Identifying the Bias

Evaluating Effectiveness of Automatic Data Collection Methods in Estimating Details of Bus Dwell Time

Emily Grisé and Ahmed El-Geneidy

Data from automated vehicle location (AVL) systems, automatic passenger counter (APC) systems, and fare box payments have been heavily used to generate dwell time models with the goal of recommending improvements in efficiency and reliability of bus transit systems. However, automatic data collection methods may result in a loss of detail with regard to the dynamics of passenger activity, which may bias the estimates associated with dwell or passenger activity time. The purpose of this study is to understand better any biases that might exist from using data from AVL–APC systems or fare box payments when estimating dwell time. Manually collected data from Montreal, Quebec, Canada, are used to estimate detailed dwell time models. This study compared those estimates to models generated by using data similar to what was reported by AVL–APC systems and fare boxes. The results reveal an overestimation in the passenger activity component of dwell time, which is mainly attributed to excess dwell time that AVL–APC data and fare box payments generally do not capture. While AVL–APC and fare box technologies provide transit agencies with rich data for analysis, adjustments to such data collection methods are warranted to reduce the overestimation of dwell time and to provide a more accurate picture of what is happening on the ground to generate better interventions that can reduce dwell times.

Data from automated vehicle location (AVL) systems, automatic passenger counter (APC) systems, and fare box payments have been used heavily in past years to generate dwell time models, with the goal of recommending or evaluating improvements in efficiency and reliability of bus transit systems, or both. These strategies can include—but are not limited to—all-door boarding strategies, smart card policy, off-board fare collection, low-floor vehicles, and use of articulated vehicles. These particular strategies are used to reduce the dwell time needed for passenger activity because dwell time can consume up to 25% of total running time (*1*). Therefore, the determinants of dwell time and time associated with its major components—including the time required to open and close doors, passenger activity time, and additional time until door closings—have been extensively studied by using data collected either manually or automatically (*2–4*).

Before AVL–APC systems, the collection of data for bus dwell time analyses was done manually through labor-intensive methods (*1, 5, 6*). Such methods, however, allow for direct observations of

School of Urban Planning, Faculty of Engineering, McGill University, Suite 400, 815 Sherbrooke Street West, Montreal, Quebec H3A 0C2, Canada. Corresponding author: A. El-Geneidy, ahmed.elgeneidy@mcgill.ca.

passenger activity per door, fare payment method, and unproductive door opening time (*4*). By the early 2000s, researchers began to study the application of AVL–APC data (*4, 7, 8*), which provided a large number of observations to aid in the development of statistical models with greater explanatory power. Automatic data collection methods result in a loss of some detail regarding the dynamics of passenger activity; such methods, however, provide a larger data set for analysis at a lower cost. This trade-off can be minimized through introducing improvements in automatic data collection methods and through identification of biases that are imposed from use of this method when compared to manual counts with detailed observations.

The purpose of this study is to estimate how accurately AVL–APC and fare box data are capturing the time associated with passenger activity. To achieve this goal, the authors used manually collected stop-level observation data from two bus routes operated by the Société de transport de Montreal (STM), the public transit service provider on the island of Montreal, Quebec, Canada. The study compared estimates of detailed dwell time by relying on manual count observations to estimates generated from models using data similar to what is reported by AVL–APC systems and fare boxes.

This paper starts with a review of previous research, followed by a methodology and data collection section, which includes a description of the data used in the analysis. A results section includes the different statistical models, followed by a discussion section. Finally, the paper ends with a conclusion and recommendations section.

LITERATURE REVIEW

Intelligent transport systems (ITS) provide transport agencies with essential information and communication technologies to make informed decisions (*9*). In the public transit sector, AVL and APC systems are major components of ITS technologies that many agencies around the world are using or are in the process of implementing because of their wide range of benefits to agencies and customers. For example the application of AVL technology has been dominated by real-time applications such as computer-aided dispatching, next-stop announcements, and next-arrival displays (*10*). Also, archived data collected through AVL and APC systems provide transit agencies with a rich and extensive database that can be analyzed in transit research for planning and operational improvements (*4, 11*). Operational improvements—namely, reductions in travel time and improvements in reliability—increase the operational efficiency for a transit provider (*12*), while such improvements may also result in the growth of patronage (*13, 14*) and increase riders' satisfaction (*15*).

Dwell time is defined as the time required for a transit vehicle to stop for the purpose of allowing passengers to board and alight

(*16*). The main component of dwell time is passenger activity at each stop (*17*); however, the *Highway Capacity Manual* in 1997 indicated that the average time for passenger activity depends on many factors, such as the height of the bus or number of steps, the fare collection system of the bus, and time associated with lift operation (*18*). In the past, manually collected data were used to estimate dwell times (*1, 19–21*). These studies focused on estimating the time associated with passenger activity. One of the first studies that analyzed transit travel time performance and dwell time variation was published in 1983 by Levinson (*1*). In this landmark study, estimates of bus dwell times revealed that each passenger boarding added 2.75 s to the constant dwell time of 5 s (includes door opening and closing times). Similarly, Guenthner and Sinha estimated that each passenger boarding or alighting added 3 to 5 s to the dwell time (*19*). Accounting for fare payment type, Zografos and Levinson estimated that a passenger boarding time took only 2 s on an uncrowded bus with an off-board fare payment system in Connecticut (*22*). Early examples of dwell time models, such as the examples presented above, were based on smaller sample sizes because of the labor-intensive data collection methods required, and resulted in lower explanatory power (*4*).

