Modeling Dwell Time And Calculating Emission Reduction By Introducing Off Board Fare Collection

Adrita Islam

CE 5090

Presentation Outline

- Introduction and motivation
- Literature review
- Data and data preparation
- Descriptive analysis
- Model formulation and results
- Model Diagnostic
- Emission reduction
- Discussion and Conclusion

Introduction and motivation

- Dwell time - interval between the opening and closing of bus doors to serve passengers at the bus stop.
- Reducing dwell time is an important factor in :
  - Improving travel time between terminals
  - Reliability of the service
  - Reducing idle emissions.
    - Idle fuel consumption per for transit bus (Diesel) is about 0.9 gal/hr. Which can go up to 1.85 gal/hr.
    - 10 minutes of dwelling is equivalent to 5miles of driving

Introduction and motivation

- In order to reduce dwell time, we first need to model it.
- No studies have been done in the state of Connecticut to model dwell time using automated data.
- The aim of this project - model dwell time of CTtransit buses in Hartford and examine off-board fare collection as a potential candidate for dwell time reduction.
- Expectation - a significant reduction in dwell time due to off-board fare collection (significant energy savings)

Literature Review:

Author	Method of Data collection	Model Specification/ Variables considered		R^2
Guenthner and Sinha (1983)	Manual	T = 5.0 - 1.2[ln (N)]		0.36
Guenthner and Hamat (1988)**	Manual	T=2.25 + 1.81A		0.667
		T= -0.27 + 5.665B		0.906
Arhin, Noel, Anderson, Williams, Ribbiso and Stinson (2015)	Manual	Boarding Parking No of lanes	Alight Bus bay length	0.72-0.92
Dueker, Kimpel, and Strathman (2004)	Automated	Boarding, $Board^2$ Type of route Lift operation	Alight, $Alight^2$ Time of day	0.348
Rajbhandari, Chien, and Daniel (2004)	Automated	DT = a + b (Total)		0.642
		DT = a + b (Ons) + c (Offs)		0.718
		DT = a + b (Total) + c (Total)(S)		0.643
		$DT = a (Total)^b$		0.753
Shockley, Salinas, and Taylor (2015)**	Automated	Board Type of fare Bus specifications	Alight Time of day Irregular Activity	0.45 0.49 (congested)

** Complex Fare Structure

Data

- APC data from CTtransit
- Duration – October 1, 2016 to November 30, 2016
- Total number of stops in the dataset is = 2273049
- Total Number of routes in the dataset– 67

Dwell time is a continuous variable – Linear Regression was chosen as appropriate modeling method

Data preparation

- Filtration:
  - Number of stops where bus door was opened = 1601987
  - Number of stops with dwell time between zero and 180 seconds (Duker Paper) is = 1516956
  - Number of stops with only local bus = 1469757

	dwell_time_express	dwell_time
count	47136.000000	1.516928e+06
mean	43.217032	1.997309e+01
std	41.830247	2.621977e+01
min	2.000000	2.000000e+00
25%	11.000000	6.000000e+00
50%	27.000000	1.000000e+01
75%	63.000000	2.100000e+01
max	180.000000	1.800000e+02

A sample of 5% of the filtered data was used for the final analysis

- Total number of stops included in analysis is = 73488
- Total Number of routes included in analysis is = 45

Variables:

- Dependent variable:
  - Dwell time – time interval (in seconds) between the opening of first door and closing of last door
- Independent Variables:
  - Board – number of people boarding the bus
  - Alight– number of people alighting the bus
  - Board2 – square of number of people boarding the bus
  - Alight2 – square of number of people alighting the bus

Variables:

- Independent Variables: (dummy)
  - on_time – $\begin{cases} 1, & Bus\ arrives\ between\ 0\ to\ 300\ sec\ (5\ mins) of\ scheduled\ time \\ 0, & any\ other \end{cases}$
  - early– $\begin{cases} 1, & Bus\ arrives\ before\ scheduled\ time \\ 0, & any\ other \end{cases}$
  - off_board_fare_col– $\begin{cases} 1, & is\ CTFastrack \\ 0, & is\ not\ CTFastrack \end{cases}$
  - weekday – $\begin{cases} 1, & is\ a\ weekday \\ 0, & is\ not\ a\ weekday \end{cases}$
  - am_peak– $\begin{cases} 1, bus\ runs\ during\ am\_peak (6{:}30 - 9{:}30am) \\ 0, bus\ runs\ during\ any\ other\ time \end{cases}$
  - pm_peak– $\begin{cases} 1, bus\ runs\ during\ pm\_peak (3{:}30 - 7{:}30pm) \\ 0, bus\ runs\ during\ any\ other\ time \end{cases}$
  - Irregular_activity– $\begin{cases} 1, if\ dwell\ time\ is\ unusually\ large (> 60sec) \\ 0,\ otherwise \end{cases}$

Descriptive analysis

	dwell_time	board	alight	board2	alight2	late	on_time	off_board_fare_col	weekday	am_peak	pm_peak	irregular_activity
count	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000
mean	19.159	1.350	1.199	6.788	5.616	0.087	0.184	0.196	0.868	0.228	0.261	0.063
std	24.978	2.228	2.044	34.255	34.490	0.283	0.388	0.397	0.338	0.420	0.439	0.243
min	2.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
25%	6.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
50%	10.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
75%	20.000	2.000	1.000	4.000	1.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000
max	180.000	47.000	43.000	2209.000	1849.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Correlation between variables

	dwell_time	board	alight	board2	alight2	late	on_time	off_board_fare_col	weekday	am_peak	pm_peak	irregular_activity
dwell_time	1.000	0.609	0.076	0.435	0.085	0.123	0.169	-0.010	-0.010	-0.022	-0.014	0.829
board	0.609	1.000	-0.095	0.822	0.011	0.117	0.151	0.009	0.007	-0.012	-0.014	0.493
alight	0.076	-0.095	1.000	0.007	0.812	0.092	0.024	0.016	0.008	-0.024	0.030	0.022
board2	0.435	0.822	0.007	1.000	0.064	0.096	0.097	-0.005	0.007	-0.021	0.010	0.381
alight2	0.085	0.011	0.812	0.064	1.000	0.064	0.018	-0.006	0.008	0.006	0.004	0.043
late	0.123	0.117	0.092	0.096	0.064	1.000	-0.147	0.056	-0.027	-0.054	0.063	0.094
on_time	0.169	0.151	0.024	0.097	0.018	-0.147	1.000	0.332	-0.031	0.022	-0.039	0.122
off_board_fare_col	-0.010	0.009	0.016	-0.005	-0.006	0.056	0.332	1.000	-0.075	-0.023	0.003	-0.022
weekday	-0.010	0.007	0.008	0.007	0.008	-0.027	-0.031	-0.075	1.000	0.047	0.000	-0.008
am_peak	-0.022	-0.012	-0.024	-0.021	0.006	-0.054	0.022	-0.023	0.047	1.000	-0.323	-0.016
pm_peak	-0.014	-0.014	0.030	0.010	0.004	0.063	-0.039	0.003	0.000	-0.323	1.000	-0.004
irregular_activity	0.829	0.493	0.022	0.381	0.043	0.094	0.122	-0.022	-0.008	-0.016	-0.004	1.000

