Abhilash C. Singh
(44) 745 933 4955
abhilash.singh19@imperial.ac.uk
https://abhilashcsingh.github.io

Summary of thesis

Endogeneity and consideration set issues in residential location choice models

Urban land use and transport policies are essential to determine where people live, work and play, enabling sustainable and healthy forms of transportation between destinations. Compact cities are clearly needed to achieve GHG targets, as 70% of the world's population is expected to live in urban areas by 2050 and 33% of total urban GHG emissions in major cities are generated by transport. However, identifying the locations and levels of land use mix and densities along with adequate transport networks and technologies to support equitable, sustainable and healthy living remains challenging. Potential trade-offs between quality of life, pressures on land uses, equitable urban growth, transport efficiencies, and urban temperatures must be addressed transparently to both inform policies and engage society in embracing sustainable solutions. This thesis presents a range of approaches used to model residential location choices within the context of land use and transport models, dealing with multi-faceted problems arising with novel data sources. As part of this effort, we discuss the model scenarios with missing or filtered alternatives during choice-making process by decision makers (consideration-set issues), simultaneous determination of modal accessibilities and trip frequencies, and omitted information during modelling choices (endogeneity issues).

Firstly, residential location choice models are often burdened with missing information for modelling the choice set formation process. The information is missing either because one or more alternatives are unavailable to decision makers or because they considered (sampled) the alternatives endogenously, a consideration process that is inaccessible to the analyst. To the best of our knowledge, this is the first application of probabilistic decision trees to model consideration process in residential location choice using London-based data. We combine probabilistic decision trees with a traditional Mixed Logit choice model to estimate the final choice decision, capturing interdependence between alternatives within specific Transport Fare Zones, thereby accounting for non-compensatory consideration of choice alternatives followed by a compensatory choice decision. The results reveal factors influencing consideration set formation, highlighting the impact of variables such as household income, borough prices, population density, and crime density on location choice probability. Ethnic variation in the effect of crime density on location choice indicates racial and ethnic segregation in London housing. The research confirms the use of a non-compensatory process followed by a compensatory one in decision-making. However, the model has data limitations and challenges in merging machine learning and econometric approaches.

Secondly, it is widely acknowledged that residential location choice is driven not just by the built environment attributes but also by the level of access to activity opportunities from the location. We simultaneously analyse employment accessibility and household trip frequency across different transport modes in the City of Toronto. Accessibility measures for auto, transit, bicycle, and walk modes were created, along with household trip frequencies for each mode, categorized into low, medium, and high levels. The study uses a multivariate ordered probit model system, considering

both observed and unobserved factors' dependence on trip frequencies and accessibility. The results reveal strong associations and direct effects between employment accessibility and household trip frequency. Complementary effects are observed between auto accessibility and transit trips, while substitution effects exist between transit accessibility and auto trips. Policymakers can use these findings to promote transit development in transit-poor neighbourhoods and reduce vehicular congestion in Toronto. The limitations of limited cycling and walking infrastructure in Toronto impact trip utility for high population density neighbourhoods, leading to inequities in transport infrastructure distribution. Moreover, certain population groups are time-impaired, limiting their activity participation and raising equity concerns regarding transport infrastructure distribution in the city.

Thirdly, endogeneity in residential location choice models occurs when any explanatory variable is correlated with the error term. This can happen due to omitted attributes, measurement or specification errors, and simultaneous determination or self-selection. Addressing endogeneity is important because it leads to inconsistent parameter estimates. This analysis develops a control function (CF) based discrete choice model, addressing endogeneity in residential location choice due to housing prices and travel behaviour. Our hypothesis highlights the prominent negative impact of housing prices by accounting for endogeneity, while also addressing the endogeneity with travel behaviour. Utilizing London Travel Demand Survey (LTDS) data, we comprehensively analyse household location choices and their relation to travel behaviour. The CF-MxL model findings underscore housing prices' negative influence on location choices, particularly affecting lower-income groups, as compared to highly subdued effect size of the incorrect MNL model. Additionally, vehicle miles travelled (VMT) significantly impacts location choices, with households having higher VMT showing aversion to postcodes with a higher population density. Furthermore, we observe a positive association of non-white households with locations infested with a higher crime density. This research offers insights into the interplay between residential location and travel behaviour, isolating and revealing the impact of various postcode-attributes, socioeconomic and demographic characteristics.

This thesis comprises three distinct studies investigating residential location choice models and their interactions with external factors. The first study employs probabilistic decision trees to analyse choice set formation, revealing the impact of variables like income, borough prices, population density, and crime density on location decisions. The second study explores the relationship between employment accessibility, household trip frequency, and transport modes in Toronto, uncovering associations and disparities in infrastructure distribution. The third study addresses endogeneity using a control function-based model, highlighting housing price effects while accounting for travel behaviour endogeneity. These findings contribute insights into residential location dynamics, offering valuable implications for policy and understanding the multifaceted aspects of location choice.