Identifying the bias: Evaluating the effectiveness of automatic data collection methods in estimating the details of bus dwell time

Emily Grisé
School of Urban Planning
McGill University
Suite 400, 815 Sherbrooke St. W.
Montreal, Quebec, H3A 0C2
Canada
Tel.: 514-398-4058
Fax: 514-398-8376
E-mail: emily.grise@mail.mcgill.ca

Ahmed El-Geneidy
School of Urban Planning
McGill University
Suite 400, 815 Sherbrooke St. W.
Montreal, Quebec, H3A 0C2
Canada
Tel.: 514-398-4058
Fax: 514-398-8376
E-mail: ahmed.elgeneidy@mcgill.ca

October 2016

Word count: 5600 words, 2 figures, 3 tables

Paper accepted for presentation at the 96th Transportation Research Board Annual Meeting

For citation please use: Grisé, E. & El-Geneidy, A. (2017). *Identifying the bias: Evaluating the effectiveness of automatic data collection methods in estimating the details of bus dwell time.* Paper to be presented at 96th Transportation Research Board Annual Meeting, Washington D.C., USA.

ABSTRACT

Automatic Vehicle Location (AVL), Automatic Passenger Counters (APC), and fare box payments data 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 regarding the dynamics of passenger activity, which may bias the estimates associated with dwell or passenger activity time. The purpose of this study is to better understand any biases that might exist from using AVL/APC or fare box payment data when estimating dwell time. Manually collected data from Montreal, Canada is used to estimate detailed dwell time models. The study compares these estimates to models generated using data similar to what AVL/APC and fare box reports. The results reveal an overestimation in the passenger activity component of dwell time, which is mainly attributed to excess dwell time that AVL/APC and fare box payment generally do not capture. While AVL/APC and fare box technology provide transit agencies with rich data for analysis, adjustments to these 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 in order to generate better interventions that can reduce dwell times.

INTRODUCTION

Automatic vehicle Location (AVL), Automatic Passenger Counters (APC), and fare box payments data have been heavily used in the past years to generate dwell time models with the goal of recommending and/or evaluating improvements in efficiency and reliability of bus transit systems. These strategies can include, but are not limited to, all-door boarding strategies, smart card policy, off-board fare collection, low-floor vehicles, and the use of articulated vehicles. These particular strategies are used to reduce the dwell time needed for passenger activity, as dwell time can consume up to 25 percent of the total running time (*1*). Therefore, understanding the determinants of dwell time and time associated with its major components, including time required to open and close doors, passenger activity, and additional time until the door closing, have been extensively studied using either manual or automatically collected data (*2-4*).

Prior to 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 passenger activity per door, fare payment method, and unproductive door opening time (*4*). However, by the early 2000s, research began to study the application of AVL/APC data (*4; 7; 8*), where this data 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, however such methods provide a larger dataset for analysis at a lower cost. This tradeoff, can be minimized through introducing improvements in automatic data collection methods and through identifications of biases that are imposed from the 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, we use 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, Canada. The study compares estimates of detailed dwell time relying on manual count observations to estimates generated from models using data similar to what AVL/APC and fare box reports.

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

LITERATURE REVIEW

Intelligent Transport Systems (ITS) provides transport agencies with essential information and communication technologies to make informed decisions (*9*). In the public transit sector, AVL and APC are major components of ITS technologies that many agencies around the world are using or in the process of implementing due to their wide range of benefits to the 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 these 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 (*18*) indicated that the average time for passenger activity depends on many factors, such as the height of the bus or number of steps, fare collection system, and time associated with lift operation. In the past, manually collected data was 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 by Levinson in 1983. In this landmark study, estimates of bus dwell times revealed that each passenger boarding added 2.75 seconds to the constant dwell time of 5 seconds (includes door opening and closing time). Similarly, Guenthner and Sinha (*19)* estimated that each passenger boarding or alighting added 3-5 seconds to the dwell time. Accounting for fare payment type, Zografos and Levinson (*22)* estimated that a passenger boarding time took only 2 seconds on an uncrowded bus with an off-board fare payment system in Connecticut. Early examples of dwell time models such as the examples presented above, were based on smaller sample sizes due to the labour-intensive data collection methods required, resulting in lower explanatory power (*4*).