Linear Model Formulation

$$dwell\ time = \beta_0 + \beta_1 * \text{board} + \beta_2 * \text{alight} + \beta_3 * \text{board}^2 + \beta_4 * \text{alight}^2 + \beta_5 * \text{ontime} + \beta_6 * \text{early} + \beta_7 * \text{off\_board\_fare\_col} + \beta_8 * \text{weekday} + \beta_9 * \text{am\_peak} + \beta_{10} * \text{pm\_peak} + \beta_{11} * \text{irregular\_activity} + \varepsilon$$

OLS Regression Results

Dep. Variable:	dwell_time	**R-squared:**	0.761
Model:	OLS	**Adj. R-squared:**	0.761
Method:	Least Squares	**F-statistic:**	2.130e+04
Date:	Mon, 01 May 2017	**Prob (F-statistic):**	0.00
Time:	15:07:49	**Log-Likelihood:**	-2.8812e+05
No. Observations:	73488	**AIC:**	5.763e+05
Df Residuals:	73476	**BIC:**	5.764e+05
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
Intercept	7.4349	0.146	50.960	0.000	7.149 7.721
board	4.8286	0.039	123.655	0.000	4.752 4.905
alight	1.8584	0.039	47.741	0.000	1.782 1.935
board2	-0.1303	0.002	-55.554	0.000	-0.135 -0.126
alight2	-0.0449	0.002	-19.763	0.000	-0.049 -0.040
on_time	2.9247	0.127	23.060	0.000	2.676 3.173
early	3.2976	0.228	14.486	0.000	2.851 3.744
off_board_fare_col	-1.3795	0.123	-11.247	0.000	-1.620 -1.139
pm_peak	-0.4786	0.109	-4.409	0.000	-0.691 -0.266
am_peak	-0.6083	0.114	-5.352	0.000	-0.831 -0.385
weekday	-0.4916	0.134	-3.679	0.000	-0.753 -0.230
irregular_activity	69.5193	0.215	323.800	0.000	69.099 69.940

Comparison with constrained model

OLS Regression Results

Dep. Variable:	dwell_time	**R-squared:**	-0.000
Model:	OLS	**Adj. R-squared:**	-0.000
Method:	Least Squares	**F-statistic:**	-inf
Date:	Sun, 30 Apr 2017	**Prob (F-statistic):**	nan
Time:	17:12:50	**Log-Likelihood:**	-2.3196e+05
No. Observations:	50000	**AIC:**	4.639e+05
Df Residuals:	49999	**BIC:**	4.639e+05
Df Model:	0		
Covariance Type:	nonrobust		

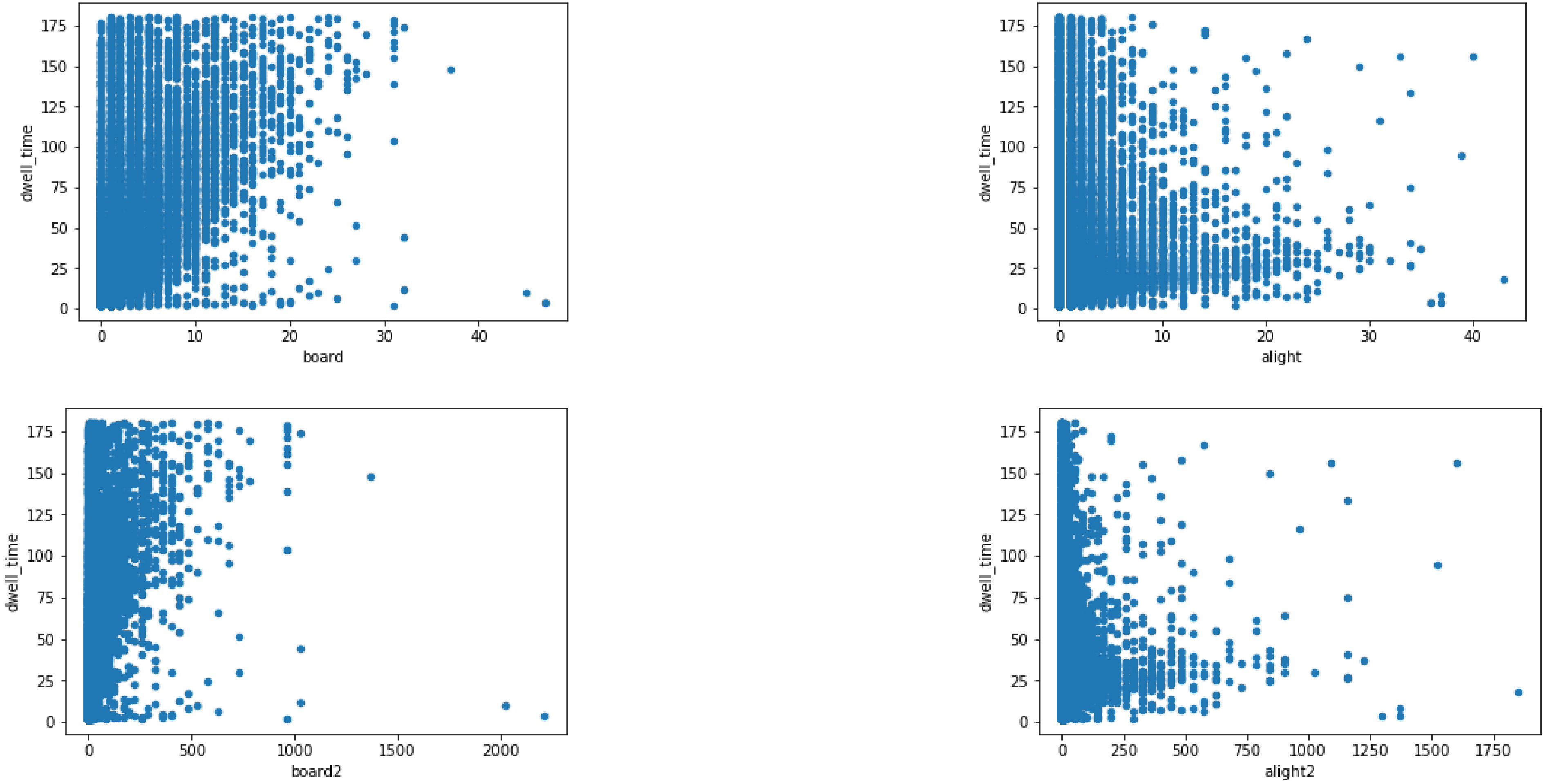
	coef	std err	t	P>|t|	[95.0% Conf. Int.]
constant	19.0776	0.112	170.406	0.000	18.858 19.297