By the late 1990s, technological advancements allowed for on-board implementation and data collection from AVL–APC systems, which allowed researchers the ability to study the application of data collected by AVL–APC systems for analyses on service reliability and enhanced route planning. Strathman et al. presented a baseline analysis of data collected by Tri-Met, the transit provider in Portland, Oregon, to assess service reliability on selected routes (*7*). Bertini and El-Geneidy focused on the dwell time of a single route in Portland, Oregon, and found significant model improvements from early studies that measured the number of passengers boarding and alighting together (*8*). The authors estimated that each passenger boarding adds approximately 3.6 s, while each passenger alighting adds only 0.85 s to the dwell time, and approximately 5.8 s is attributable to time required for door opening or closing. In a similar study of AVL–APC data collected by Tri-Met, Dueker et al. found that each passenger boarding adds 3.5 s and each passenger alighting adds 1.7 s to the constant dwell time of 5.1 s (*4*).

More recently, Diab and El-Geneidy evaluated the impact of bus stop location on dwell time and found longer dwell times associated with bus stops located on the nearside of an intersection (*23*). Furthermore, the authors estimated that passenger boarding added 3.3 s and each passenger alighting added 1.9 s to the total dwell time. Evaluating the dwell time-savings of operating articulated buses, El-Geneidy and Vijayakumar found an average passenger boarding time of 4.1 s and alighting time of 2.7 s; however, time savings are found at the second and third doors where passengers are not required to scan their fare cards (*3*). From these select studies using AVL–APC data, it can be seen that on average, the first passenger boarding takes between 3.3 and 4.1 s. However, as shown in Levinson's 1983 estimate of dwell times using manually collected data, each passenger boarding added approximately 2.75 s (*1*). Such higher estimates of passenger boarding time may be attributed to the method by which AVL–APC data record dwell time. Dwell time is recorded by an AVL–APC system as the total time that the door is open; therefore, additional time when the door remains open after passenger activity may result in an overestimate of the impacts of passenger activity on dwell time. Thus, while AVL–APC technology provides transit agencies with rich data for analysis and operational improvements, manual data collection methods remain vital for verification of AVL–APC data to understand better the reasons for a discrepancy in the estimates.

Last, literature has explored the passenger boarding time associated with various fare payment methods. Improvements to both the method and location of fare payment can result in significant time-savings in dwell times (*24*). Research has found different boarding times associated with different methods of fare payment (*25, 26*). In some cases, because of the absence of AVL–APC data, automatically collected fare box data are used to estimate passenger activity time (*6*). While some studies relied on manual counts to understand impacts of fare collection methods on dwell time due to limitations in automated data collection systems (*27*). Despite significant advances in knowledge of determinants of dwell time through the use of AVL–APC and fare box data, more detailed analysis is required to assess the accuracy of automatic data collection methods used to estimate dwell time and to explain the discrepancy the authors noticed between earlier models generated from manual counts and current ones generated from automatically collected data.

METHODOLOGY AND DATA COLLECTION

The objective of this analysis is to assess how well data from AVL–APC systems and fare boxes are able to measure passenger activity through the generation of dwell time models. Dwell time is defined as the time between doors opening and doors closing, including passenger activity time. Figure 1 represents the elements of dwell time that are captured through an automated or a manual system. An AVL system measures the time from door opening to door closing, while the APC system, which relies on two infrared beams in most cases, counts the number of passengers crossing these beams to identify the number of boardings and alightings during this period. However as shown in Figure 1, additional time with no passenger activity before the door closes may be captured in AVL–APC data, which might be generating an overestimate of the time required for passenger activity. This excess dwell time can be present for various reasons: for example, bus holdings at time points or a red light in the case of a nearside bus stop. Such excess dwell time can be needed in some cases, while in other circumstances this time can be removed from the schedule to save passenger and operating time. This study was focused on estimating the extent to which this excess dwell time—time with no passenger activity—impacts the estimate of passenger activity in dwell time models. The authors therefore employed operations data collected through field measurement to model dwell time, from stop-level observations collected from two bus routes. This was done by observing the total dwell time, as well as the difference in time between the end of passenger activity and time until the door closed, which this study referred to as excess dwell time.

On the basis of previous research (*3, 4, 23, 28*), a typical dwell time model is defined as follows:

dwell time (s) = f (boardings, alightings, total passenger activity^2, passenger load, friction, direction, time of day, lift usage, stop location, weather condition, fare payment method).

Variables directly related to passenger activity include: total boardings and alightings, passenger load, a passenger friction factor, and passenger activity squared. Total boardings are the number of people who boarded the bus, while passenger alightings are the number of people who exited the bus. The squared term of passenger activity is included to capture the potential nonlinear relationship of passenger activity, suggesting that each additional passenger requires
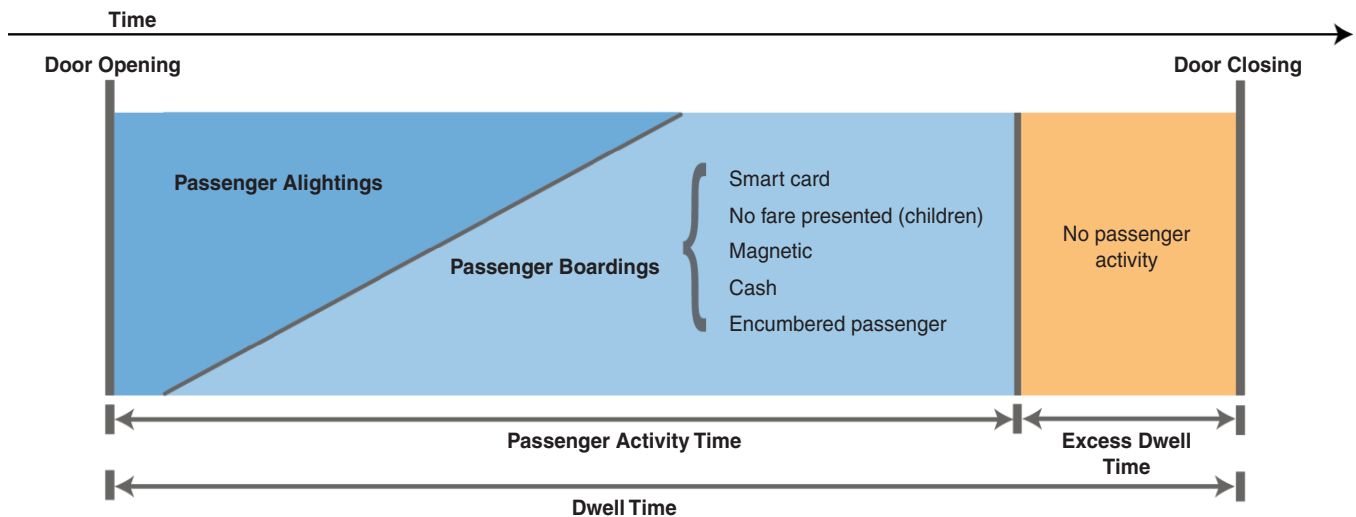
FIGURE 1   Elements of dwell time.