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

More recently, Diab and El-Geneidy (*23)* evaluated the impact of bus stop location on dwell time, and found longer dwell times associated with bus stops located on the near side of an intersection. Furthermore, the authors estimated that passenger boarding added 3.3 seconds and each passenger alighting added 1.9 seconds to the total dwell time. Evaluating the dwell time-savings of operating articulated buses, El-Geneidy and Vijayakumar (*3)* found an average passenger boarding time of 4.1 seconds and alighting time of 2.7 seconds, however time savings are found at the second and third doors where passengers are not required to scan their fare cards. From these select studies using AVL/APC data, we see that on average, the first passenger boarding takes between 3.3 – 4.1 seconds. However, returning back to Levinson's 1983 estimate of dwell times using manually collected data, each passenger boarding added approximately 2.75 seconds. These higher estimates of passenger boarding time may be attributed to the method of which AVL/APC data records 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 better understand the reasons for a discrepancy in the estimates.

Lastly, 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, due to 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 the knowledge of the 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 we noticed between earlier models generated from manual counts and current ones generated from automatically collected data.

METHODOLOGY

The objective of this analysis is to assess how well AVL/APC and fare box data are able to measure passenger activity through the generation of dwell time models. Dwell time is defined as the time between door open to door close, 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 open to door close, 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 due to various reasons, for example, bus holdings at time points, or a red light in the case of a near side 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. In this study we are interested in estimating the extent to which this excess dwell time, time with no passenger activity, impacts the estimate of passenger activity in dwell time models. We 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 end of passenger activity and time until door close, which we here-on refer to as excess dwell time.

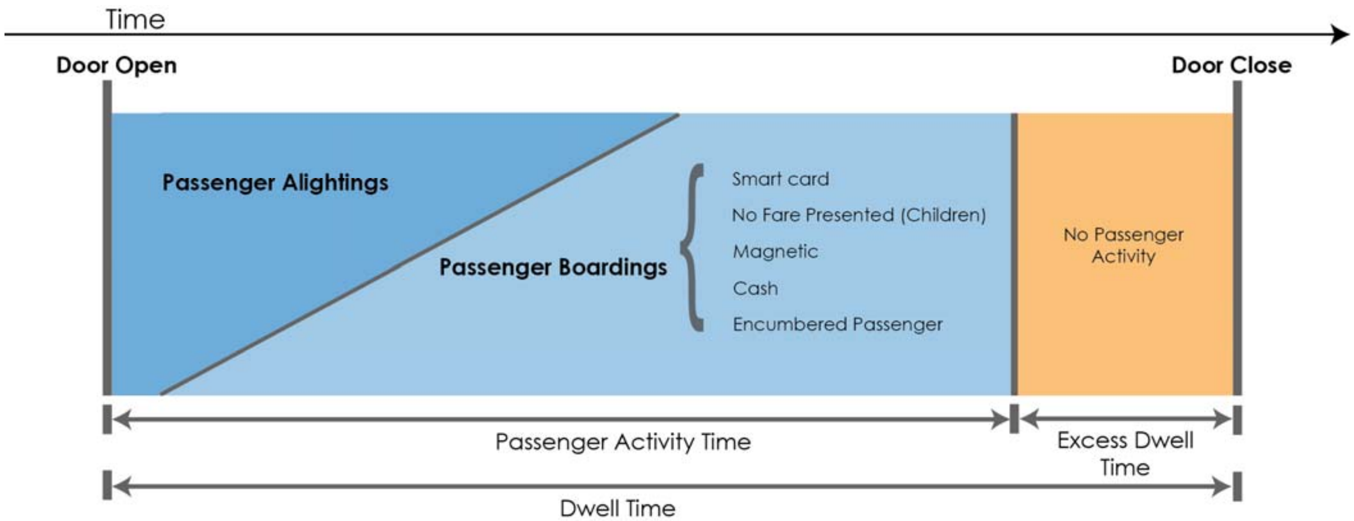

Figure 1 Elements of dwell time.

