

Can an Aggregate Transport Mode Choice Model for London Buses, based on Ward Attributes, be used for Policy Making and Intervention?

Fred Shone

UCL - Centre for Advanced Spatial Analysis - Spatial Data Science and Visualization

Quantitative Methods 2017 Final Assessment Presentation – 8th December 2017

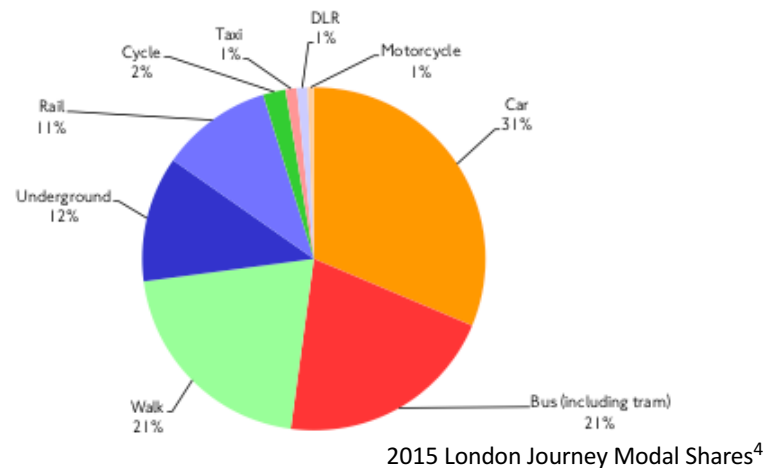
Part 1: Transport Mode Choice Models

- Transport mode choice models are not new
 - Recent analysis is predominantly for cycling and walking¹
 - Models are increasingly complex²
 - Often incorporate segmentation³ ie clustering and then linear or logistic regression⁴
 - Primarily based on individuals, ie dis-aggregate

1. Estimation of the determinants of bicycle mode share for the journey to work using census data. Parkin et al.
2. Spatial Autoregressive Error Components in Travel Flow Models: An Application to Aggregate Mode Choice. Denis Bolduc
3. Using Segmentation Approaches for Better Prediction and Understanding from Consumer Mode Choice Models. Imran S. Currim
4. Sirakaya, E., Woodside, A. G., 2005. Building and testing theories of decision making by travelers. *Tourism Management*, 26 815-832.

Part 1: Why Bus Mode Share?

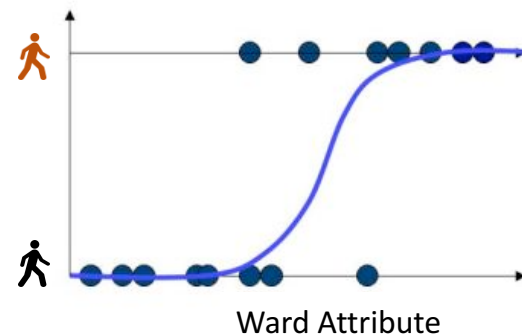
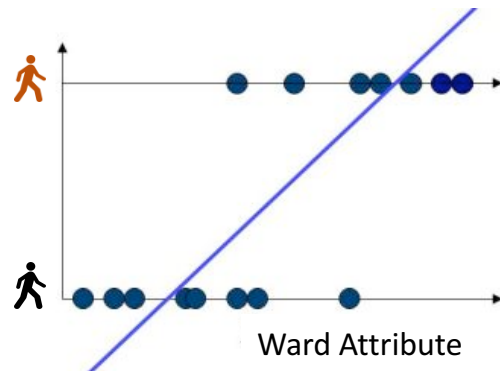
- The TFL bus network carries ~2.3bn passengers per year¹
- The current aim is to increase Active and Public Transport by 16% by 2041²
- TFL spends ~3bn pa. operating, maintaining and investing in buses³



1. Transport for London Website [<https://www.tfl.gov.uk/modes/buses/improving-buses>]
2. Mayors Transport Strategy [<https://consultations.tfl.gov.uk/policy/mayors-transport-strategy>]
3. Transport for London Budget [[transport-for-london-budget-2017-18.pdf](https://www.tfl.gov.uk/transport-for-london-budget-2017-18.pdf)]
4. TFL, Travel in London Report, Report 9

Part 1: Modelling Bus Mode Share

- Dependent Variable = London Bus Mode Share by London Ward
 - Data from the London Travel Demand Survey (LTDS)
 - Combined over ~three years to get adequate sample size
 - The share of each ward population choosing to travel by bus
- Non Linear! Use Logistic Regression, eg Logit