less time to board than the passenger ahead (*3, 4, 23*). The passenger friction factor is used to capture the effect of an overcrowded bus with standees on dwell time and in this study was calculated as the sum of passenger activity and number of standees. Passenger load represents the number of passengers onboard the bus at the departure of the stop, where higher passenger loads are expected to increase dwell time (*4*).

The direction of the route and time of day are included in the model to estimate any associated differences in dwell time. Weather condition, a variable outside the control of the transit agency, may affect dwell times as seen in previous studies (*3*) but was not included in this study's models because of the lack of adverse or varying weather conditions during the data collection period. A dummy variable for lift operation, used to board passengers with a disability, accounts for the additional dwell time at such stops (*4*). The bus lift was not used during the data collection period; however, the authors observed passengers boarding with an encumberment (such as a stroller, or large, heavy bags), as well as passengers with an observable disability or mobility limitation. Characteristics of each stop along the route were collected and tested in the authors' models, such as stops occurring on the nearside or farside of an intersection, presence of a reserved bus lane, and bus shelters; few, however, were found significant and included in the final model. Finally, fare payment method was expected to affect dwell time, as previous research observed differences in passenger boarding times by different payment types (*27*).

The data for this study were collected from two bus routes in Montreal, Quebec, Canada, which are both operated by STM, the primary transit operator in Montreal. The two routes used for this study are 121 Sauvé and 69 Gouin (as shown in Figure 2). Both routes operate east–west and have similar operational characteristics, such as similar average stop spacing. Both routes operate in neighborhoods with comparable built environments, share a connection with a metro station, and exclusively operate articulated buses. The daily weekday ridership of Route 121 is approximately 35,000 individuals, and approximately 27,000 individuals use Route 69 on weekdays. The main operational difference between these routes is that STM introduced a pilot all-door boarding policy along Route 121 in March 2016, which will be accounted for in the authors' models through a Route 121 dummy variable.

Manual observations of data were collected by three individuals who recorded important information related to passenger activity at each door, as well as the dwell time at each stop along both routes. For consistency, the same three people collected all the data. With regard to the dwell time recorded at each stop, the authors collected data on the time between door open and door close (total dwell time), the time of passenger activity (time after doors open to end of passenger activity), and excess dwell time. The excess dwell time recorded represents the time spent at stops that was not the result of passenger activity (such as time points and driver changes). Additional data collected included number of passengers boarding and alighting at each door at each stop, arrival time at each stop, as well as observations related to passenger encumberment (such as a stroller or large, heavy bags), or a passenger with an observable disability or mobility restriction which may have extended the average time of passenger activity. Finally, passenger load at every stop was calculated by counting the number of passengers on board at the beginning of the route and then adding and subtracting the boardings and alightings, respectively, throughout the route.