Based on previous research (*3; 4; 23; 28*) a typical dwell time model is 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 is the number of people who boarded the bus, while passenger alightings is the number of people who exited the bus. The squared term of passenger activity is included to capture the potential non-linear relationship of passenger activity, suggesting that each additional passenger requires 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 (*4*), and in this study is 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 our 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 these stops (*4*). The bus lift was not used during the data collection period, however we 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 our models, such as stops occurring on the near-side or far-side of an intersection, presence of a reserved bus lane, and bus shelters, however few were found significant and included in the final model. Finally, fare payment method is expected to impact dwell time, as previous research observed differences in passenger boarding times by different payment types (*27*).

Data Collection

The data for this study was collected from two bus routes in Montreal, Québec, Canada, which are both operated by the Société de transport de Montreal (STM), which is 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, operation in neighborhoods with comparable built environments, both routes share a connection with a metro station and both routes 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 the STM introduced a pilot all-door boarding policy along route 121 in March of 2016, which will be accounted for in our models through a route 121 dummy variable.

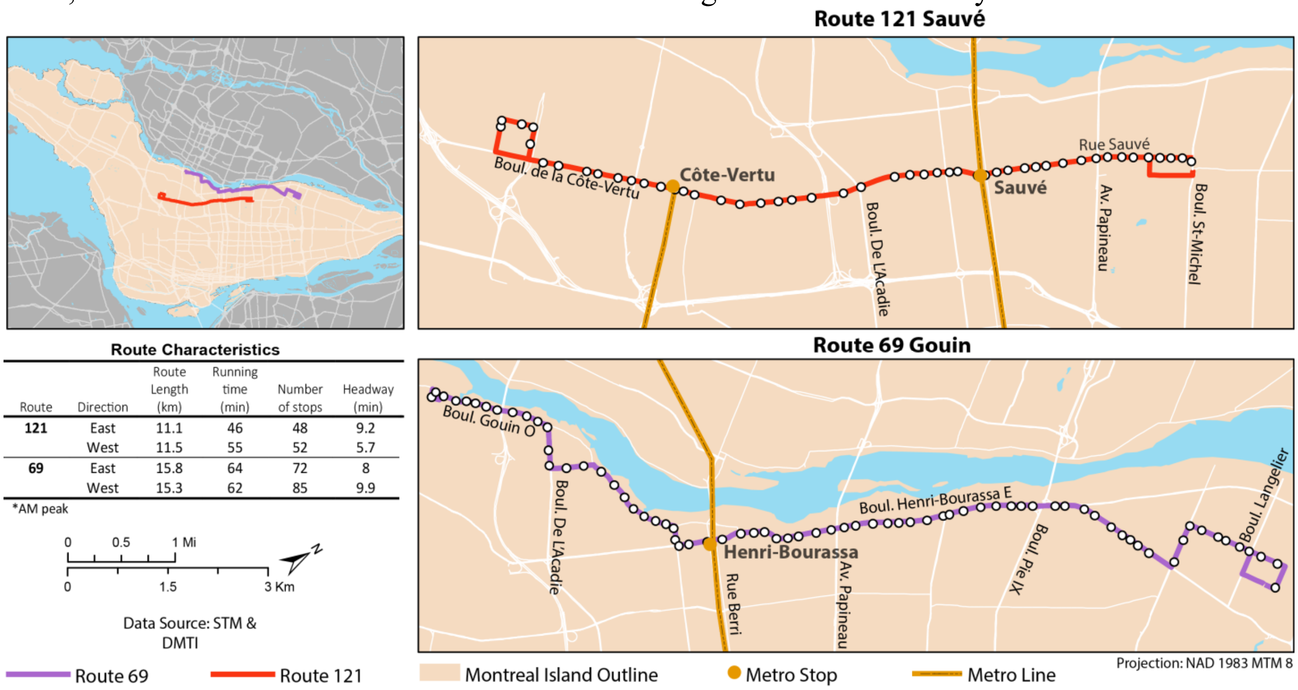

Figure 2 Context map of routes studied and relevant route characteristics.

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 the purposes of consistency, the same three people collected all the data. With regard to the dwell time recorded at each stop, we collected the time between door open to door close (total dwell time), the time of passenger activity (time after doors open to end of passenger activity), and we recorded 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, driver changes, etc.). 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 boardings and alighting throughout the route.

