Fare Payment Structure and Dwell Time

A Dwell Time Modeling Method with Use of Automatic Passenger Count and Automatic Fare Count Data

S. Kiavash Fayyaz S., Cathy Liu

Dwell time at bus stops is a major component of bus operating time and impacts its variability. Even though dwell time largely depends on the number of boarding and alighting passengers, there are secondary factors such as fare type, bus design, and stop design that may affect it. These secondary factors may strongly influence the effectiveness of different strategies used to improve service. A detailed data needs to be collected in order to determine the impact of the fare payment methods. Manual data collection provides a detailed dataset, but it is labor intensive and the sample size is small. Automatically collected data doesn't have these shortcomings, but it doesn't provide detailed data for non-electronic fare payment methods (e.g. cash payment, prepaid paper tickets). This paper introduce a new method to determine the impact of detailed fare payment methods using data from automatic passenger count and automatic fare count systems installed on Utah Transit Authority BRT line 35 M. The results show that model can estimate the impact of non-electronic fare payment methods with less than 5% error. On average boarding time for passengers who use on-board cash fare payment is estimated to be 5.1 seconds longer than passenger who buy their tickets off-board from ticket vending machines. When the bus stop is placed on the median of roadway, on average the dwell time becomes 2.6 s shorter.

Transit Capacity and Quality of Service Manual (TCQSM) defines dwell time (DT) as the sum of passenger service time, boarding lost time, and door opening and closing time (TCQSM, 3rd edition). DT is a major component of bus travel time (Rajbhandari et al. 2003). As ridership increases, passenger service time will increase leading to increase in DT. Consequently increasing bus run time variation and decreasing operation performance. Therefore understanding the nature of factors influencing DT will help transit authorities planning more effectively.

Passenger service time is understood to be the largest component of DT and influenced by passenger demand, fare payments, vehicle configuration, passenger load, door usage and platform configuration (TCQSM). For example average boarding time for passenger paying by cash is 2s more than average boarding time for passenger who has prepay pass (Fletcher and El-Geneidy 2013). There are also other factors such as atypical passenger boarding, passenger age, time of the day and fare payment issues that can affect DT significantly.

Researcher have used Electronic (Rajbhandari et al. 2003; Bertini et al. 2004; Dueker et al. 2004; Milkovits 2008) and Manual (Aashtiani and Iravani 2002; Fletcher and El-Geneidy 2013; Tirachini 2011) data collection for analysis of bus DT. However each of these data collection methods have its advantages and disadvantages. Manual data collection provides detailed and more accurate data, but it typically involves labor-intensive ride checks (Dueker et al. 2004). Electronic data collection methods such as Automatic Passenger Counter (APC), Automatic Vehicle Location (AVL) and Automatic Fare Counting (AFC), provide a large sample of data with no additional cost (if already installed on the buses) compared to manual data collection method. On the other hand they lack detailed information on non-electronic fare media (Cash and Prepaid paper tickets) and also there are concerns about data validity and reliability and loss of detail (Fletcher and El-Geneidy 2013). Previous studies on dwell time have used classic linear regression (CLR) to model DT. As in automatically collected data, the non-electronic fare payers aren't counted separately, the model comes short in estimating their impact separately. This paper tries to utilize ordinary least squares (OLS) regression and genetic algorithm (GA) to find the impact of different factors affecting dwell time including separated non-electronic fare payments impacts based on APC/EFC data.

LITERATURE REVIEW

Prior research on DT from manually collected data measures the impact of boarding and alighting passengers, fare payment method, crowding, vehicle configuration, passenger age and driver characteristic. Kraft and Bergen found that average passenger service time for off-peak hours are greater than peak hours. They also found out that DT per boarding passenger for cash and change fare is 3 seconds more than for exact fare (Kraft and Bergen 1974). Levinson related the bus DT to the total number of passengers boarding and alighting (Levinson 1983). Lin and Wilson develop a functional form that estimates the dwell time per door and find that the crowding effect is best captured nonlinearity (Lin and Wilson 1992). Ashtiani and Iravani analyzed the influence of passenger boarding and alighting, load factor (crowding) and number of doors used to model the DT for different bus types based on 3454 observations, DT model then used as an input to be used in transit assignment models (Ashtiani and Iravani 2003). Fletcher and El-Geneidy found out crowding effect significantly increased dwell time after 60% of Bus capacity was surpassed (Fletcher and El-Geneidy 2013). In this study they calculated the effects of crowding and passenger boarding/alighting by fare payment method based on 1764 observations. Cash fare payers estimated to have a 2 seconds slower transaction and boarding time than prepaid pass holders. Tirachini analyzed the effects of fare payment methods, age, bus floor level and crowding on DT based on 1604 observations (Tirachini 2011). Tirachini obtained DT per passenger boarding paying with cash exact fare to be 7 seconds less than passenger paying with cash change given. These studies data sets are limited to relatively small sample of observations due to cost and time required for manual data collection. Small sample size raise concerns about providing adequate information for an accurate analysis and also unable to address rare (atypical) conditions.