Characteristics of each stop along the route were also collected, such as stops occurring on the nearside or farside of an intersection, presence of a reserved bus lane, and bus shelters. Furthermore, the authors observed which payment type was used by each passenger boarding that involved a fare transaction with the driver (because passengers were allowed to board at the middle and rear doors on Route 121). There are three options of payment method, including cash, a magnetic fare card, and a smart card. Furthermore, the authors recorded the number of passengers under the age of 6 years who boarded at the front door of the bus accompanied by an adult and were categorized as no-fare payment because such users ride transit free of charge. Payments by smart card on the STM bus network only require commuters to tap their passes upon boarding. A total of 1,036 stop observations were collected from 17 unique trips aboard Route 121 and, additionally, data from four unique trips were collected from Route 69. Data collection occurred between 6:30 a.m. and 6:30 p.m. on a Tuesday, a Wednesday, and a Thursday during the month of May 2016. The dates and times of data collection were chosen to collect an equal distribution of data between peak and off-peak hours. Throughout the data collection, weather conditions were normal (dry) and no events impeded ordinary operations of the bus route.
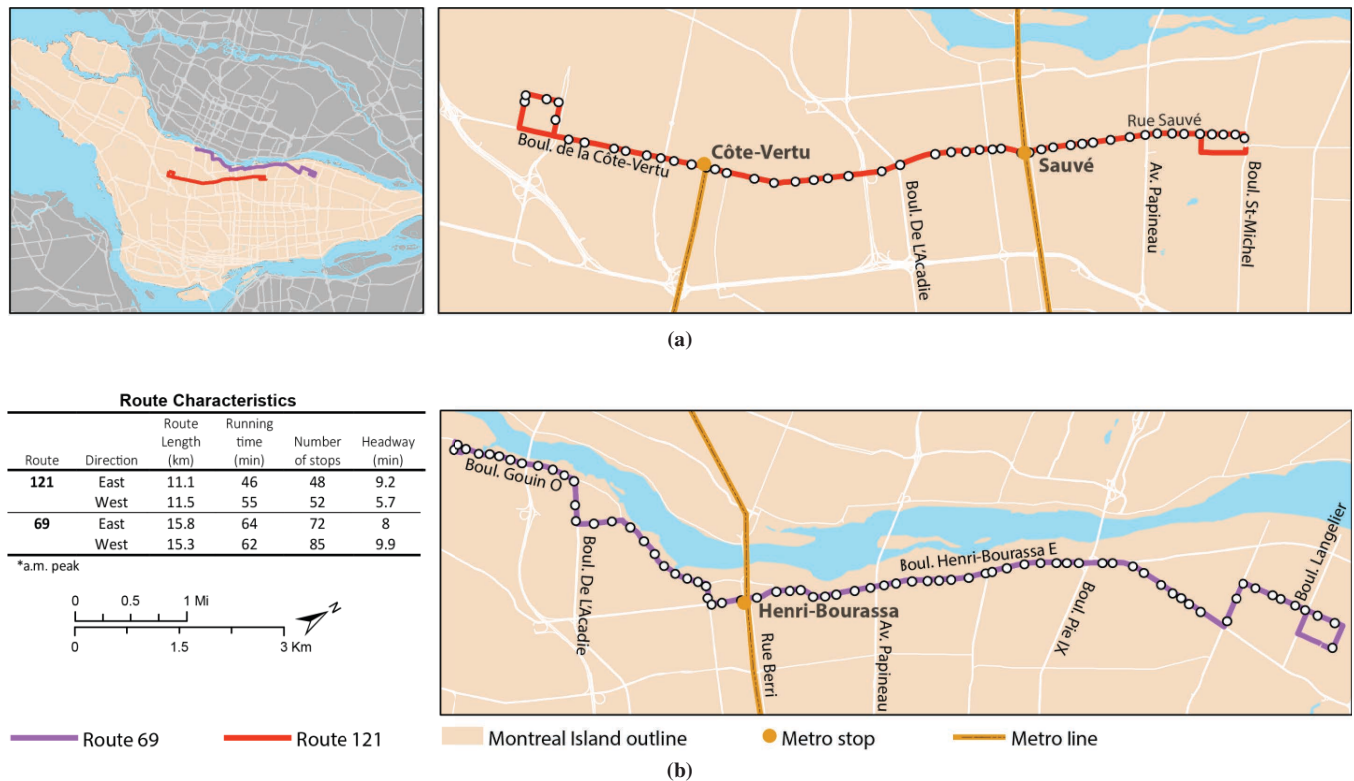
(a)

Route	Direction	Route Length (km)	Running time (min)	Number of stops	Headway (min)
121	East	11.1	46	48	9.2
	West	11.5	55	52	5.7
69	East	15.8	64	72	8
	West	15.3	62	85	9.9

*a.m. peak

(b)

FIGURE 2   Context map of routes studied and relevant route characteristics: (*a*) Route 121 Sauvé and (*b*) Route 69 Gouin. (Data source: STM and DMTI; projection source: North American Datum 1983 Modified Transverse Mercator 8.)

RESULTS

Model 1 is a traditional dwell time model, which employs a linear regression method—where the times required for each passenger boarding and alighting are estimated—with a series of independent variables meant to capture variations in dwell time between stops and the two routes. This model represents typical information collected by AVL–APC systems. In Model 2, the authors added the exact amount of excess dwell, which was noticed through the manual observation to understand the bias imposed on each coefficient when such a variable is not included and only AVL–APC-like data are used. Model 3 includes a dummy variable for stops where the authors observed an encumbered passenger boarding. The model predicted the additional dwell time that can be expected to serve encumbered passengers, data currently not captured by AVL–APC data. Model 4 builds off the first three models by including detailed information on fare payment methods used by each front-door boarding. Models 5 and 6 expand on Model 4 similarly by adding the excess dwell variable and passenger encumberment variable to understand better the impacts of not including such information in a model derived from information from AVL–APC systems and fare box payments. A further description of variables included in such models is presented in Table 1. Layover stops (the first and final stops of each trip) are not included.

Dwell Time Model

Table 2 shows the estimates and 95% confidence intervals for Models 1 through 3. After first examining the traditional dwell time model (Model 1), the independent variables included in the model explain approximately 52% of the variation in dwell time. The constant variable (4.8 s) in this model represents a fixed amount of time that is associated with door openings and door closings. On average, each passenger boarding adds 4.3 s to the base dwell time while keeping all other variables at their mean values. Each passenger alighting adds 2.1 s to the total dwell time, keeping everything else constant. Similar to previous research, the authors noticed a diminishing impact of the squared term (passenger activity squared). This means that time used for every additional passenger movement is lower than the previous one. Furthermore, the positive coefficient of the passenger friction factor shows the additional dwell time added on overcrowded buses with standees. These coefficients are consistent with findings from previous research (*3, 4, 23*).

In the expanded dwell time model (Model 2), which includes the amount of excess dwell time, the model variables explain approximately 95% of the variation in dwell time. The excess dwell time variable controls for the amount of additional dwell time that occurred after the end of passenger activity. The constant variable reports 3.3 s compared to 4.8 s in the first model, while each passenger boarding on average adds 1.8 s to the total dwell time keeping all other variables at their mean values. While each passenger alighting adds 0.8 s to the constant dwell time of 3.3 s. The estimated time of the first passenger boarding in this model is approximately 2.4 times less than the traditional model. Similar to Model 1 and previous research, additional passengers use less time, as noticed from the passenger activity square coefficient. On average, excess dwell time adds an additional 1 s to the total dwell time, indicating an overestimation of time required for passenger activity.