Characteristics of each stop along the route were also collected, such as stops occurring on the near-side or far-side of an intersection, presence of a reserved bus lane, and bus shelters. Furthermore, we observed which payment type was used by each passenger boarding that involved a fare transaction with the driver (as 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, we recorded the number of passengers under the age of 6 that boarded the front door of the bus accompanied by an adult, and were categorized as no fare payment as these users' ride transit free of charge. Payments by smart card on the STM bus network only require commuters to tap their pass 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 AM and 6:30 PM on a Tuesday, Wednesday, and Thursday in 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.

RESULTS

Model 1 is a traditional dwell time model, which employs a linear regression method, where we estimate the time required for each passenger boarding and alighting, with a series of independent variables meant to capture variations in dwell time between stops and the two routes. This model represents typical AVL/APC collected information. In Model 2 we add the exact amount of excess dwell which is noticed through the manual observation to understand the bias imposed on each coefficient when we do not include such a variable and only use AVL/APC like data. In Model 3 we then include a dummy variable for stops where we observed an encumbered passenger boarding, to predict the additional dwell time which can be expected to serve encumbered passengers, data which is currently not captured by AVL/APC data. Model 4 builds off the first three models by including detailed information on fare payment method 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 better understand the impacts of not including such information on a model derived from AVL/APC and fare box payment information. A further description of variables included in these models is presented in Table 1. Please note that layover stops (the first and final stops of each trip) are not included.

TABLE 1 Descriptive Statistics

Variable	Description	Mean	Std dev	Count
Total Boardings	Total number of passengers that boarded at all doors at a single bus stop during a trip.	1.95	4.64	NA
Smart card	Number of passengers that paid with a smart card.	1.54	2.74	1456
Magnetic card	Number of passengers that paid with a magnetic pass.	0.03	0.27	28
Cash	Number of passengers that paid with cash.	0.05	0.24	48
No fare presented	Number of passengers that boarded the bus without presenting fare.	0.05	0.28	44
Total Rear Door Boardings	Total number of passengers that boarded the 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 that alighted at all doors at a single bus stop during a trip.	2.23	5.41	NA
Total Passenger Activity ^2	The square of the 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 (sec).	6.17	25.76	NA
Route 121	A dummy variable equal to 1 if the stop occurs on Route 121.	0.77	0.42	NA
Friction	The total number of standees plus the sum of the total boarding and alightings at a stop.	4.67	9.74	NA
AM	A dummy variable equal to 1 if the trip took place between 6:30 am and 9:30 am.	0.34	0.48	NA
PM	A dummy variable equal to 1 if the trip took place between 3:30 pm and 6:30 pm.	0.24	0.43	NA
Midday	A dummy variable equal to 1 if the trip took place between 9:30 am and 3:30 pm.	0.42	0.49	NA
Passenger Load	The total number of passengers on a bus at the departure of a stop.	27.21	15.96	NA
Eastbound Trip	A dummy variable equal to 1 if the stop occurs on an eastbound trip.	0.49	0.50	NA
Metro Station	A dummy variable equal to 1 if the stop occurs at a metro station.	0.05	0.21	NA
Encumbered Passenger	A dummy variable equal to 1 if a passenger with an encumberment, disability or mobility limitation boarded or alighted the bus.	0.03	0.18	NA
Signalized intersection	A dummy variable equal to 1 if the stop occurs at a signalized intersection.	0.66	0.47	NA

Dwell Time Model

Table 2 shows the estimates and 95% confidence intervals for Models 1 – 3. First examining the traditional dwell time model (Model 1), the independent variables included in the model explain approximately 52 percent of the variation in dwell time. The constant variable (4.8 seconds) in this model represents a fixed amount of time that is associated with door opening and door closing. On average each passenger boarding adds 4.3 seconds to the base dwell time while keeping all other variables at their mean values. Each passenger alighting adds 2.1 seconds to the total dwell time, keeping everything else constant. Similar to previous research we notice a diminishing impact of the square term (passenger activity square). 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, the model variables explain approximately 95 percent 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 seconds compared to 4.8 in the first model, while each passenger boarding on average adds 1.8 seconds to the total dwell time keeping all other variables at their mean values. While each passenger alight adds 0.8 seconds to the constant dwell time of 3.3 seconds. 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 it is noticed from the passenger activity square coefficient. On average excess dwell time adds an additional one second 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 approximately 96 percent 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 seconds for an encumbered passenger boarding or alighting the bus, while keeping all other variables constant at their mean value. The constant variable and estimate of passenger alighting are consistent with Model 2, while each passenger boarding on average adds 1.7 seconds compared to 1.8 in Model 2, when all variables are controlled for. Remaining variables related to passenger activity behave consistent to Models 1 and 2.