$$dwell\ time_{constrained} = \beta'_0$$

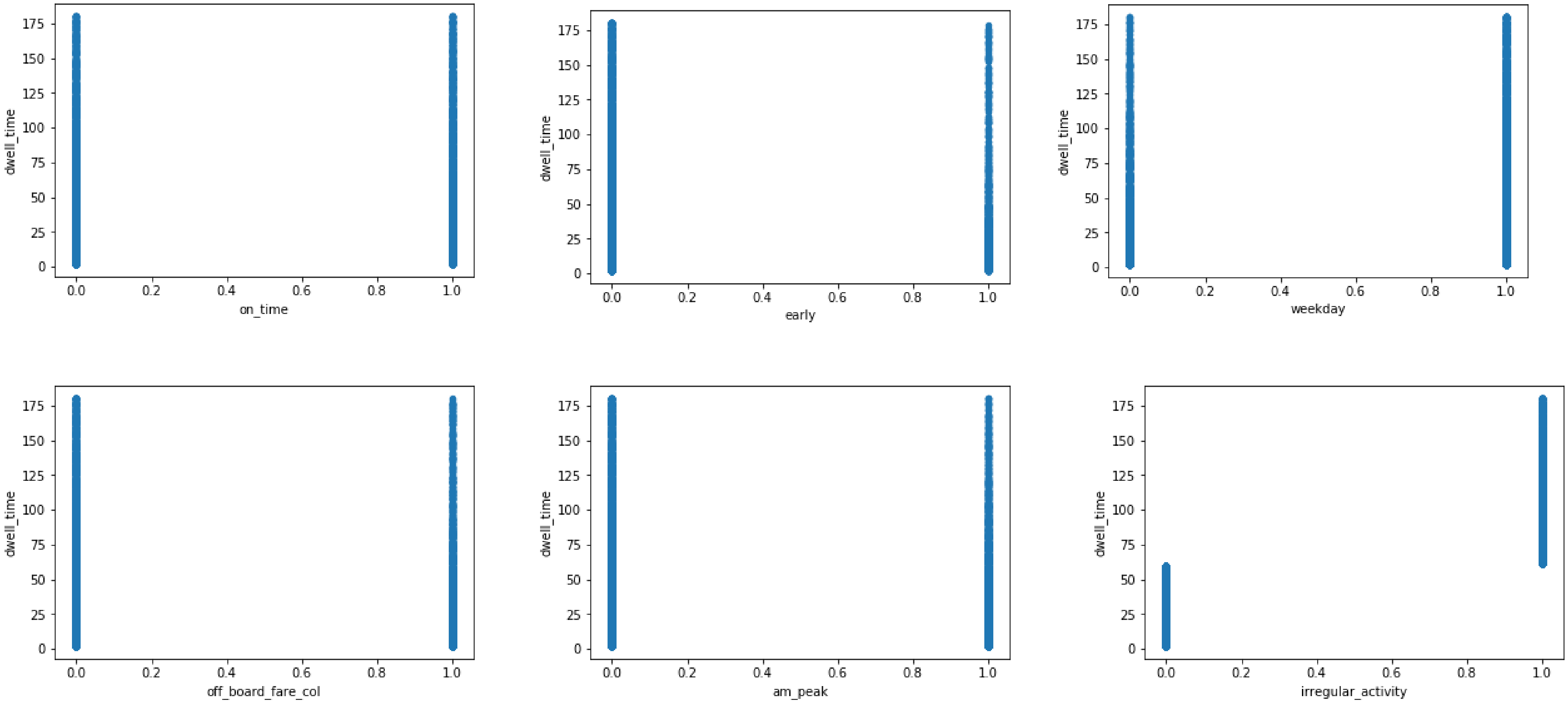
F-statistic = 3402.4124777047923

P_value = $1.1102230246251565*10^{-16}$

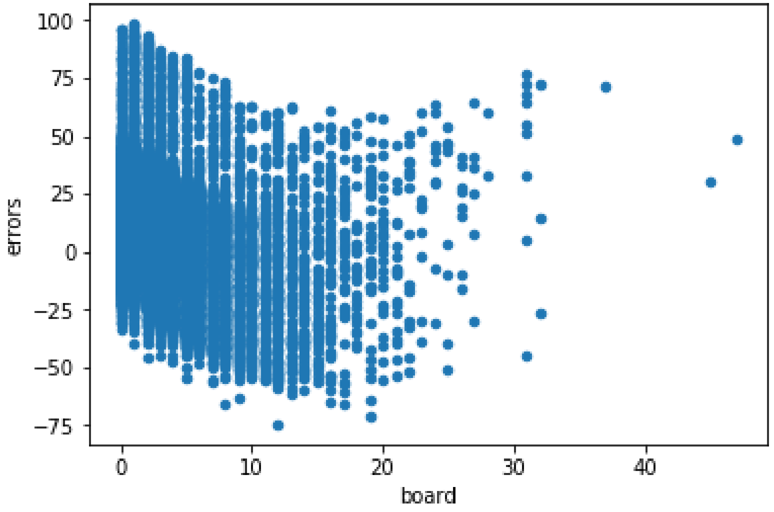
Model Diagnostic- Correct Specification

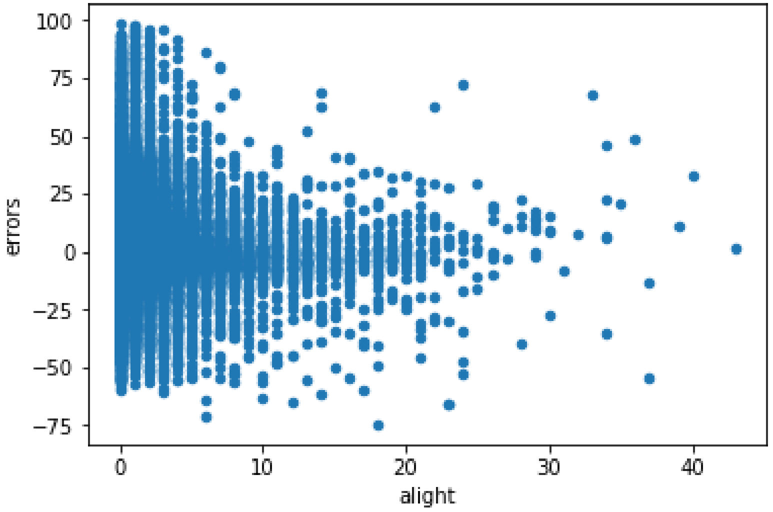


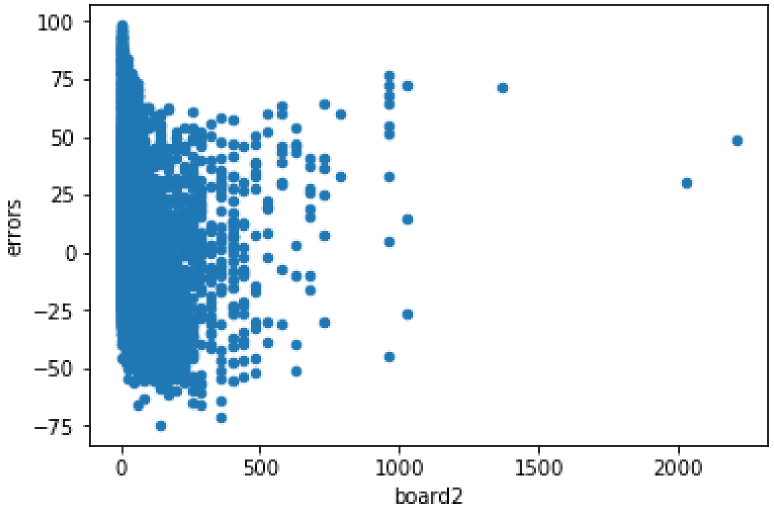
