Understanding the Role of Transportation Network Companies (TNC) in Addressing Transportation Demand: A Chicago Case Study

Dewan Ashraful Parvez[1]
Graduate Research Assistant

Dr. Sudipta Dey Tirtha[1]
Post-doctoral Scholar

Dr. Tanmoy Bhowmik[2]
Assistant Professor

Dr. Srinivas Peeta[3]
Frederick R. Dickerson Chair & Professor

Dr. Naveen Eluru[1]
Professor

1 UNIVERSITY OF CENTRAL FLORIDA

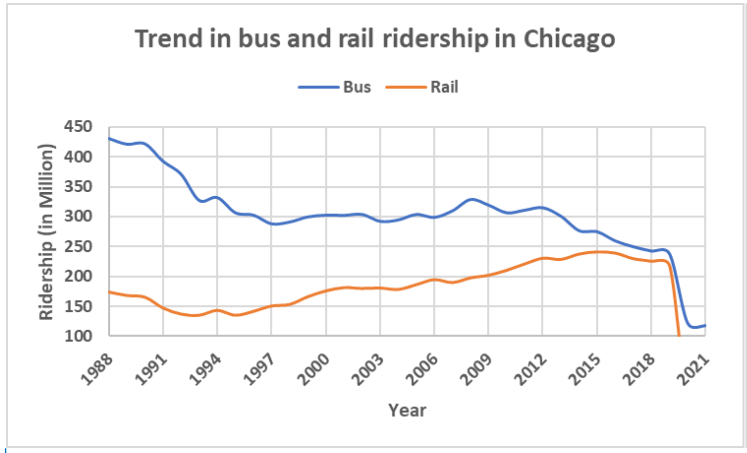
2 Portland State UNIVERSITY

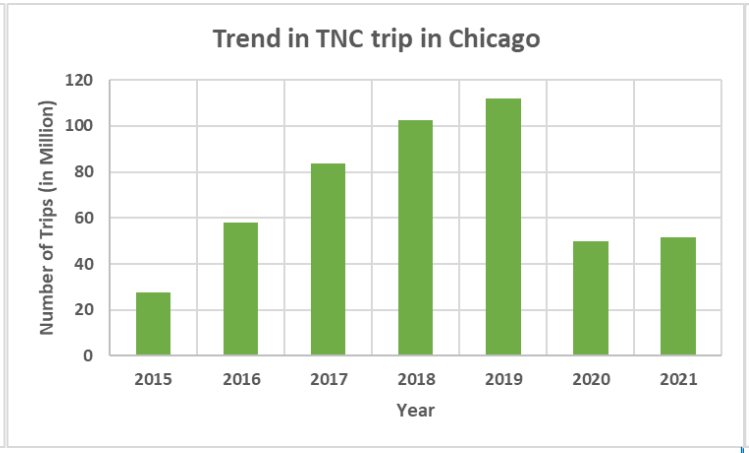
3 Georgia Tech

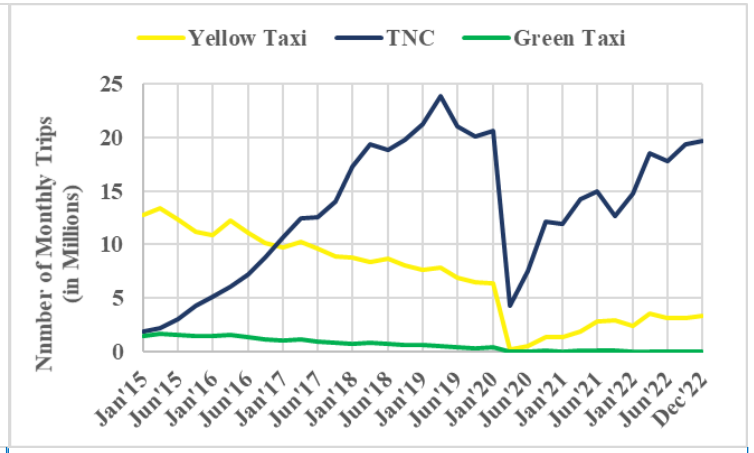
Outline

Introduction

- Transportation Network Companies (TNCs):
  - Uber: 10,000 cities
  - Lyft: 656 cities
  - Grab: 500 cities
- Ride share market is valued 85.8 billion in 2021 and expected to rise up to 185 billion by 2026
- In 2021, 118 million users used Uber at least once in a month
- Makeshift public transportation options







