period

The difference in average travel time between the two groups was 57 seconds. When compared to the maximum red times shown in Table 4, the coordinated signal system behaved as expected. Travel times differences were slightly longer than the average red time, and many slow population travel time differences exceeded 70 seconds, which is slightly longer than the maximum red time at Hydraulic Road of 65 seconds.

TABLE 4 Maximum Red Times for US-29 SB, AM Peak Period

Cross Street	Maximum Red Time *(s)*
Hydraulic Road	65
Seminole Court	48
Greenbrier Drive	54
Lenox Avenue	25
Average	48

CONCLUSIONS

Travel time is a widely-used performance metric in transportation. Travel times on signalized intersections and corridors often exhibit a bimodal distribution. While researchers have suggested that traffic signal control impacts the form of these distributions, no research has investigated the specific role of timing plans on bimodal distributions. This study proposed a theoretical framework for signal control and the differences in average travel time between fast and slow populations. For isolated intersections, travel time differences are similar to the average red phase time, but slightly higher given deceleration and acceleration delay. For coordinated signals, vehicles traveling in platoons will encounter a red phase at the beginning of the phase. Vehicles in coordinated systems must therefore wait out almost all of the red phase. The difference in travel time between this slow group and a faster group should be equal to the red time, plus some additional time for acceleration and deceleration. When a corridor has multiple signals with different red times, then the slow population will have a wide variance reflecting the experiences of vehicles stopping at different signals.

These assumptions were tested in a microscopic simulation, on empirical vehicle trajectories from the NGSIM dataset, and from sampled vehicle travel times collected with Bluetooth reidentification sensors. Results appeared to confirm the assumptions, although additional vehicle trajectory data would be needed to completely confirm that vehicles are behaving as expected.

This paper contributes to a general understanding of the effect of signal control on bimodal travel time distributions. Differences between average travel times between fast and slow groups can be predicted with some certainty, and any deviation from a prediction can be investigated. Unimodal distributions over short segments, for example, suggests overcapacity, as no vehicles are able to get through without experiencing delay. Travel time differences may also indicate which signals experience disproportionately more stops. If the travel time histogram shows travel time differences of twenty seconds along a corridor, then the signal with a twenty second red phase is likely the bottleneck.

Future research efforts should deploy Bluetooth, WiFi, video, or other sensors at multiple points along the corridor to determine how individual vehicles travel within a corridor, rather than relying exclusively on end-to-end travel time. Signal control affects how many vehicles stop at a signal, and for how long. Without intracorridor monitoring, it is difficult to develop and validate models.

ACKNOWLEDGEMENTS
For the US-29 travel time data, Michael Pajewski assisted in data collection and Mengxuan Zhang performed preliminary analysis.

REFERENCES
1. Bhaskar, A., and E. Chung. Fundamental Understanding on the Use of Bluetooth Scanner as a Complementary Transport Data. *Transportation Research Part C: Emerging Technologies*, Vol. 37, 2013, pp. 42–72. https://doi.org/10.1016/j.trc.2013.09.013.
2. Goodall, N. J. Fundamental Characteristics of Wi-Fi and Wireless Local Area Network Re-Identification for Transportation. *IET Intelligent Transport Systems*, Vol. 11, No. 1, 2017, pp. 37–43. https://doi.org/10.1049/iet-its.2016.0087.
3. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers*, Vol. 2, No. 1, 1952, pp. 325–379.
4. Richardson, A. J., and M. Taylor. Travel Time Variability on Commuter Journeys. *High Speed Ground Transportation Journal*, Vol. 12, No. 1, 1978.
5. Polus, A. A Study of Travel Time and Reliability on Arterial Routes. *Transportation*, Vol. 8, No. 2, 1979, pp. 141–151. https://doi.org/10.1007/BF00167196.
6. Al-Deek, H., and E. B. Emam. New Methodology for Estimating Reliability in Transportation Networks with Degraded Link Capacities. *Journal of Intelligent Transportation Systems*, Vol. 10, No. 3, 2006, pp. 117–129. https://doi.org/10.1080/15472450600793586.
7. Susilawati, S., M. A. P. Taylor, and S. V. C. Somenahalli. Distributions of Travel Time Variability on Urban Roads. *Journal of Advanced Transportation*, Vol. 47, No. 8, 2013, pp. 720–736. https://doi.org/10.1002/atr.192.
8. Davis, G. A., and H. Xiong. *Access to Destinations: Travel Time Estimation on Arterials*. Publication MnDOT 2007-35 1381. Minnesota Department of Transportation, Minneapolis, MN, 2007.
9. Susilawati, M. A. Taylor, and S. Somenahalli. Travel Time Reliability and the Bimodal Travel Time Distribution for an Arterial Road. Presented at the ARRB Conference, 24th, 2010, 2010.
10. Ji, Y., S. Jiang, Y. Du, and H. M. Zhang. Estimation of Bimodal Urban Link Travel Time Distribution and Its Applications in Traffic Analysis. *Mathematical Problems in Engineering*, Vol. 2015, 2015, pp. 1–11. https://doi.org/10.1155/2015/615468.
11. Zheng, F., and H. Van Zuylen. Uncertainty and Predictability of Urban Link Travel Time. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2192, 2010, pp. 136–146. https://doi.org/10.3141/2192-13.
12. Yang, F., M.-P. Yun, and X.-G. Yang. Travel Time Distribution Under Interrupted Flow and Application to Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2466, 2014, pp. 114–124. https://doi.org/10.3141/2466-13.
13. Ji, Y., and H. M. Zhang. Travel Time Distributions on Urban Streets: Estimation with Hierarchical Bayesian Mixture Model and Application to Traffic Analysis with High-Resolution Bus Probe Data. Presented at the Transportation Research Board 92nd Annual Meeting, 2013.

14. Young, S. E., E. Sharifi, C. M. Day, and D. M. Bullock. Visualizations of Travel Time Performance Based on Vehicle Re-Identification Data. Presented at the 96th Annual Meeting of the Transportation Research Board, Washington, DC, 2017.
15. Jr, D., R. W, E. Curtis, and P. Olson. The National Traffic Signal Report Card. *ITE Journal*, Vol. 82, No. 6, 2012.
16. Eaton, J. W., D. Batemen, S. Hauberg, and R. Wehbring. GNU Octave Version 4.2.0 Manual: A High-Level Interactive Language for Numerical Computations. https://www.gnu.org/software/octave/doc/interpreter/Citing-Octave-in-Publications.html. Accessed Aug. 2, 2017.
17. Federal Highway Administration. Next Generation SIMulation Fact Sheet, FHWA-HRT-06-135. https://www.fhwa.dot.gov/publications/research/operations/its/06135/index.cfm. Accessed Aug. 1, 2017.