Model Diagnostic- Correct Specification

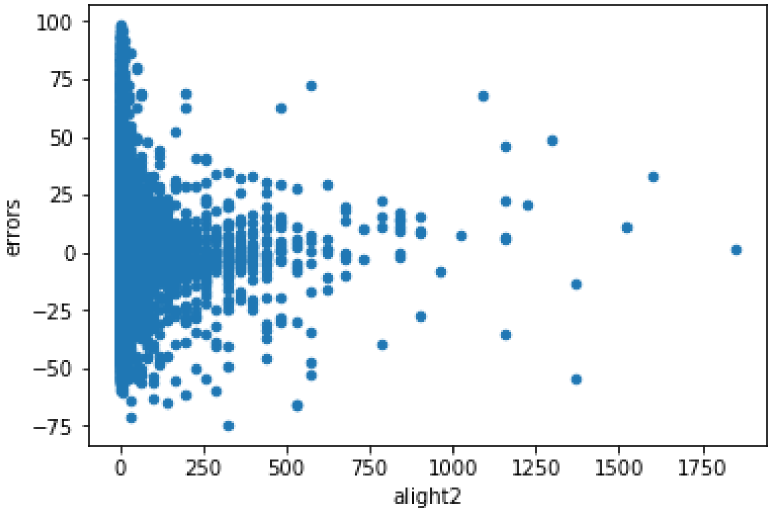


Model Diagnostic- Exoginity

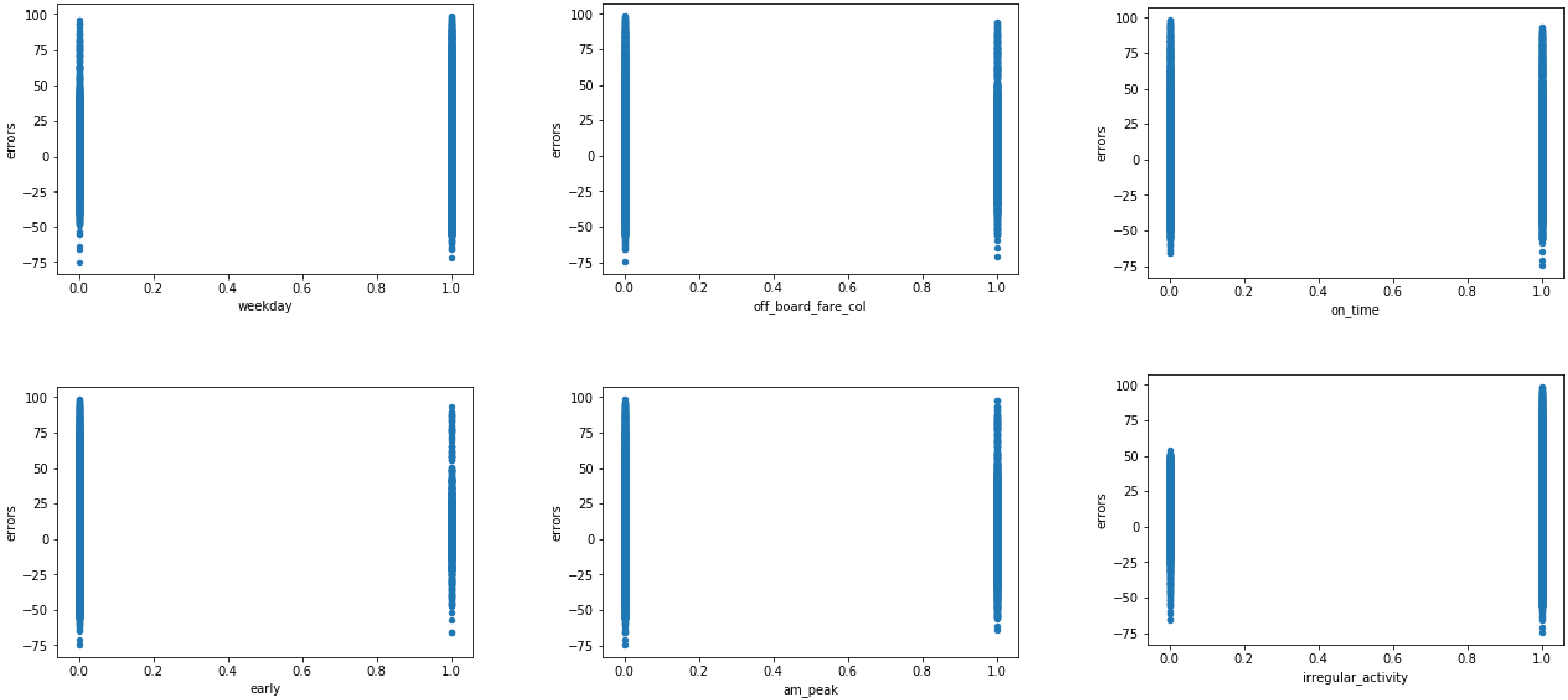






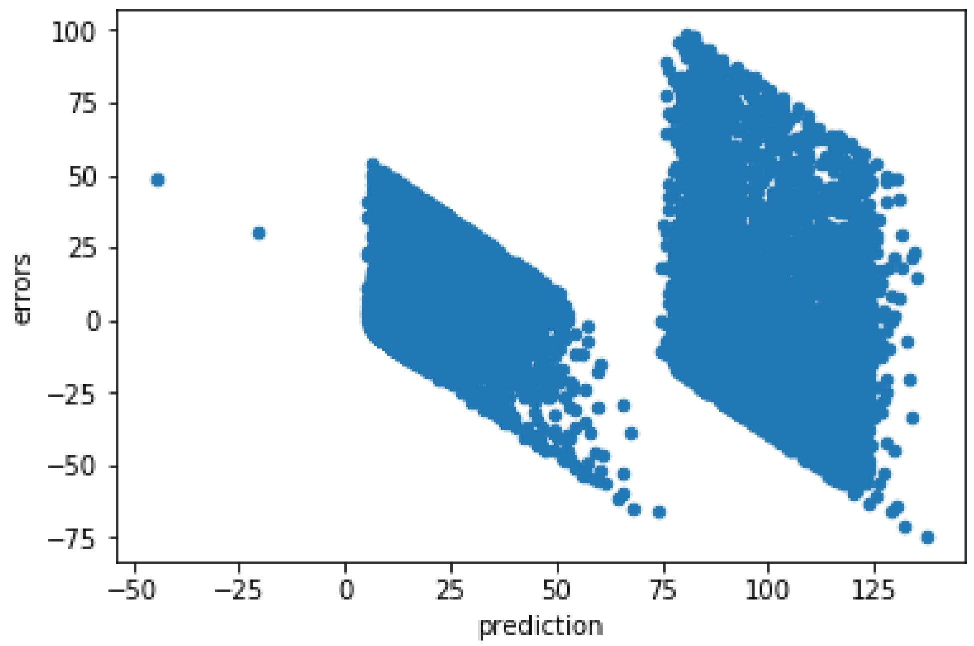


Model Diagnostic- Exoginity

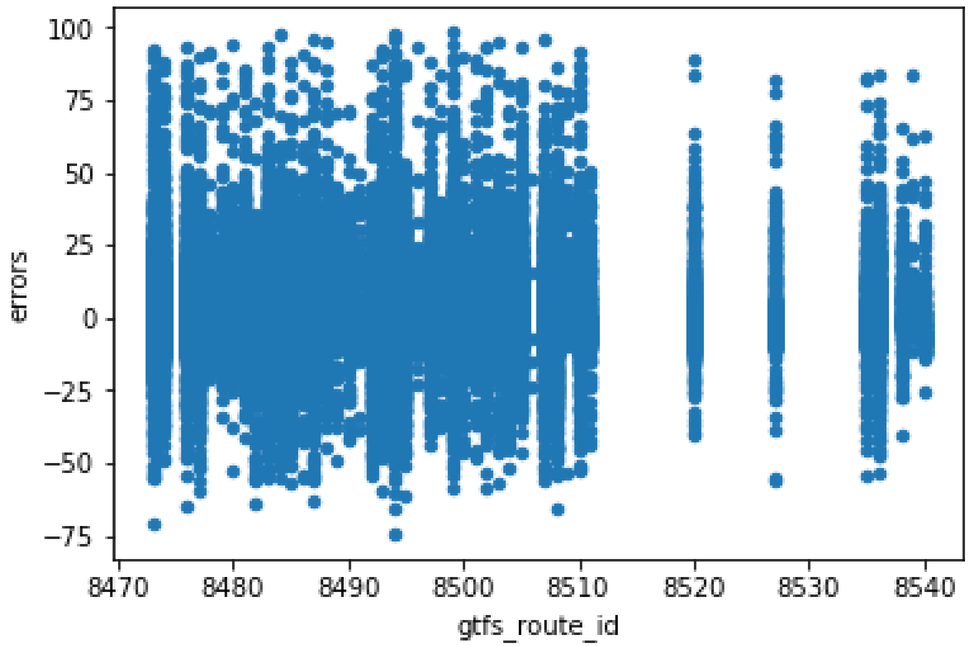


Model Diagnostic