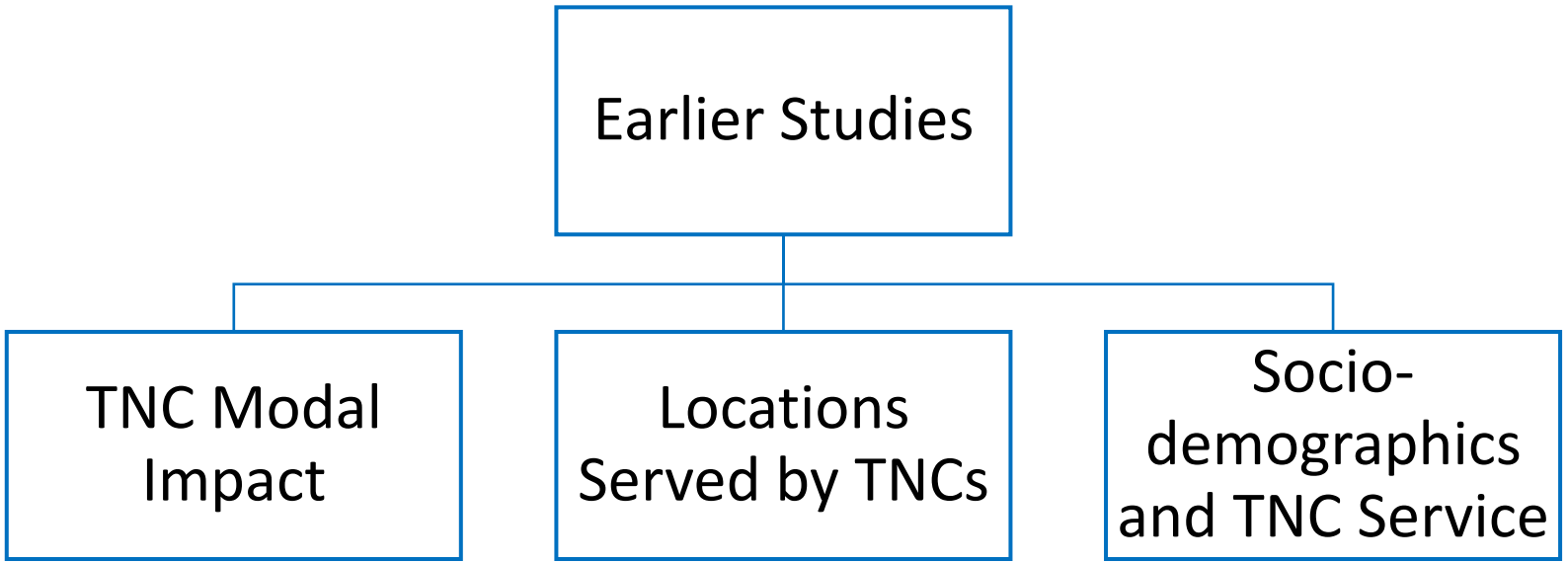
Chicago

New York City

Georgia Tech

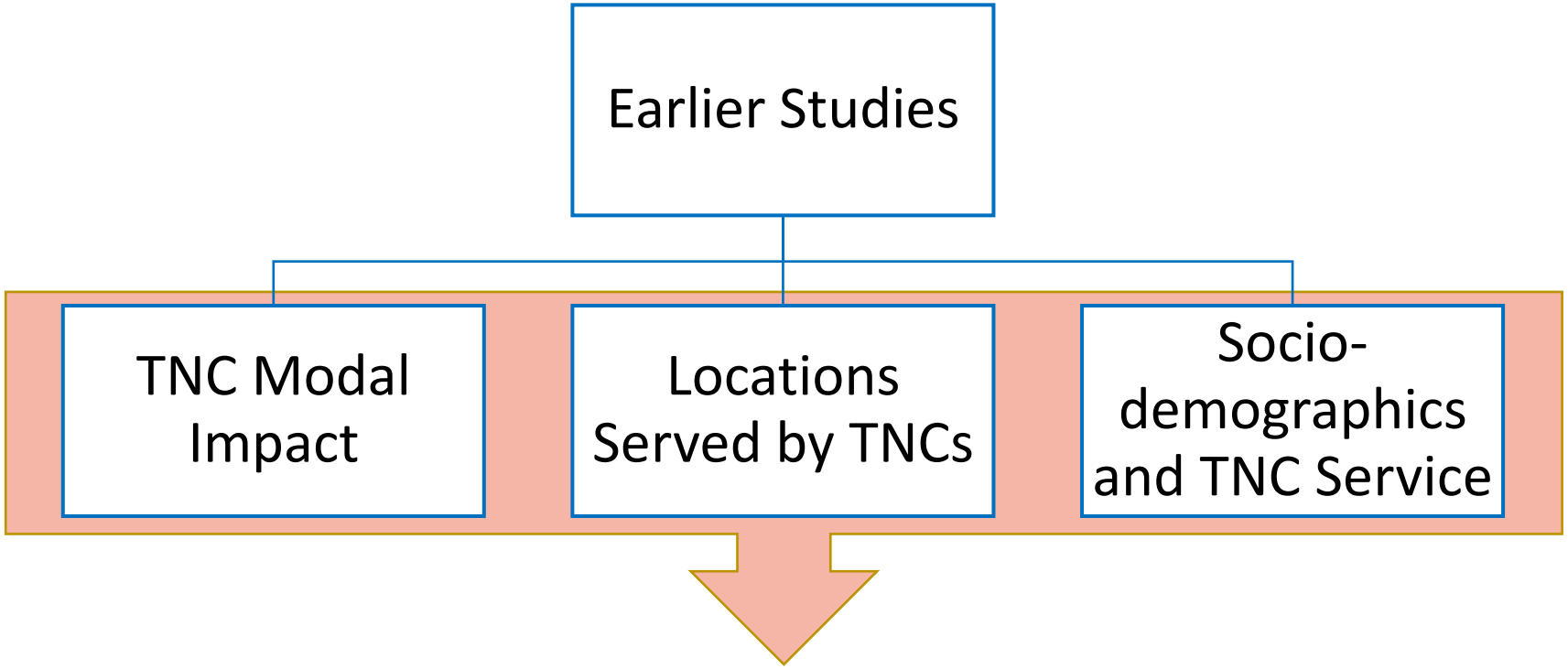
Background and Motivation

Findings from existing literature



Background and Motivation

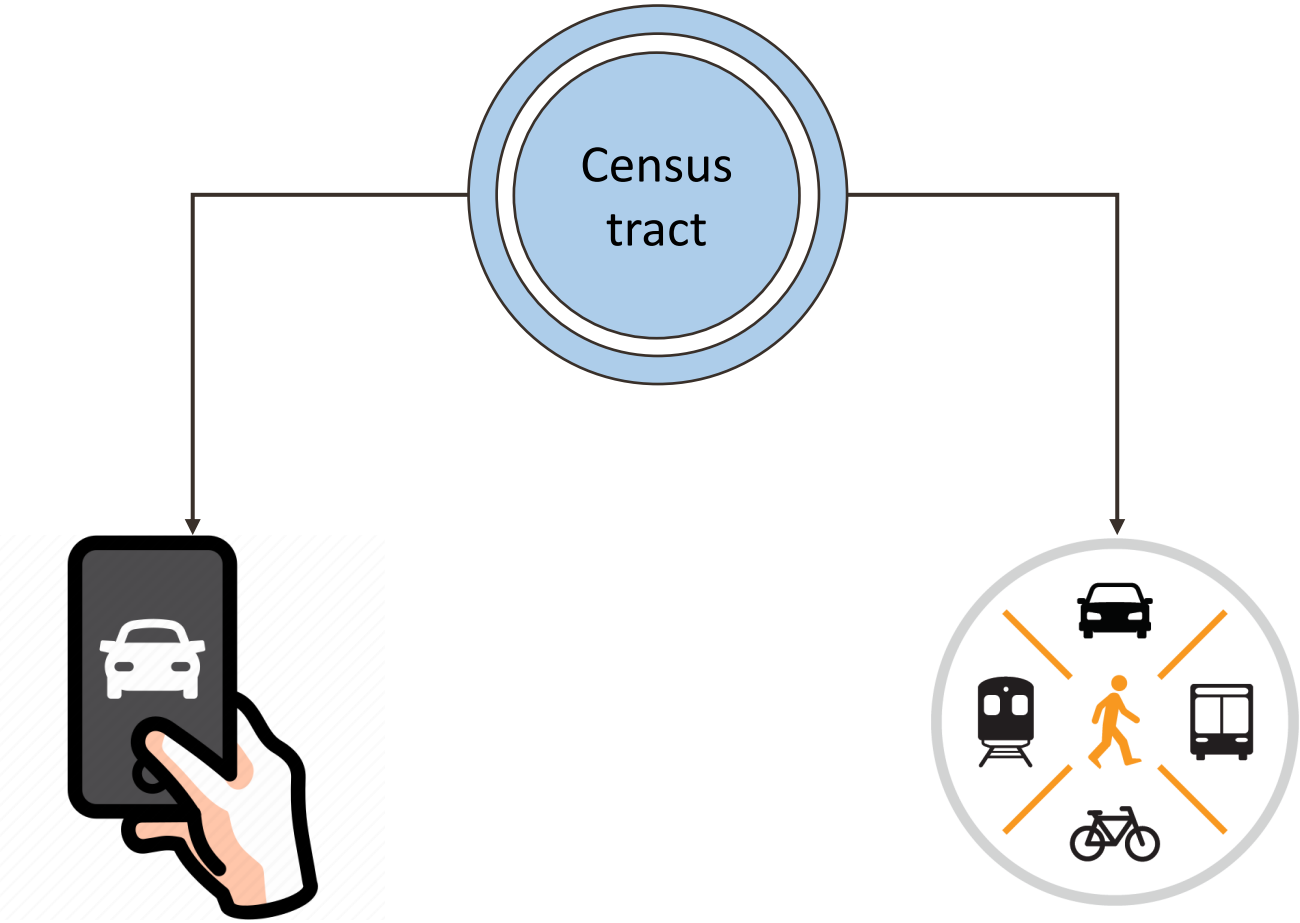
Findings from existing literature

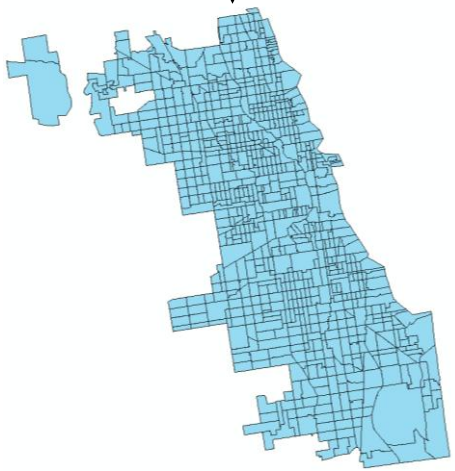


Current Study in Context

- Evaluating TNCs in the context of overall transportation demand.
- Compute a novel metric to identify the potential imbalance between overall transportation demand and TNC demand.
- Using a comprehensive set of explanatory variables.
- Framework developed can be applied for any urban region to identify spatial pockets underserved or overserved by TNCs.

Data Preparation & Dependent Variable





TNC Demand

Overall transportation demand

Data Preparation & Dependent Variable

TNC demand

1st Decile	1st Decile
2nd Decile	2nd Decile
3rd Decile	3rd Decile
4th Decile	4th Decile
5th Decile	5th Decile
6th Decile	6th Decile
7th Decile	7th Decile
8th Decile	8th Decile
9th Decile	9th Decile
10th Decile	10th Decile