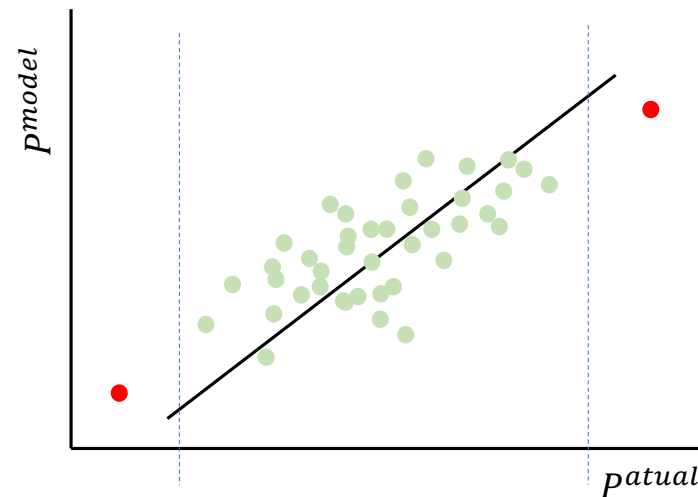


Part 1: Aggregate Mode Share

- In aggregate can a categorical variable reasonably be approximated as linear?
- Large sample allows Normal Approximate to Binomial¹



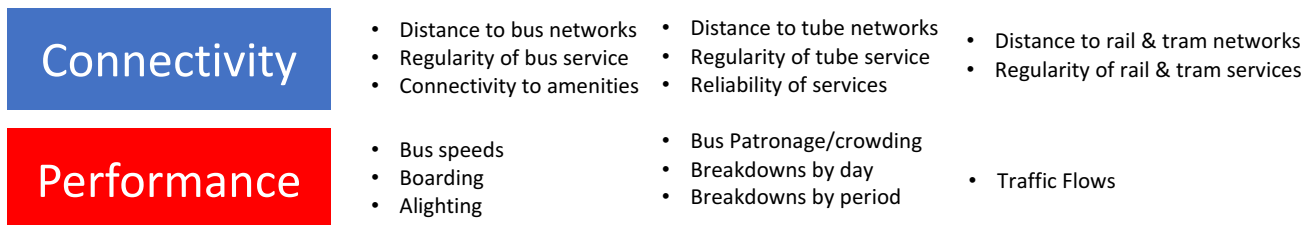
- Sensible results due to avoiding extremes



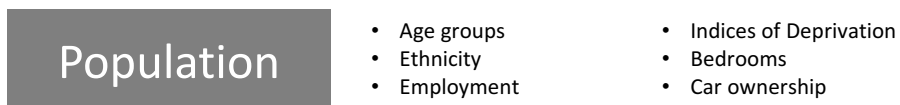
1. Penn State College Website. <https://onlinecourses.science.psu.edu/stat414/node/179>

Part 2: Explanatory Variables

- Explanatory variables (predictors) are aggregated to London Ward level
- Aggregation allows for use of TFL data to build model

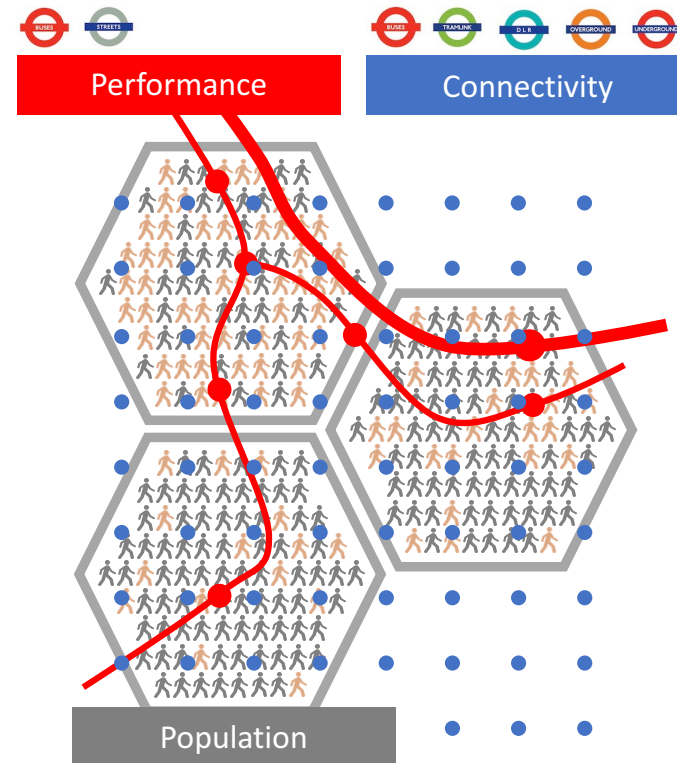


- Loss of individuals attributes compared to dis-aggregate model



Part 2: Methodology Discussion

- Challenges
 - Data cleaning and manipulation
 - Avoid over-fitting
 - Deal with correlated predictors
 - Capture complex interactions
 - Keep it interpretable
- Solutions
 - Algorithmic (Ridge Regression or LASSO methodologies)
 - **vs** common sense
 - Segmentation using clustering
 - Try and keep it linear and interpretable