Homoscedasticity of disturbances

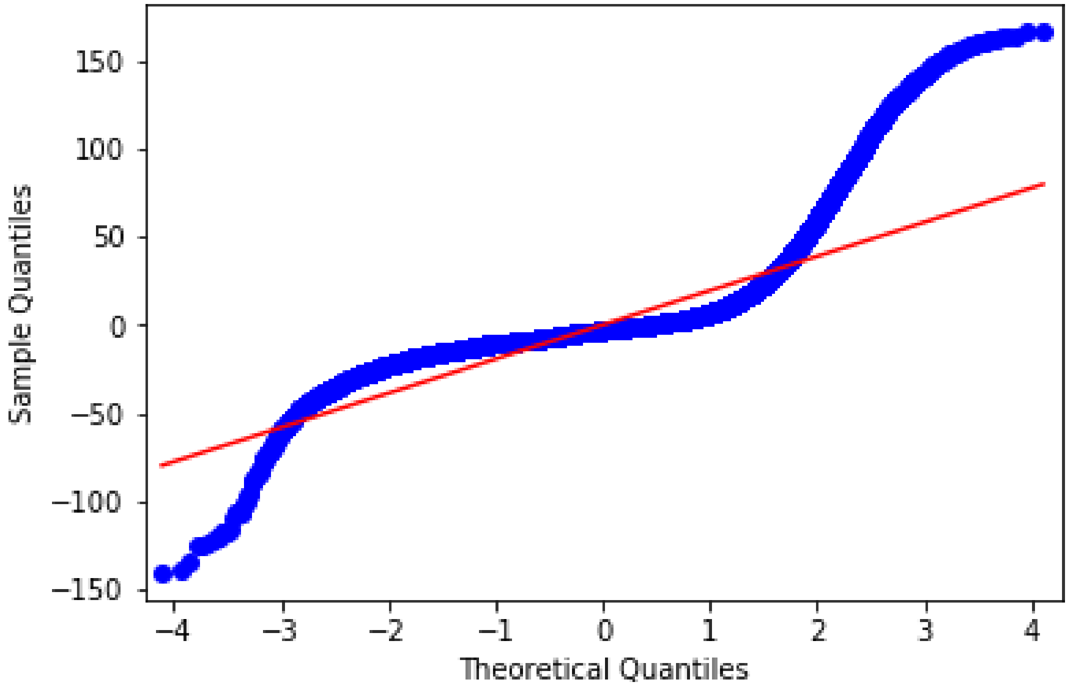


No Autocorrelation of disturbances



Model Diagnostic

Normality of disturbances



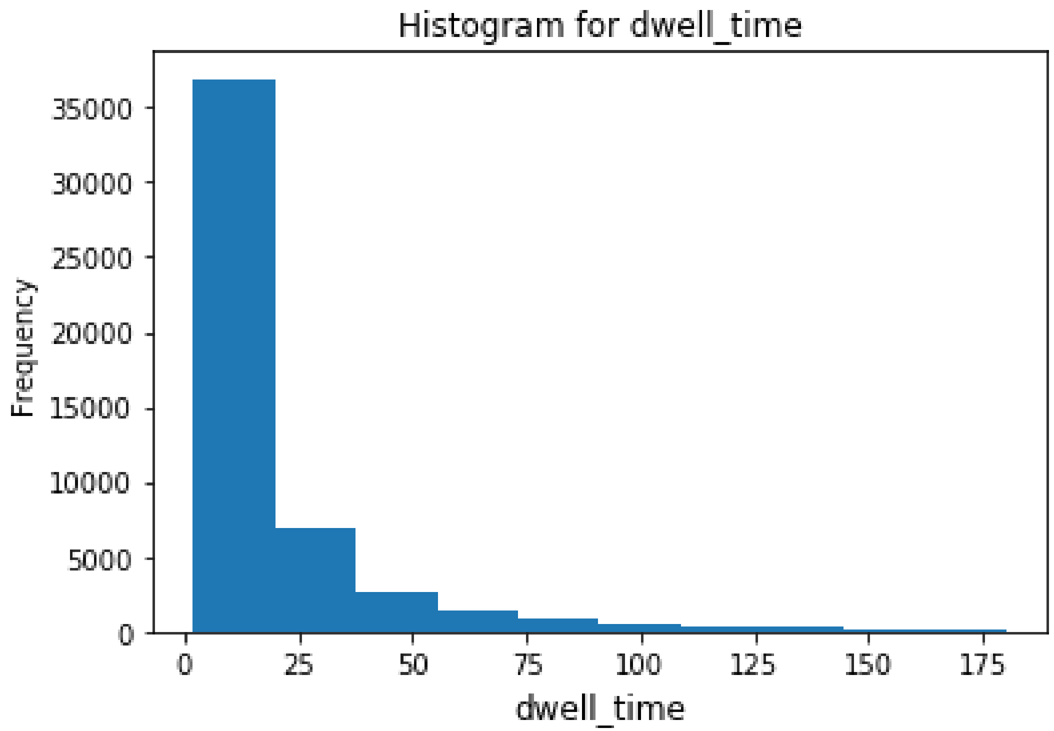
Logarithmic transformation of data

Cox-Box Transformation (δ = 0.001)

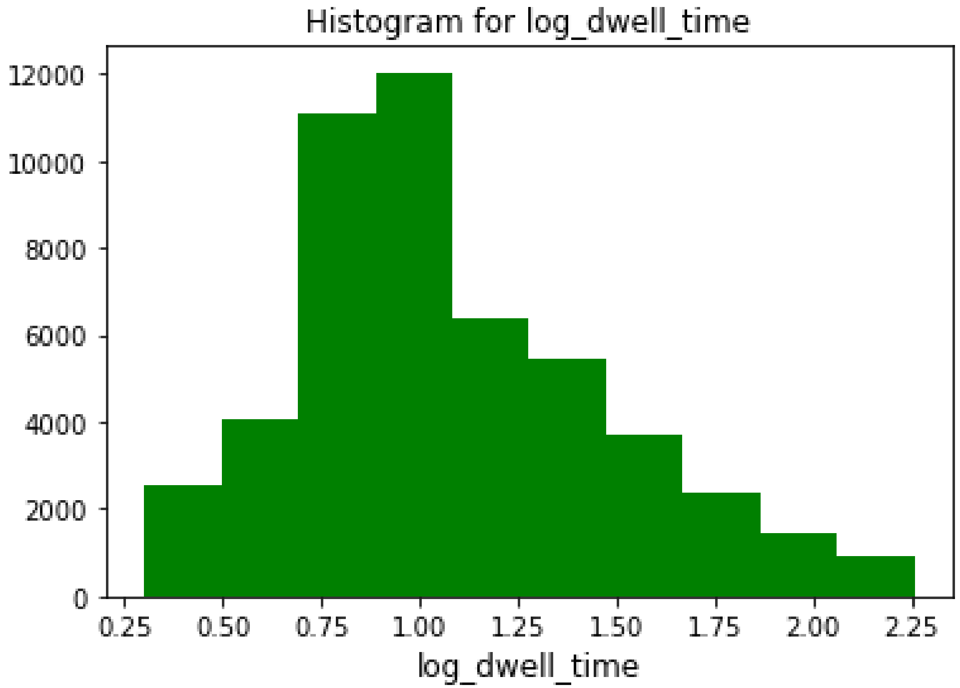
- Dependent variable:
  - dwell_time – is transformed to log_dwell_time
- Independent Variables:
  - board – is transformed to log_board
  - alight– is transformed to log_alight
  - board2 – is transformed to log_board2
  - alight2 – is transformed to log_alight2
  - All dummy variables are kept untransformed