For the remaining control variables in Models 1-3, dwell times at signalized intersections are slower compared to midblock stops or non-signalized 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 AM peak are estimated to have faster dwell times than midday trips, this can be attributed to a greater proportion of regular riders who may board using passes and ask fewer questions (*4*). Lastly, dwell times of eastbound bus stops were on average one second faster compared to westbound trips. The variable controlling for dwell time variation between bus routes 121 and 69 did not show a statistically significant effect on dwell time in the sample, although route 121 allowed all door boardings.

TABLE 2 Dwell Time Models

	Traditional Dwell Time (Model 1)			**Expanded Model (Model 2)**			**Expanded Model (Model 3)**		
Variable	Coefficient	95% Conf. Interval Lower Bound	Upper Bound	Coefficient	95% Conf. Interval Lower Bound	Upper Bound	Coefficient	95% Conf. Interval 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
AM	-4.55***	-7.81	-1.29	-0.70	-1.72	0.31	-0.76*	-1.74	0.23
PM (*ref= midday*)	-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
***=p<0.01, **=p<0.05, *=p<0.1	R-Squared 0.52			R-Squared 0.95			R-Squared 0.96		

Dwell Time Model with Fare Payment Type

Table 3 shows the estimates and 95% confidence intervals of Models 4-6 which estimate the variation in dwell time by each fare payment type, while otherwise keeping the models consistent with the previous dwell time model. Examining first the dwell time model with fare payment type (Model 4), the independent variables included in the model explain approximately 54 percent of the variation in dwell time. The constant variable in this model reports 3.3 seconds that is associated with door opening and door closing time 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 adds 4.7 seconds, while each passenger boarding with a magnetic card adds 21.8 seconds, on average to the dwell time 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 seconds on average. Finally, each passenger who pays with cash to the driver adds 8.7 seconds to the dwell time. Since route 121 allowed all-door boarding, passengers' who boarded at the second and third door increased the dwell time by 1.5 seconds. Consistent with Models 1-3, the time required for each additional boarding is lower than the previous one. 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 percent of the variation in dwell time. The constant variable reports 2.5 seconds, which is lower than the estimated constant of 3.3 seconds 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 are observed when including the excess dwell variable. Each passenger boarding with a smart card adds 2.5 seconds compared to 4.7 seconds in Model 4. Passenger boardings with a magnetic card are 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. Therefore we are unable to attain a statistically significant estimate of the average boarding time associated with magnetic card payments. Young children boarding with an adult add on average 4.2 seconds to the dwell time. Lastly, individuals who paid cash add 7.1 seconds to the total dwell time, which is 22% lower than the estimated time of 8.7 seconds in Model 4. Similar to other studies, cash transactions add more dwell time than electronic payment methods (*27*). On average excess dwell time adds an additional one second to the total dwell time, which is consistent with Models 2 and 3.

TABLE 3 Dwell Time Model with Fare Payment Types

	Traditional Fare Payment (Model 4)			Expanded Fare Payment (Model 5)			Expanded Fare Payment (Model 6)		
		95% Conf. Interval			95% Conf. Interval			95% Conf. 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
AM	-3.49**	-6.69	-0.30	-0.66	-1.57	0.26	-0.70	-1.60	0.20
PM (*ref= midday*)	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
***=p<0.01, **=p<0.05, *=p<0.1	R-squared 0.54			R-squared 0.96			R-squared 0.96		

Model 6 shows the expanded fare payment model, including a variable for an encumbered passenger boarding. The variables in this model are consistent with 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 boarding is 2.4 compared to 2.5 seconds in Model 5 and 4.2 seconds 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 seconds for an encumbered passenger boarding or alighting the bus, keeping all other variables constant at their mean value. Consistent with Model 5 and the expanded dwell time models, excess dwell time adds an additional one second on average to the total dwell time, keeping all other variables at their mean values.