Model 3 shows the variation in dwell time associated with an encumbered passenger boarding. The model variables explain

TABLE 1   Descriptive Statistics

Variable	Description	Mean	SD	Count
Total boardings	Total number of passengers who boarded at all doors at a single bus stop during a trip	1.95	4.64	NA
Smart card	Number of passengers who paid with a smart card	1.54	2.74	1,456
Magnetic card	Number of passengers who paid with a magnetic pass	0.03	0.27	28
Cash	Number of passengers who paid with cash	0.05	0.24	48
No fare presented	Number of passengers who boarded bus without presenting fare	0.05	0.28	44
Total rear-door boardings	Total number of passengers who boarded bus at either the middle doors or the rear doors at a single bus stop during a trip	0.46	2.71	NA
Total alightings	Total number of passengers who alighted at all doors at a single bus stop during a trip	2.23	5.41	NA
Total passenger activity^2	Square of total number of passenger boardings and alightings at all doors at a stop during a trip	105.02	676.02	NA
Excess dwell	Additional dwell time after end of passenger activity (s)	6.17	25.76	NA
Route 121	Dummy variable equal to 1 if stop occurred on Route 121	0.77	0.42	NA
Friction	Total number of standees plus sum of total boarding and alightings at a stop	4.67	9.74	NA
a.m.	Dummy variable equal to 1 if trip took place between 6:30 and 9:30 a.m.	0.34	0.48	NA
p.m.	Dummy variable equal to 1 if trip took place between 3:30 and 6:30 p.m.	0.24	0.43	NA
Midday	Dummy variable equal to 1 if trip took place between 9:30 a.m. and 3:30 p.m.	0.42	0.49	NA
Passenger load	Total number of passengers on a bus at departure of a stop	27.21	15.96	NA
Eastbound trip	Dummy variable equal to 1 if stop occurred on an eastbound trip	0.49	0.50	NA
Metro station	Dummy variable equal to 1 if stop occurred at a metro station	0.05	0.21	NA
Encumbered passenger	Dummy variable equal to 1 if a passenger with an encumberment, disability, or mobility limitation boarded or alighted from bus	0.03	0.18	NA
Signalized intersection	Dummy variable equal to 1 if stop occurred at a signalized intersection	0.66	0.47	NA

NOTE: NA = not available.

TABLE 2   Dwell Time Models

	Traditional Dwell Time (Model 1)[a]			Expanded Model (Model 2)[b]			Expanded Model (Model 3)[c]		
		95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
Variable	Coefficient	Lower Bound	Upper Bound	Coefficient	Lower Bound	Upper Bound	Coefficient	Lower Bound	Upper Bound
Constant	4.82**	0.84	0.98	3.33***	2.09	4.57	3.26***	2.05	4.46
Total boardings	4.33***	3.73	8.81	1.84***	1.65	2.03	1.74***	1.55	1.93
Total alightings	2.14***	1.47	4.93	0.78***	0.57	0.99	0.76***	0.56	0.97
Total passenger activity^2	−0.011***	−0.02	2.82	−0.010***	−0.01	−0.01	−0.0096***	−0.01	−0.01
Excess dwell	NA	NA	NA	0.96***	0.94	0.98	0.96***	0.94	0.98
Friction	−0.66**	−1.19	−0.14	0.32***	0.16	0.49	0.32***	0.16	0.48
Eastbound trip	−1.19	−3.96	1.58	−0.76**	−1.62	0.10	−0.83**	−1.67	0.00
a.m.[d]	−4.55***	−7.81	−1.29	−0.70	−1.72	0.31	−0.76*	−1.74	0.23
p.m.[d]	−0.49	−4.19	3.20	0.17	−0.98	1.32	0.14	−0.97	1.25
Metro station	26.54***	18.46	34.62	−3.82***	−6.41	−1.23	−3.23**	−5.74	−0.71
Encumbered passenger	NA	NA	NA	NA	NA	NA	9.19***	6.84	11.54
Signalized intersection	5.52***	2.50	8.54	1.67***	0.72	2.61	1.42***	0.51	2.34
Route 121	−1.43	−4.96	2.10	0.45	−0.65	1.54	0.65	−0.42	1.71

[a]R^2 = .52.
[b]R^2 = .95.
[c]R^2 = .96.
[d]Reference category is midday.
*$p < .1$; **$p < .05$; ***$p < .01$.

approximately 96% of the variation in dwell time. Models 2 and 3 show that the addition of excess dwell time and encumbered passenger boarding variables to Model 1 result in a significant increase in the model fitting, from 0.52 to 0.95 and 0.96, respectively. This indicates the importance of these two variables for estimating the variation in dwell time. The model indicates that dwell time is expected to increase by 9.2 s for an encumbered passenger boarding or alighting the bus, while keeping all other variables constant at their mean values. The constant variable and estimate of passenger alighting are consistent with those of Model 2, while each passenger boarding on average adds 1.7 s compared to 1.8 s in Model 2, when all variables are controlled for. Remaining variables related to passenger activity behave consistently with those of Models 1 and 2.

For the remaining control variables in Models 1 through 3, dwell times at signalized intersections are slower compared to midblock stops or nonsignalized stops, which is consistent with previous studies (23). Dwells taking place at a bus stop with a connection to the metro are generally faster compared to a non-metro stop. Also, trips during the a.m. peak period are estimated to have faster dwell times than those of midday trips, this can be attributed to a greater proportion of regular riders who may board by using passes and ask fewer questions (4). Lastly, dwell times of eastbound bus stops were on average 1 s faster compared to westbound trips. The variable controlling for dwell time variation between Routes 121 and 69 did not show a statistically significant effect on dwell time in the sample, although Route 121 allowed all door boardings.