Dwell time

Linear Regression



Log-transformed Linear Regression



Descriptive analysis – Logarithmic Model

	log_board	log_alight	log_board2	log_alight2	log_dwell_time	late	on_time	off_board_fare_col	weekday	am_peak	pm_peak	irregular_activity
count	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000	73488.000
mean	-1.513	-1.553	-1.376	-1.438	1.072	0.087	0.184	0.196	0.868	0.228	0.261	0.063
std	2.096	2.079	2.245	2.206	0.395	0.283	0.388	0.397	0.338	0.420	0.439	0.243
min	-4.000	-4.000	-4.000	-4.000	0.301	0.000	0.000	0.000	0.000	0.000	0.000	0.000
25%	-4.000	-4.000	-4.000	-4.000	0.778	0.000	0.000	0.000	1.000	0.000	0.000	0.000
50%	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
75%	0.301	0.000	0.602	0.000	1.301	0.000	0.000	0.000	1.000	0.000	1.000	0.000
max	1.672	1.633	3.344	3.267	2.255	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Logarithmic Model Formulation

$$\log\_dwell\_time = \beta_0 + \beta_1 * \log\_board + \beta_2 * \log\_alight + \beta_4 * \log\_board^2 + \beta_5 * \log\_alight^2 + \beta_7 * ontime + \beta_8 * early + \beta_9 * off\_board\_fare\_col + \beta_{10} * weekday + \beta_{11} * am\_peak + \beta_{12} * pm\_peak + \beta_{13} * irregular\_activity + \varepsilon$$

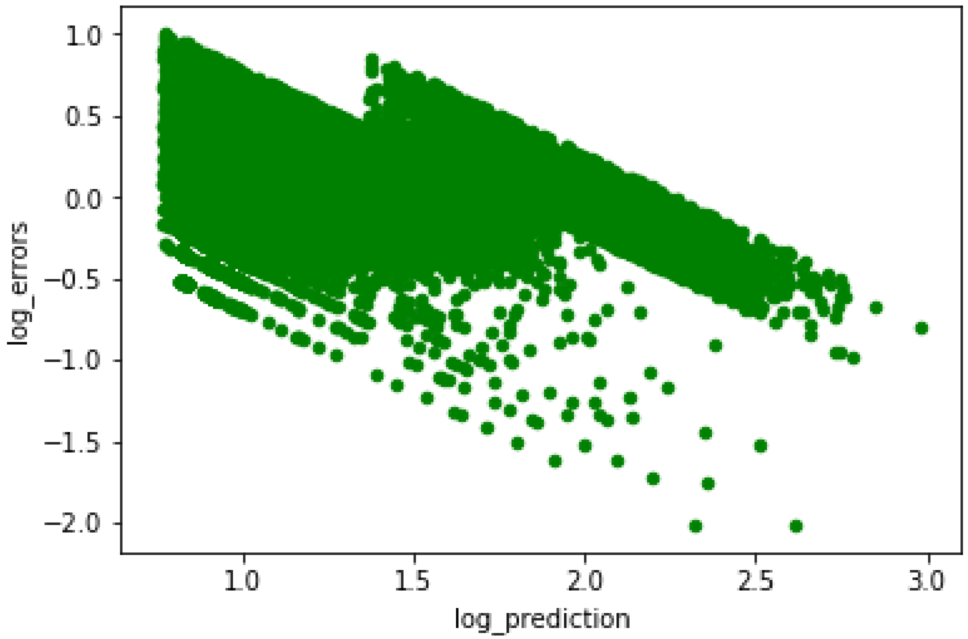
OLS Regression Results

Dep. Variable:	log_dwell_time	**R-squared:**	0.584
Model:	OLS	**Adj. R-squared:**	0.584
Method:	Least Squares	**F-statistic:**	1.030e+04
Date:	Mon, 01 May 2017	**Prob (F-statistic):**	0.00
Time:	15:46:40	**Log-Likelihood:**	-3830.2
No. Observations:	73488	**AIC:**	7682.
Df Residuals:	73477	**BIC:**	7784.
Df Model:	10		
Covariance Type:	nonrobust		

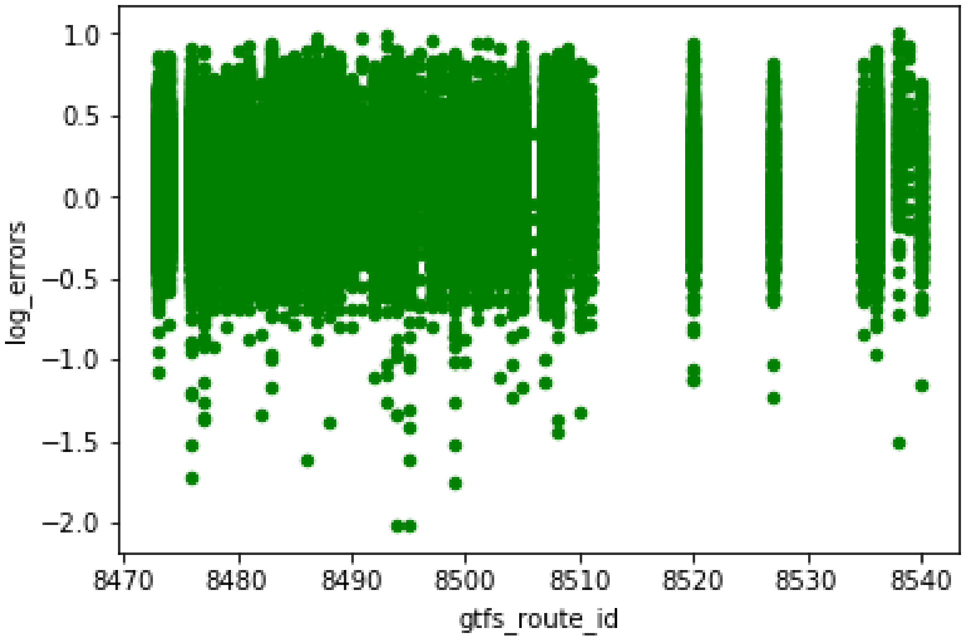
	coef	std err	t	P>|t|	[95.0% Conf. Int.]
Intercept	0.9747	0.003	342.823	0.000	0.969 0.980
log_board	-0.5595	0.005	-105.484	0.000	-0.570 -0.549
log_alight	-0.3543	0.005	-67.818	0.000	-0.365 -0.344
log_board2	0.5929	0.005	120.229	0.000	0.583 0.603
log_alight2	0.3674	0.005	75.900	0.000	0.358 0.377
on_time	0.0690	0.003	26.048	0.000	0.064 0.074
early	0.0885	0.005	18.598	0.000	0.079 0.098
off_board_fare_col	-0.0083	0.003	-3.254	0.001	-0.013 -0.003
am_peak	-0.0167	0.002	-7.041	0.000	-0.021 -0.012
pm_peak	-0.0171	0.002	-7.519	0.000	-0.022 -0.013
irregular_activity	0.6009	0.004	137.476	0.000	0.592 0.609