In regards to the remaining control variables in Models 4-6, the coefficients generally follow the same sign and statistical significance and similar magnitude as Models 1-3, 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

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 AVL/APC and fare box reports. While each of the models revealed coefficients and statistical significance of key variables expected to impact dwell time, the traditional model using data similar to what AVL/APC reports overestimated the additional time of the first passenger boarding by approximately 2.5 times. These dwell time estimates of the traditional model are comparable to estimated boarding times of 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 using manually collected data. This excess dwell time occurs after passenger activity has commenced before the door closes. The manual data collection process employed in this study allowed us 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, for a route where 34,000 boardings are taking place per day the lack of knowledge about excess dwell time can add a total of 24 operation hours to ensure enough time is placed in the schedule to account for all passenger activities. At the system level, such a number is expected to vary between different routes due to 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. We then 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 an example 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, such as off-board payment methods or eliminating cash transactions are recommended. This model adds to the literature, as in previous studies (*6; 24*) disaggregate data regarding passenger boardings by fare type was not available. Therefore, at this time, studies using manually collected passenger activity, similar to (*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 operating time and delays to on-board 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 used to identify the end of passenger activity. This information when combined with the door closing time can enable transit agencies to identify the amount of excess dwell at every stop and adjust schedules accordingly. It is important to note that some excess dwell time does need to be added to schedules intentionally to work as a cushion for route interruptions and delays, yet such 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. The 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 it can decrease due to 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 these 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.

ACKNOWLEDGEMENT

We would like to thank Dea van Lierop and Geneviève Boisjoly from the Transportation Research at McGill (TRAM) group for their efforts in the data collection. This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada as well as the Social Sciences and Humanities Research Council.

REFERENCES

[1] Levinson, H. *Analyzing transit travel time performance*. 1983.
[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,* Vol. 2143, 2010, pp. 59-66.
[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. 4.
[4] Dueker, K., T. Kimpel, J. Strathman, and S. Callas. Determinants of bus dwell time. *Journal of Public Transportation,* Vol. 7, No. 1, 2004, p. 2.
[5] Kathuria, A., M. Parida, C. 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. 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.
[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,* Vol. 1666, 1999, pp. 28-36.
[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.
[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*, No. 1887, 2004, pp. 62-70.
[11] EI-Geneidy, A., J. Strathman, T. Kimpel, and D. Crout. Effects of bus stop consolidation on passenger activity and transit operations. *Transportation Research Record*, No. 1971, 2006, pp. 32-41.
[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.
[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, 2001, pp. 191-229.
[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.
[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.
[16] Highway Capacity Manual. Special report 209. *Transportation Research Board, Washington, DC,* Vol. 1, 1985, p. 985.
[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.
[18] Transportation Research Board. Transit Capacity: Chapter 12.In *Highway Capacity Manual, Spec. Rep. 209, 2nd Ed.* , National Research Council, Washington, D.C., 1997.
[19] Guenthner, R., and K. Sinha. *Modeling bus delays due to passenger boardings and alightings*. 1983.

[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.
[21] Lin, T., and N. Wilson. Dwell time relationships for light rail systems. *Transportation Research Record*, No. 1361, 1992.
[22] Zografos, K., and H. Levinson. *Passenger service times for a no-fare bus system*. 1986.
[23] Diab, E., and A. El-Geneidy. The far side story: Measuring the benefits of bus stop location on transit performance. *Transportation Research Record,* Vol. 2538, 2015, pp. 1-10.
[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.
[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,* Vol. 2072, 2008, pp. 125-130.
[26] Jara-Diaz, S., and A. Tirachini. Open all doors for boarding. *Journal of Transport Economics and Policy,* Vol. 47, No. 1, 2013, p. 15.
[27] Fletcher, G., and A. El-Geneidy. Effects of fare payment types and crowding on dwell time: Fine-grained analysis. *Transportation Research Record,* Vol. 2351, 2013, pp. 124-132.
[28] Tirachini, A. Bus dwell time: The effect of different fare collection systems, bus floor level and age of passengers. *Transportmetrica A: Transport Science,* Vol. 9, No. 1, 2013, pp. 28-49.