Dwell Time Model with Fare Payment Type

Table 3 shows the estimates and 95% confidence intervals of Models 4 through 6 which estimate the variation in dwell time by each fare payment type, while otherwise keeping the models consistent with the previous dwell time model. After first examining the dwell time model with fare payment type (Model 4), the independent variables included in the model explain approximately 54%t of the variation in dwell time. The constant variable in this model reports 3.3 s that is associated with door opening and door closing times for each dwell, keeping all other variables at their mean values. Consistent with previous studies, certain fare types have larger impacts on dwell time (27). Each passenger boarding with a smart card added 4.7 s, while each passenger boarding with a magnetic card added 21.8 s, on average, to the dwell time while keeping everything else constant. A magnetic card must be validated upon boarding, and the long boarding time associated with this payment type is largely associated with observed difficulties or confusion of passengers with the correct method to insert and validate the magnetic strip card. Boarding time associated with passengers categorized as no-fare payment (young children) increased dwell time by 4.2 s, on average. Finally, each passenger who paid with cash to the driver added 8.7 s to the dwell time. Because Route 121 allowed all-door boarding, passengers who boarded at the second and third doors increased the dwell time by 1.5 s. Consistent with Models 1 through 3, the time required for each additional boarding was lower than the time of the

TABLE 3   Dwell Time Models with Fare Payment Types

Variable	Traditional Fare Payment (Model 4)[a]	95% Confidence Interval		Expanded Fare Payment (Model 5)[b]	95% Confidence Interval		Expanded Fare Payment (Model 6)[b]	95% Confidence Interval	
	Coefficient	Lower Bound	Upper Bound	Coefficient	Lower Bound	Upper Bound	Coefficient	Lower Bound	Upper Bound
Constant	3.30*	−0.62	7.21	2.53***	1.41	3.66	2.50***	1.40	3.60
Fare Payment Type									
Smart card	4.71***	3.96	5.46	2.50***	2.28	2.71	2.42***	2.20	2.64
Magnetic card	21.77***	16.23	27.31	0.42	−1.23	2.06	0.19	−1.42	1.79
No fare presented	4.23*	−0.83	9.28	4.17***	2.72	5.62	2.50***	1.00	4.01
Cash	8.66***	2.56	14.76	7.07***	5.32	8.82	6.93***	5.22	8.65
Total alightings	1.73***	1.05	2.40	0.64***	0.45	0.84	0.64***	0.45	0.83
Total rear-door boardings	1.48**	0.30	2.67	0.03	−0.31	0.37	0.054	−0.28	0.39
Total passenger activity^2	−0.0047**	−0.01	0.00	−0.0047***	−0.01	0.00	−0.0047***	−0.01	0.00
Friction	−0.51**	−1.03	0.01	0.27***	0.12	0.42	0.28***	0.13	0.42
Eastbound trip	−0.87	−3.58	1.83	−0.79**	−1.57	−0.02	−0.84**	−1.60	−0.08
a.m.[c]	−3.49**	−6.69	−0.30	−0.66	−1.57	0.26	−0.70	−1.60	0.20
p.m.[c]	0.21	−3.41	3.83	0.51	−0.53	1.55	0.53	−0.49	1.54
Metro station	27.55***	19.53	35.57	−1.12	−3.49	1.25	−0.74	−3.06	1.58
Encumbered passenger	NA	NA	NA	NA	NA	NA	7.58***	5.32	9.85
Signalized intersection	5.48***	2.53	8.43	1.46***	0.61	2.31	1.29***	0.46	2.12
Route 121	−1.04	−4.50	2.41	0.60	−0.39	1.60	0.70	−0.27	1.67
Excess dwell	NA	NA	NA	0.97***	0.95	0.99	0.97***	0.95	0.99

[a]R^2 = .54.
[b]R^2 = .96.
[c]Reference category is midday.
*$p < .1$; **$p < .05$; ***$p < .01$.

previous boarding. Furthermore, additional dwell time is expected on heavily loaded buses.

In the expanded fare-payment model (Model 5), which includes the amount of excess dwell, the model variables explain approximately 96% of variation in dwell time. The constant variable reports 2.5 s, which is lower than the estimated constant of 3.3 s in Model 4. Rear-door boardings in this model revealed no statistically significant effect on dwell time as such boarding happens simultaneously with front-door boardings, which requires interaction with the fare box. Similar to the previous expanded model (Model 2), significant reductions in the estimates of each fare method were observed when the excess dwell variable was included. Each passenger boarding with a smart card added 2.5 s, compared to 4.7 s in Model 4. The time required for passengers who boarded by using a magnetic card was not statistically significant in the expanded model. This may be attributed to the small number of passengers boarding with a magnetic card (1.8% of observed boardings), and likely the inconsistent time associated with this method. The authors were therefore unable to attain a statistically significant estimate of the average boarding time associated with magnetic card payments. Young children boarding with an adult added, on average, 4.2 s to the dwell time. Lastly, individuals who paid cash added 7.1 s to the total dwell time, which is 18% lower than the estimated time of 8.7 s in Model 4. Similar to other studies, cash transactions add more dwell time than do electronic payment methods (27). On average, excess dwell time adds an additional 1 s to the total dwell time, which is consistent with that of Models 2 and 3.

Model 6 shows the expanded fare payment model, including a variable for an encumbered passenger boarding. The variables in this model are consistent with those of Model 5, as well as the time reported for passengers boarding with a smart card and cash. However, in this model, the dwell time added for a passenger boarding is 2.4 s compared to 2.5 s in Model 5 and 4.2 s in Model 4. The lower time estimate for boardings observed in Model 6 can be explained by the addition of the encumbered passenger variable, which would account for the additional time required to board a young child in a stroller. The model indicates that dwell time is expected to increase by 7.6 s for an encumbered passenger boarding or alighting from the bus, while keeping all other variables constant at their mean values. Consistent with Model 5 and the expanded dwell time models, excess dwell time adds an additional 1 s, on average, to the total dwell time, while keeping all other variables at their mean values.