On the other hand automatically collected data avoid these problems but with a loss of detail. Prior research into DT from automatically collected data measures the impact of boarding and alighting passengers, fare payment method, crowding, vehicle configuration, time of the day, schedule adherence and atypical passenger boarding. Rajbhandari et al. found that nonlinear model of total boarding and alighting together explains the DT better than the linear model, when there is no detailed information on payment methods and door used (Rajbhandari et al. 2003). Dueker et al. explored the impact of lift operation (estimating 62 seconds per lift operations), time of day (finding that the morning peak has the shortest dwell time), low floor bus effect, friction (crowding) and schedule adherence on dwell time based on APC data of 350,000 observations (Dueker et al. 2004). However this study doesn’t address the effects of fare payment methods (probably due to lack of AFC data). Milkovits analyzed the impact of fare payment methods (finding that smart media cards are roughly 1.5 seconds faster to process than magnetic stripe tickets) on DT based on 173,750 observation from APC, AFC and AVL data (Milkovits 2008). However this study doesn’t estimate impacts of fare payment methods that do not have AFC record associated with them separately, for instance based on Tirachini’s study a passenger paying with cash change given will have about 11 seconds longer boarding time than a school students (that doesn’t require to pay or show any proof). Measuring the impact of cash payment becomes an important issue considering public transit agencies looking into eliminating cash payments (bus fare box) to improve bus lines (e.g. to BRT lines).

TCQSM suggests average passenger service time of 4.5 seconds with exact change into fare box, 2 seconds with visual inspection and 1.75 seconds for no fare payment. If the passenger doesn’t have the exact change the service time will be more than 4.5 seconds, Tirachini found it to be about 12 seconds (Tirachini 2011). Thus, it is important to find out the number of passengers paying with each of these methods (where no AFC record available) and their impact on dwell time. This paper aims to define a systematic method to calculate the impact of detailed fare payment methods (including no AFC record associated methods) based on APC and AFC data. This paper aims to define a systematic method to calculate the impact of detailed fare payment methods (including no AFC record associated methods) based on APC and AFC data.

DATA COLLECTION METHODS AND VARIABLES

Data used in this study are collected both automatically through the APC and AFC systems installed on UTA buses for bus rapid transit (BRT) line 35M, and manually for limited time. Automatic observations were collected from UTA bus route 35M during the entire month of May 2014 with a total of 65536 observations for 28 stops. AFC and APC data were then matched based on the following conditions; a) AFC and APC records has the same date, b)same station and c) the time difference between records is less than 2 minutes (to consider any measurement error), if the two line matched then an AFC record added to related APC observation. In addition limited observations were collected both manually and automatically to compare the modeling results with reality. These observations collected during February 10th 2015 for UTA bus route 35M. Manual collected data consist of 120 observations during 7 to 9 AM, 11 AM to 1 PM and 4 to 6 PM. Automatic data were collected using APC and AFC system.

Preliminary filtration were applied based on the following constraints; APC data collectors malfunction were considered as Dwell time more than 3 minutes or Dwell time divided by the number of boarding passengers be less than 1 second. Terminal observations were also excluded because it included the drivers' layover time. DT is generally controlled by number of passenger boarding (BC), number of passenger alighting (AC), or atypical activity. For each of these situation a separate model is needed. So the observations were categorized in these three categories (the categorization process is shown in appendix). Finally 7725 observations found to be controlled by boarding (BC).

BRT 35M buses have low floors, are equipped with smart card readers at all doors, have 28 seats (total capacity is 60 passengers), and have three doors. All three doors are available for both boarding and alighting, but passengers who want to buy their ticket from the driver (exact change into the fare box) are limited to board from the front door. The drivers are instructed to utilize all doors for boarding-alighting and no fare visual inspection, but still some drivers do not follow. The summary statistics of the data are presented in Table 1.