Part 3: Applications in Policy Making and Intervention (1)

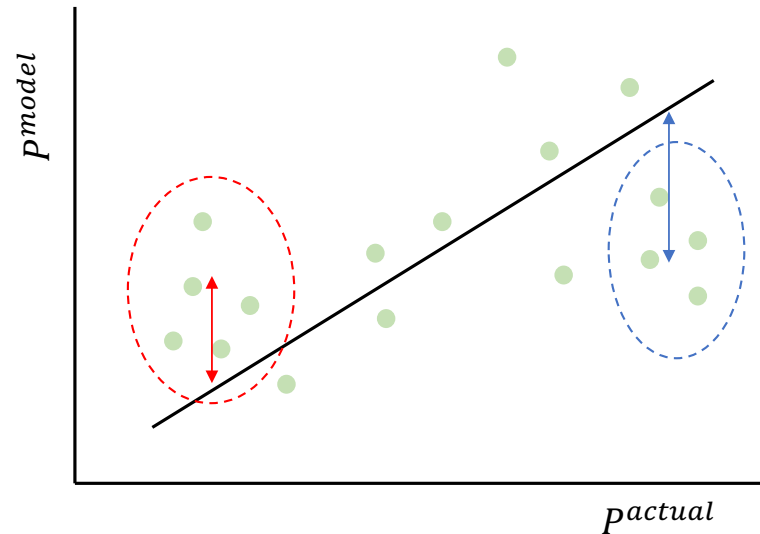
- Understanding predictors in the model (significance testing with hypotheses)

$$P^{ward} = \beta_0 + \beta_{0 \rightarrow n} [Population]^{ward} + \beta_{n \rightarrow m} [Connectivity]^{ward} + \beta_{m \rightarrow p} [Performance]^{ward}$$

- Allows prediction of future changes
- Suggests levers for change

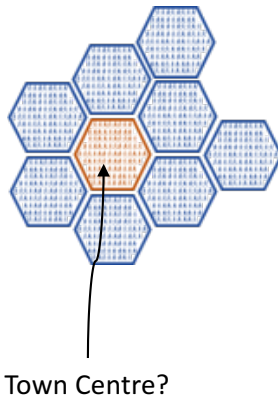
Part 3: Applications in Policy Making and Intervention (2)

- Identifying outliers
 - More data?
 - Improve the model?
 - Identify patterns

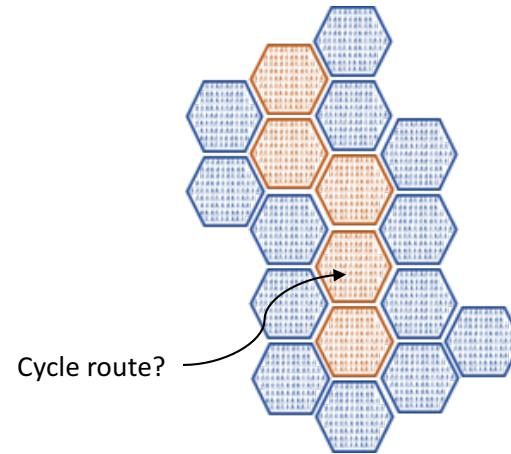


Part 3: Some Predictions for Policy Making and Intervention

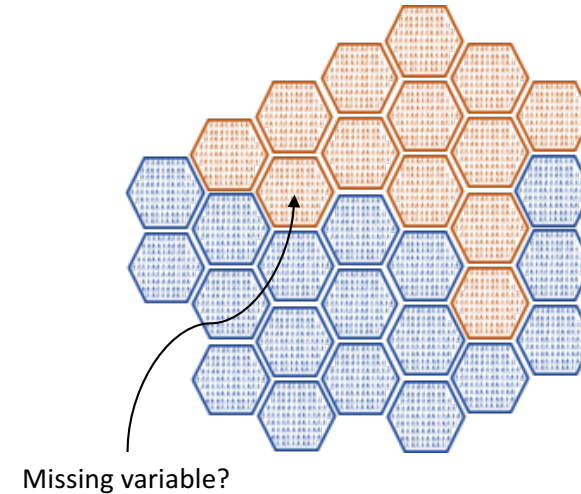
Intervention identification A:



Intervention identification B:



Policy identification:



Questions Please