In regard to the remaining control variables in Models 4 through 6, the coefficients generally follow the same sign and statistical significance and have a similar magnitude as those of Models 1 through 3, with the exception of the lack of statistical significance associated with dwells occurring at a bus stop with a connection to the metro. This observation requires further research in the future.

DISCUSSION OF RESULTS

Reducing dwell time at bus stops is expected to decrease overall running time and can improve reliability and speed (20). Dwell time can make up to 25% of the total running time of a bus (1). Passenger activity is a major component of dwell time; however, without careful knowledge of average time needed to serve passengers, transit agencies may be overestimating the scheduled running time. The purpose of this study was to estimate how accurately AVL–APC and fare box data are capturing the time associated with passenger activity from stop-level observations of dwell time. A series of dwell time

models were estimated from manually collected stop-level observation data to compare estimates from detailed dwell time models to models using data similar to what is reported by AVL–APC systems and fare boxes. While each of the models revealed coefficients and statistical significance of key variables expected to affect dwell time, the traditional model using data similar to that reported by AVL–APC systems overestimated the time of the first passenger boarding by approximately 2.5 times. The dwell time estimates of the traditional model are comparable to estimated boarding times in previous studies using AVL–APC data (3, 4, 23). This overestimation of time required for passenger activity was a result of excess dwell time likely captured by AVL–APC data. After accounting for excess dwell time, the estimated passenger activity time resembles Levinson's 1983 estimate of dwell time by using manually collected data (1). This excess dwell time occurs after passenger activity has commenced before the door closes. The manual data collection process employed in this study allowed the authors to capture details regarding the dynamics of passenger activity, details that are not currently well captured by AVL–APC and fare box data.

Identifying and reducing this bias imposed by AVL–APC data is critical for the improvement of automatic data collection methods. Schedulers use estimates from dwell time models to build route schedules. Accordingly, lack of knowledge about excess dwell time can result in an overestimation of the designated operational hours of a bus route. At the system level, such a number is expected to vary between different routes as a result of variation in the number of passenger activities, yet it will contribute a significant overestimation of operating costs that can be translated to thousands of dollars per day. The authors also added a variable for passenger movements involving encumbered passengers, allowing for a more detailed and accurate dwell time model. Passengers boarding with various encumberments, mobility restrictions, or strollers are examples of the dynamics of passenger activity which AVL–APC data fail to capture.

Fare payment methods have a substantial effect on dwell time. Comparing dwell time estimates of fare box data to stop-level observations of dwell time revealed an overestimation of the effect of each payment type on total dwell. When controlling for excess dwell, smart card payments had the least effect on dwell time, while cash transactions were associated with the highest additional dwell time. To reduce dwell time associated with passenger boarding, alternative fare collection methods are recommended, such as off-board payment methods or eliminating cash transactions. This model adds to the literature, because in previous studies disaggregate data regarding passenger boardings by fare type were not available (6, 24). Therefore, at this time, studies using manually collected passenger activity, similar to Fletcher and El-Geneidy (27), are vital to the understanding of dwell time associated with fare payment methods.

CONCLUSION AND RECOMMENDATIONS

The results indicate that without adjustments to the automatic data collection of dwell time and passenger activity, transit agencies are not capturing the full benefit of policies which aim to reduce the running time of bus routes. Estimates of boarding times derived from AVL–APC and fare box data can be misleading and can add a substantial amount of operating time to the schedules, leading to additional running time and causing delays for onboard passengers. To address this issue, improvements to AVL–APC data collection are recommended to capture excess dwell time. The time stamp of

each passenger boarding can be collected and recorded by the APC system, especially for the last boarding passenger, and can be used to identify the end of passenger activity. When combined with the door closing time, this information can enable transit agencies to identify the amount of excess dwell at every stop and adjust schedules accordingly. Some excess dwell time does need to be added to schedules intentionally to work as a cushion for route interruptions and delays, yet such an amount should be carefully calculated after fully understanding the dynamics of the dwell time along a route through detailed analysis, similar to the dwell time analysis conducted in this study. Furthermore, addressing the bias in dwell time resulting from encumbered passengers is also recommended. For example, an operator-facilitated button to record passenger activity involving a rider with a physical disability or various encumberments that does not require the use of the lift would increase the accuracy of dwell time estimates derived from AVL–APC and fare box data. Knowledge of the composition of patronage along a bus route, such as a route serving a high proportion of elderly passengers, can inform schedulers with the required modifications to the schedule.

Researchers argue that growth in public transport patronage can result from service reliability improvements, while growth can decrease because of unreliable service (*13, 14*). While AVL–APC and fare box systems provide transit agencies with a rich database for analysis for operational improvements, minor adjustments to such systems that can help in capturing excess dwell time are warranted to provide a more accurate picture of what is happening on the ground. Future research using manually collected data will remain a key resource for planners and researchers to enable advancements in the technology of automatically collected data for improved utilization.