Variables

- DT: dependent variable that measures the time (in seconds) between door open and close.
- BT: Total number of boarding passengers.
- AT: Total number of alighting passengers.
- B-EFC: independent variable for the number of passengers boarding using electronic fare payment
- A-EFC: independent variable for the number of passengers alighting using electronic fare payment
- B-TVM*: variable for the number of passengers boarding using tickets (bought from ticket vending machines) or passengers who do not pay. Note that the boarding and alighting time for the TVM and Non-Payers are almost the same as there is no visual inspection.
- B-Cash*: variable for the number of passengers boarding who buy ticket from bus driver. Note that cash payers are limited to board from the front door.
- B-CTVM: Sum of B-TVM and B-Cash within the dwell time observation. This variable is equal to BT – B-EFC.
- A-CTVM: independent variable for the number of passengers alighting who paid fare by TVM or Cash. There is no inspection or transaction related activity associated with Cash and TVM payers while alighting. This variable is equal to AUT – A-EFC.
- Weekend: indicator variable that shows whether the observation is collected on weekend (equal 1) or on weekdays (equal to zero).
- Door-Cycle: variable that shows how many time bus doors were opened and closed in the observation.
- Stop-[station name]: indicator variables shows the station name that the observation were collected.

* B-TVM and B-Cash variables are not collected with APC or AFC systems, only sum of them (B-CTVM) can be extracted from automatically collected data.

TABLE 1 Summary Statistics

Variable	Obs.	Mean	Std. Deviation	Min	Max	Sum
DT	7725	17.055	13.270	4.8	169.8	-
Weekend	7725	0.084	0.277	0	1	-
B-EFC	7725	0.263	0.610	0	8	2033
B-CTVM	7725	2.689	2.282	0	18	20770
B-TVM[a]	7725	1.743	2.163	0	18	13467
B-Cash[a]	7725	0.935	1.605	0	12	7225
A-EFC	7725	0.051	0.235	0	3	396
A-CTVM	7725	0.917	1.443	0	16	7087
Door-Cycle	7725	1.288	0.496	0	5	-
Fair-Mall stop (Magna direction)	7725	0.083	0.276	0	1	-
3575 W stop	7725	0.035	0.185	0	1	-

3955 W stop	7725	0.059	0.235	0	1	-
Fair-Mall stop (TRAX direction)	7725	0.035	0.183	0	1	-
1685 W stop	7725	0.049	0.215	0	1	-

[a] B-Cash and B-TVM's summary statistics are based on estimated results of model estimation and testing section.

METHODOLOGY

Assume that DT can be calculated as a linear function of vector of collected independent variables, plus a disturbance term.

(1) $$DT_i = \beta * X_i + \varepsilon_i$$

Where DT_i is DT for i^{th} observation, β is the vector of estimable coefficients, X_i is the vector of measurable characteristic that determine DT for i^{th} observation and ε_i is the disturbance term. Then assuming that all classical linear regression (CLR) model assumptions holds, the ordinary least squares estimator is the optimal estimator.

Let's construct X_i as the independent variables including B-CTVM and not including B-Cash and B-TVM. In this case estimated coefficient for B-CTVM captures part of the combined effect of B-Cash and B-TVM variables, and the rest will be captured by disturbance terms,

(2) $$\beta_{B\_CTVM} * B\_CTVM_i + \lambda_i * \varepsilon_i = \alpha_{B\_Cash} * B\_Cash_i + \alpha_{B\_TVM} * B\_TVM_i$$

Where α_{B_Cash} and α_{B_TVM} are the specific boarding time for B-Cash and B-TVM, and λ_i is the portion of disturbance term caused by excluding B-Cash and B-TVM variables from the model for i^{th} observation. Assuming that X_i include all the independent variables that impact dwell time and there is no other factors that cause disturbances (e.g. measurement errors), then most of the disturbances are caused by exclusion of B-Cash and B-TVM. In other words $\lambda_i = 1$ for all observations. So equation (2) will reduce to,

(3) $$\beta_{B\_CTVM} * B\_CTVM_i + \varepsilon_i = \alpha_{B\_Cash} * B\_Cash_i + \alpha_{B\_TVM} * B\_TVM_i$$