Model Diagnostic

Homoscedasticity of disturbances

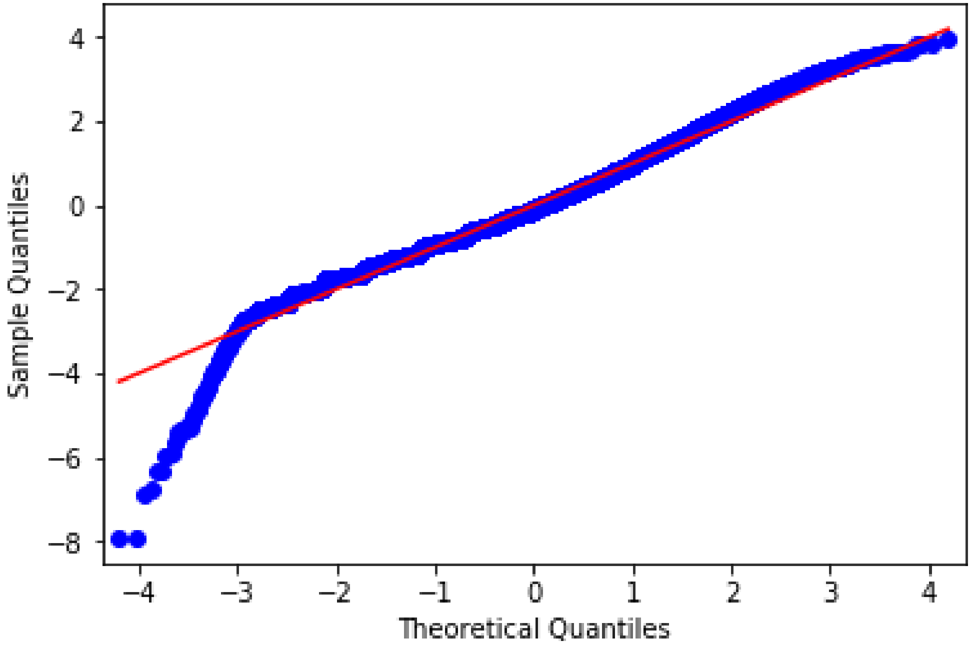


No Autocorrelation of disturbances



Model Diagnostic

Normality of disturbances



Model Comparisons

	Constrained Model	Linear Model	Logarithmic Linear Model
R^2	0.000	0.761	0.584
Adjusted R^2	0.000	0.761	0.584

Reducing dwell time and emission by introducing off-board fare collection

- Total dwell time in the month October and November of 2016 is = 67976537 seconds
- Average daily dwell time is = 1114369 seconds or 309.54 hours
- Dwell time saved is = -1.3795
- Total idling saved in a year = 3673 hours
- Total fuel saved in a year = 6740.116 gallons
- Total equivalent VMT in a year = 101101.745 miles
- Total Green House Gas saved in a year = 223.856 Metric Tons of CO_2 equivalent

Conclusion and Discussion

- The linear model performs better than logarithmic model
- A small amount of dwell time can be saved by using off-board fare collection in Hartford Buses
- If all buses in Hartford had off-board fare collection system – a total of 223.856 Metric tons of CO_2 equivalent of Green House emission can be reduced

- Further study:
  - Modelling CTFastrak separately
  - Using fare collection method information
  - Exclude zero passenger activity data
  - Use different threshold for irregular activity

References:

- *Data Source: t-HUB, http://thub-gis.engr.uconn.edu/realtime_apc_data/*
- *Guenthner, R. P., and K. C. Sinha, Modeling Bus Delays Due to Passenger Boardings and Alightings, Transportation Research Record 915, TRB, National Research Council, Washington, D.C., 1983, pp. 7–13.*
- *Milkovits MN (2008), Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Dat, Transportation Res Record 2072: 125-130*
- Daniel B. Shockley, Julia Salinas, and Brian D. Taylor, Making Headways: An Analysis of Smart Cards and Bus Dwell Time in Los Angeles,
- *Rajbhandari R, Chien S, Daniel J (2003) Estimation Dwell Times with Automatic Passenger Counter Information. Transportation Res Record: J Transportation Research Board 1841.*
- *Alejandro Tirachini, Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers, Transportmetrica A: Transport Science, pp. 30-49, 2004*
- *Dueker, K.J., Kimpel, T.J., and Strathman, J.G., 2004. Determinants of bus dwell time. Journal of Public Transportation, 7 (1), 21–39.*
- *Stephen A Arhin, Errol C Noel, Melissa Anderson, Lakeasha Williams, Asteway Ribbiso and Regis Stinson, Predicting Dwell Time by Bus Stop Type and Time of the Day, Journal of Civil & Environmental Engineering, 2015*
- *S.M. Ashrafur Rahman, H.H. Masjuki, M.A. Kalam, M.J. Abedin, A. Sanjid, H. Sajjad, Impact of idling on fuel consumption and exhaust emissionsand available idle-reduction technologies for diesel vehicles – A review, Journal of Energy Conversion and Management, 2013*
- *Rajat Rajbhandari, Steven I. Chien, and Janice R. Daniel, Estimation of Bus Dwell Times with Automatic Passenger Counter Information, Transportation Research Record 1841, Paper No. 03-2675, pp. 120-127*
- *Han Lim,Study of Exhaust emission from idling heavy duty vehicles and Commercially available Idle Reducing Devices, Technical report Office of Transportation and Air Quality, EPA, 2002*
- *EPA – Simplified GHG Calculator, EPA Center for Corporate Climate Leadership, 2017*
- *APTA – Transit GHG Calculator, Statistics of American Public Transportation Association, 2015*

Thank you!

Questions?

Interpretation:

If independent variable (X) is not transformed:

1 unit change in X_1 is equal to $(exp^{\beta_1} - 1) * 100$ percent change in Y

If independent variable (X) is not transformed:

1 % change in X_2 is equal to $100 * (1 - 1.01^{\beta_2})$ percent change in Y