ACKNOWLEDGMENTS

REFERENCES

1. Levinson, H. S. Analyzing Transit Travel Time Performance. *Transportation Research Record,* No. 915, 1983, pp. 1–6.
2. Fernández, R., P. Zegers, G. Weber, and N. Tyler. Influence of Platform Height, Door Width, and Fare Collection on Bus Dwell Time: Laboratory Evidence for Santiago de Chile. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2143, 2010, pp. 59–66. https://dx.doi.org/10.3141/2143-08.
3. El-Geneidy, A., and N. Vijayakumar. The Effects of Articulated Buses on Dwell and Running Times. *Journal of Public Transportation,* Vol. 14, No. 3, 2011, p. 63–86. https://doi.org/10.5038/2375-0901.14.3.4.
4. Dueker, K. J., T. J. Kimpel, J. G. Strathman, and S. Callas. Determinants of Bus Dwell Time. *Journal of Public Transportation,* Vol. 7, No. 1, 2004, p. 21–40. https://doi.org/10.5038/2375-0901.7.1.2.
5. Kathuria, A., M. Parida, C. R. Sekhar, and M. Pathak. Examining Bus Lost Time Dynamics for a Bus Rapid Transit Station. *Journal of Public Transportation,* Vol. 19, No. 2, 2016, p. 168–182. https://doi.org/10.5038/2375-0901.19.2.10.
6. Sun, L., A. Tirachini, K. Axhausen, A. Erath, and D. Lee. Models of bus boarding and alighting dynamics. *Transportation Research Part A: Policy and Practice,* Vol. 69, 2014, pp. 447–460. https://doi.org/10.1016/j.tra.2014.09.007.
7. Strathman, J., K. Dueker, T. Kimpel, R. Gerhart, K. Turner, P. Taylor, S. Callas, D. Griffin, and J. Hopper. Automated Bus Dispatching, Operations Control and Service Reliability: Baseline Analysis. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1666, 1999, pp. 28–36. https://dx.doi.org/10.3141/1666-04.
8. Bertini, R., and A. El-Geneidy. Modeling Transit Trip Time Using Archived Bus Dispatch System Data. *Journal of Transportation Engineering,* Vol. 130, No. 1, 2004, pp. 56–67. https://doi.org/10.1061/(ASCE)0733-947X(2004)130:1(56).
9. ITS Society of Canada. *ITS in Society—An Integration of Technologies.* https://www.itscanada.ca/it/society/index.html.
10. Furth, P., T. Muller, J. Strathman, and B. Hemily. Designing Automated Vehicle Location Systems for Archived Data Analysis. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1887, 2004, pp. 62–70. http://dx.doi.org/10.3141/1887-08.
11. El-Geneidy, A., J. Strathman, T. Kimpel, and D. Crout. Effects of Bus Stop Consolidation on Passenger Activity and Transit Operations. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1971, 2006, pp. 32–41. https://dx.doi.org/10.3141/1971-06.
12. Diab, E., M. Badami, and A. El-Geneidy. Bus Transit Service Reliability and Improvement Strategies: Integrating the Perspectives of Passengers and Transit Agencies in North America. *Transport Reviews,* Vol. 35, No. 3, 2015, pp. 292–328. https://doi.org/10.1080/01441647.2015.1005034.
13. Bates, J., J. Polak, P. Jones, and A. Cook. The Valuation of Reliability for Personal Travel. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 37, No. 2–3, 2001, pp. 191–229. https://doi.org/10.1016/S1366-5545(00)00011-9.
14. Noland, R., and J. Polak. Travel Time Variability: A Review of Theoretical and Empirical Issues. *Transport Reviews,* Vol. 22, No. 1, 2002, pp. 39–54. https://doi.org/10.1080/01441640010022456.
15. Hensher, D., P. Stopher, and P. Bullock. Service Quality—Developing a Service Quality Index in the Provision of Commercial Bus Contracts. *Transportation Research Part A: Policy and Practice,* Vol. 37, No. 6, 2003, pp. 499–517. https://doi.org/10.1016/S0965-8564(02)00075-7.
16. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.
17. Chien, S., S. Chowdhury, K. Mouskos, and Y. Ding. Enhancements of CORSIM Model in Simulating Transit Operations. *Journal of Transportation Engineering,* Vol. 126, No. 5, 2000, pp. 396–404. https://doi.org/10.1061/(ASCE)0733-947X(2000)126:5(396).
18. Transit Capacity. In *Special Report 209: Highway Capacity Manual,* 3rd ed. (1997 update). TRB, National Research Council, Washington, D.C., 1998.
19. Guenthner, R. P., and K. C. Sinha. Modeling Bus Delays Due to Passenger Boardings and Alightings. *Transportation Research Record,* No. 915, 1983, pp. 7–13.
20. Levine, J., and G. Torng. Dwell-Time Effects of Low-Floor Bus Design. *Journal of Transportation Engineering,* Vol. 120, No. 6, 1994, pp. 914–929. https://doi.org/10.1061/(ASCE)0733-947X(1994)120:6(914).
21. Lin, T., and N. Wilson. Dwell Time Relationships for Light Rail Systems. *Transportation Research Record,* No. 1361, 1992, pp. 287–295.
22. Zografos, K., and H. Levinson. Passenger Service Times for a No-Fare Bus System. *Transportation Research Record,* No. 1051, 1986, pp. 42–48.
23. Diab, E., and A. El-Geneidy. The Far Side Story: Measuring the Benefits of Bus Stop Location on Transit Performance. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2538, 2015, pp. 1–10. https://dx.doi.org/10.3141/2538-01.
24. Tirachini, A. Estimation of Travel Time and the Benefits of Upgrading the Fare Payment Technology in Urban Bus Services. *Transportation Research Part C: Emerging Technologies,* Vol. 30, 2013, pp. 239–256. https://doi.org/10.1016/j.trc.2011.11.007.
25. Milkovits, M. Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Data. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2072, 2008, pp. 125–130. https://dx.doi.org/10.3141/2072-13.
26. Jara-Diaz, S., and A. Tirachini. Urban Bus Transport: Open All Doors for Boarding. *Journal of Transport Economics and Policy,* Vol. 47, No. 1, 2013, p. 91–106.
27. Fletcher, G., and A. El-Geneidy. Effects of Fare Payment Types and Crowding on Dwell Time: Fine-Grained Analysis. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2351, 2013, pp. 124–132. https://dx.doi.org/10.3141/2351-14.
28. Tirachini, A. Bus Dwell Time: The Effect of Different Fare Collection Systems, Bus Floor Level and Age of Passengers. *Transportmetrica A: Transportation Science,* Vol. 9, No. 1, 2013, pp. 28–49.