Where,

(4) $$B\_CTVM_i = B\_Cash_i + B\_TVM_i$$

It is expected that the boarding time for passengers who buy tickets from the driver paying with cash be more than boarding time for passengers using tickets bought from TVM, in other words $average\ (\alpha_{B\_Cash}) > average\ (\alpha_{B\_TVM})$. Also as mentioned above the estimated coefficients for B-CTVM variable captures a combined (or average) effects of B-Cash and B-TVM variables, so,

(5) $$average\ (\alpha_{B\_Cash}) > \beta_{B\_CTVM} > average\ (\alpha_{B\_TVM})$$

Based on equation (5), previous studies and field observations, thresholds for fastest and slowest boarding time for cash and TVM payment methods can be set,

(6) $$C_{slowest} > \alpha_{B\_Cash} > C_{fastes}$$

(7) $$TVM_{slowest} > \alpha_{B\_TVM} > TVM_{fastes}$$

Considering that B-Cash and B-TVM are integers, genetic algorithm can be used to solve the equations (3), (4), (6) and (7) to estimate the B-Cash and B-TVM variables. Then OLS is utilized to estimate new coefficients, in the model where B-CTVM is replaced by these variables. Estimation concerns relating to this method and how to test the model validation are discussed in model estimation and testing section of this paper.

MODEL ESTIMATION AND TESTING

The first step in implementing our method is to find the best model specification. Best model specification must include as many independent variables as possible, so that λ's value become closer to 1. The results is shown in table 2, B-CTVM model.

TABLE 2 Model Results

Model	B-CTVM Model		B-TVM & B-Cash Model	
DT	Coefficient	t-stat	Coefficient	t-stat

Weekend	1.411	4.04	1.087	6.30
B-EFC	4.992	30.64	5.279	65.57
B-CTVM	3.329	71.46		
B-TVM			1.803	73.03
B-Cash			6.917	211.68
A-EFC	2.623	6.28	2.020	9.79
A-CTVM	1.741	23.14	1.611	43.40
Door-Cycle	1.580	8.06	1.509	15.60
Fair-Mall stop indicator (Magna dir.)	2.478	6.36	2.116	11.00
3575 W stop indicator	-2.598	-4.95	-2.588	-9.98
3955 W stop indicator	2.222	5.38	1.617	7.92
Fair-Mall stop indicator (Trax dir)	3.766	6.72	3.432	12.40
1685 W stop indicator	2.479	5.44	2.287	10.15
Constant	2.411	8.32	2.026	14.15
Adjusted R-Squared =	0.5937		0.9009	

TCQSM 3rd edition, in exhibit 6-4, recommends a range of 1.75 to 2.5 seconds for average boarding time for passengers with no fare payment (off-board collected fare) and a range of 3.1 to 8.4 seconds for average boarding time for passenger using exact change into fare-box. Note that, here no fare payment and exact change into fare-box are referring to B-TVM and B-Cash. The best model estimated 3.3 seconds average boarding time for B-CTVM. Based on these, the thresholds for equations (6) and (7) are set as follows; $3.3\ seconds > \alpha_{B\_Cash} > 7.3\ second$, $1.3\ seconds > \alpha_{B\_TVM} > 4.3\ seconds$. Then equations (3) and (4) were solved together using GA, and $B\_Cash_i\ and\ B\_TVM_i$ were estimated. The new model estimation where B-CTVM was replaced by these results, is shown in table 2, B-TVM & B-Cash model.

We recommend three different tests to check the validation of the results. First test checks the consistency of coefficients of common variables in two models, second test check the estimated B-Cash and B-CTVM with the manually collected data, and finally the third test check the effect of estimation error in coefficients.

The seemingly unrelated estimation in STATA (Weesie 1999) were used to test the consistency of the coefficients of common variables in two model. The null hypothesis were that all the coefficients of common variables in two model being equal. The results ($\chi^2 = 8.09$, Pr > $\chi^2 = 0.62$) shows that null Ho is accepted, thus the coefficients of common variables are consistent in two models.

Equations (3) and (4) based on the same thresholds were solved for automatically collected observations, then the estimated value for B-Cash and B-TVM were compare with manually collected observations (real values). The estimated B-Cash were exact match with real values 92% times, and the total number of B-Cash were 10% more than total number of real cash payers. The errors are probably caused by small sample size of manually collected data and the fact that not all the cash boarding follow the thresholds set in the model.

Finally, to test the impact of estimation errors on coefficients, error range of [-10%, 15%] of total B-Cash were chosen based on second test's results. To implement the error range to the data set, random probabilities were generated and when the random probabilities were less than our error range percentage, one passenger were subtracted or added to B-Cash. Then the new B-Cash and B-TVM were put in the model. This process has been repeated 20 times. The estimation errors of B-TVM and B-Cash's coefficients were fairly small (less than 0.1 second) in the context of bus dwell time modeling, and thus the results are acceptable. In other words, the coefficients of the model are biased, but the biasedness of them are fairly small and unimportant in dwell time modeling.

INTERPRETATION OF THE RESULTS

The model shows good statistical fit with adjusted R-squared of 0.90. All variable coefficients included in the specification are statistically significant and have plausible signs. The results show that on weekends DT is on average one second longer than weekdays. This is caused by the fact that usually passengers are not in hurry on weekends and thus take their time to board or alight. The model predicts 1.5 seconds for each time that a bus door-cycle. Dead-time can be estimated as sum of the constant term and door-cycle time, so $Dead\ time = 1.5 * door_{cycle} + 2\ seconds$.

Average boarding time for passengers who use on-board fare collection methods (i.e. EFC and Cash) are much larger than off-board fare collection method (i.e. TVM). Average boarding time for EFC, Cash, and TVM fare payments are estimated to be 5.3, 6.9, and 1.8 seconds respectively. Our field observations shows that smart card readers installed on the bus takes about 1 second to read each card, this will cause longer boarding time for EFC payers than TVM. In addition, long DT for cash payers is caused by transaction time associated with each cash payers.

Average alighting time for TVM and Cash payers are the same as they don't have any transaction associated with this activity, on the other hand EFC payers have check-out transaction time. Average alighting time for passenger who use TVM or Cash is estimated to be about 1.6 seconds, and 2 seconds for EFC payers.

Finally, the impact of built environment and stop design is captured by stop indicator variables in the model. DT for Fair-mall stop is estimated to be on average 2.1 and 3.4 seconds (for each direction) longer than other stops, this is probably caused by the longer service time required for passengers who are carrying shopping bags. DT for 3955 W and 1685 W stops are estimated to be on average 2.2 and 2.5 seconds (respectively) longer than other stops. 3955 W stop is placed in front of the hospital, it is expected to observe longer DT because of sick passenger require longer service time. 1685 W is a transfer stop, between bus route 35M and bus route 217. It is expected that the driver will delay at this stop for passenger to complete their transfer from route 217 to 35M. Finally the impact of stop design can be properly seen in 3575 stop (estimated 2.6 seconds shorter on average). Shorter dwell time for 3575 W stop is probably caused by the fact that, 3575 W is the only stop on bus route 35M that is placed on the median of roadway. This is probably caused by the fact that it is easier for driver to handle the stopping process as he has no concerns about traffic.

CONCLUSION

In this paper we introduced a new method to estimate the number of passenger boarding that have no AFC record associated with them and use different fare payment method. Then estimate the effects of each fare payment method using classic linear regression. The advantage of this method is that it does not require manual data collection, which is a labor and time intensive procedure. Another advantage of this methods is using large sample of data provided by APC and AFC. We also recommended three different tests to check for the result validation.

The model results shows the on-board fare collection methods can take 3.5 to 5.1 seconds longer boarding time per passenger. In addition, the built environment (i.e. stop indicators) and median placed stop have noticeable impact on dwell time. The model result can be used to predict dwell times for future year with increased ridership. For future research, we recommend to further explore the impact of built environment, socioeconomic parameters of passenger, and stop design model on bus dwell time.

ACKNOWLEDGEMENT

Funding for this research was provided by the Utah Transit Authority. The authors thank Dr. RJ Porter and Dr. Milan Zlatkovic for their guidance throughout this study.

REFERENCES

1. Transit Capacity and Quality of Service Manual, Third ed. Transportation Research Board of the National Academics, Washington, D.C.
2. Fletcher, G., El_Geneidy, A. Effects of Fare Payment Types and Crowding on Dwell Time (Fine-Grained Analysis), Transportation Research Record: Journal of the Transportation Research Board, No. 2351, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 124–132.
3. Tirachini, A. Bus Dwell Time: The Effect of Different Fare Collection Systems, Bus Floor Level and Age of Passengers. Transportmetrica, 2011, pp. 1–22.
4. Milkovits, M. N. Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Data. In Transportation Research Record: Journal of the Transportation Research Board, No. 2072, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 125–130.
5. Bertini, R. L. and A. M. El-Geneidy. 2004. Modeling transit trip time using archived bus dispatch system data. Journal of Transportation Engineering, 130 (1): 56-67.
6. Dueker, K. J., T. J. Kimpel and J. G. Strathman (2004). Determinants of bus dwell time. Journal of Public Transportation 7(1): 21-39.
7. Rajbhandari, R., S. Chien and J. Daniel (2003). Estimation of Bus Dwell Times with Automatic Passenger Counter Information. Transportation Research Record 1841: 120-127.
8. Aashtiani, H. and H. Iravani (2002). Application of Dwell Time Functions in Transit Assignment Model. Transportation Research Record 1817: 88-92.
9. Moore, J. E., II, G. Giuliano, and J. G. March. Performance of Automatic Passenger Counters in the Ventura/Lompoc Smart Card Demonstration Project. Journal of Intelligent Transportation Systems, Vol. 7, No. 2, 2002, pp. 131–150.
10. Lin, T., and N. H. M. Wilson. Dwell Time Relationships for Light Rail Systems. In Transportation Research Record 1361, TRB, National Research Council, Washington, D.C., 1992, pp. 287–295.
11. Levinson, H. S. (1983). Analyzing Transit Travel Time Performance. Transportation Research Record 915: 1-6.
12. Kraft W. and T. Bergen. 1974. Evaluation of passenger service times for street transit systems. Transportation Research Record, 505: 13-20.
13. Baltagi, B. H. 2013. Econometric Analysis of Panel Data. 5th ed. Chichester, UK: Wiley.
14. Hausman, J. A. 1978. Specification tests in econometrics. Econometrica 46: 1251–1271.
15. Weesie, J. 1999. sg121: Seemingly unrelated estimation and the cluster-adjusted sandwich estimator. Stata Technical Bulletin 52: 34–47. Reprinted in Stata Technical Bulletin Reprints, vol. 9, pp. 231–248. College Station, TX: Stata Press.

APPENDIX

Observation Categorization Process

Bus DT is generally controlled by the number of people boarding-alighting or atypical activity (e.g. bike boarding, disable passenger boarding-alighting). Whether the all door boarding-alighting (simultaneously) is allowed on the bus or not the DT can change. The automatically collected data available for this study is aggregated boarding and alighting for all doors which limits the study to assume one door boarding-alighting model. So the DT can be modeled as;

$$DT = \max\left\{\sum_{i=1}^{A} t_i^a, \sum_{i=1}^{B} t_i^b, Atypical\right\} + deadtime$$

Where A is the number of passengers alighting, B the number of passengers boarding, and t_i^a and t_i^b the time that each passenger takes to alight and board, respectively (Tirachini 2011 suggest the same model without Atypical factor). This model formulation shows that three separate models are needed for BC, AC and Atypical situations. When Dwell time divided by number of boarding passengers were bigger than 10 seconds or Dwell time divided by number of alighting passengers were bigger than 5 seconds, it was considered as the Atypical passenger activity. BC and AC categories were divided using GA based on following formulation;

$$GA: \min(abs(DT_{estimated} - DT_{actual})$$

Where:

$$DT_{estimated} = \max\left\{\sum_{i=1}^{A} t_i^a, \sum_{i=1}^{B} t_i^b\right\} + deadtime$$

$$\sum_{i=1}^{A} t_i^a = B_{EFC} * T_{EFC} + B_{Non-EFC} * T_{Non-EFC}$$

$$\sum_{i=1}^{B} t_i^b = A * T_{alighting}$$

With the following constraints:

- B_{EFC} is total boarding passengers using electronic fare payment
- T_{EFC} is average time associated with EFC boarding per each passenger (2 s =< T_{EFC} =< 8 s)
- $B_{Non-EFC}$ is total boarding passengers using non-electronic fare payment (including exact change into the fare box and prepaid paper tickets) or not paying. Note that the actual number of passengers using each non-electronic fare payment methods is not provided in automatically collected data as they don't have any electronic foot prints.
- $T_{Non-EFC}$ is average time associated with Non-EFC boarding per each passenger (2 s =< $T_{Non-EFC}$ =< 10 s)
- A is total number of alighting passengers
- $T_{alighting}$ is average time associated per each alighting passenger (1 s =< $T_{alighting}$ =< 5 s)

GA will find feasible estimates for T_{EFC}, T_{Non_EFC}, $T_{alighting}$ and consequently whether the observation DT is boarding or alighting controlled. Finally 7725 observations found to be controlled by boarding (BC).