Overall transportation demand

Data Preparation & Dependent Variable

TNC demand

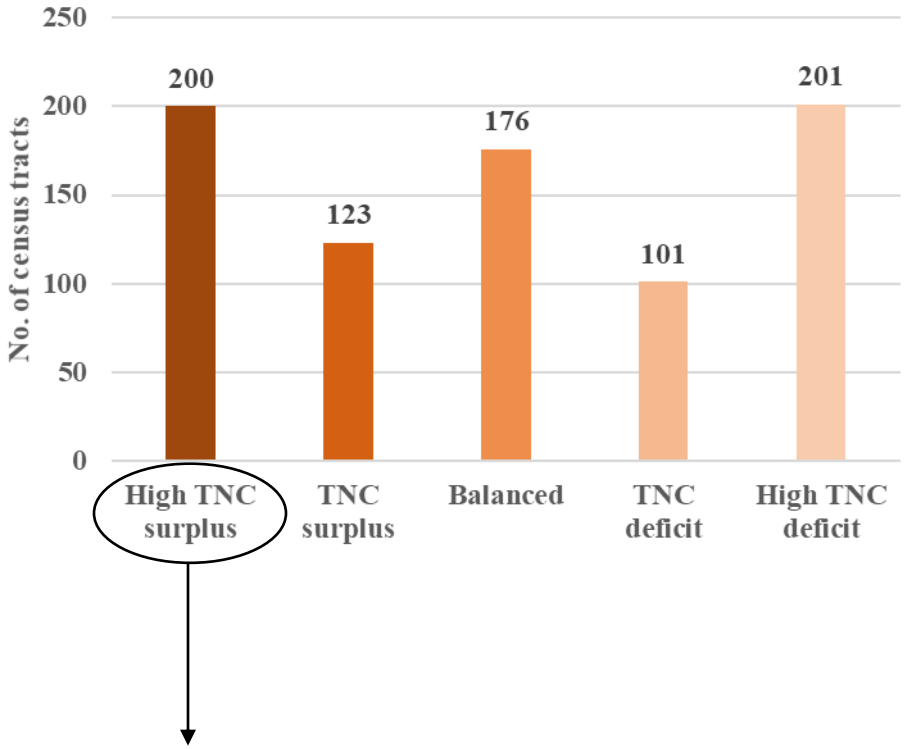
Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
Rank 8
Rank 9
Rank 10

Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
Rank 8
Rank 9
Rank 10

Overall transportation demand

Data Preparation & Dependent Variable

TNC demand

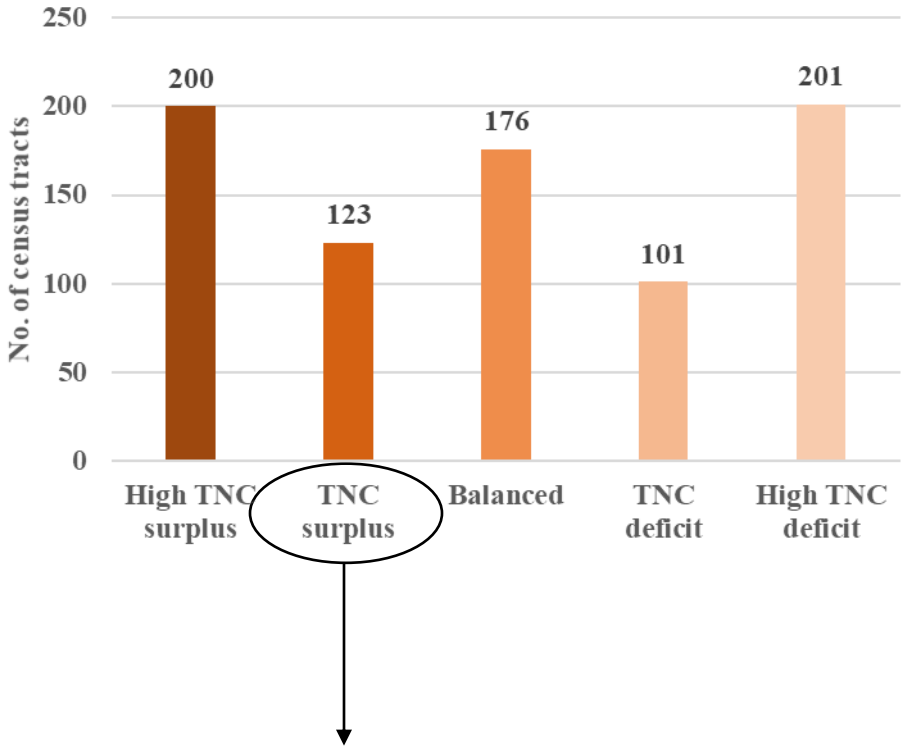


Overall transportation demand

Rank of overall transportation demand – Rank of TNC demand ≤ -2

Data Preparation & Dependent Variable

TNC demand

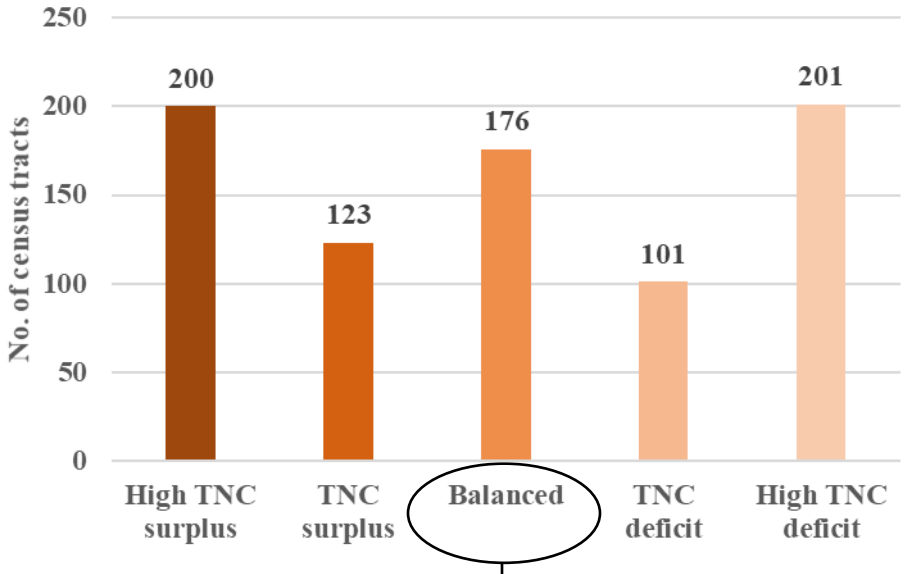


Overall transportation demand

Rank of overall transportation demand – Rank of TNC demand = -1

Data Preparation & Dependent Variable

TNC demand

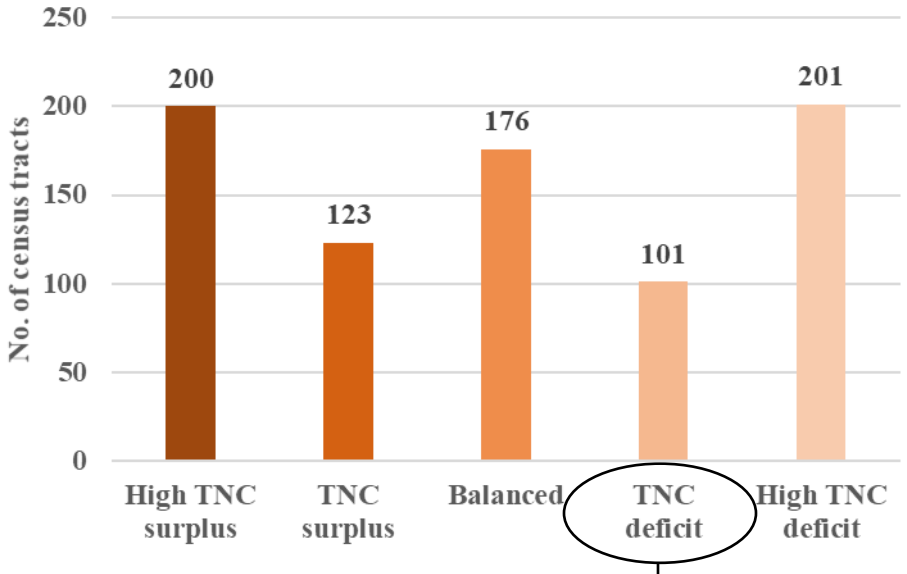


Overall transportation demand

Rank of overall transportation demand – Rank of TNC demand = 0

Data Preparation & Dependent Variable

TNC demand

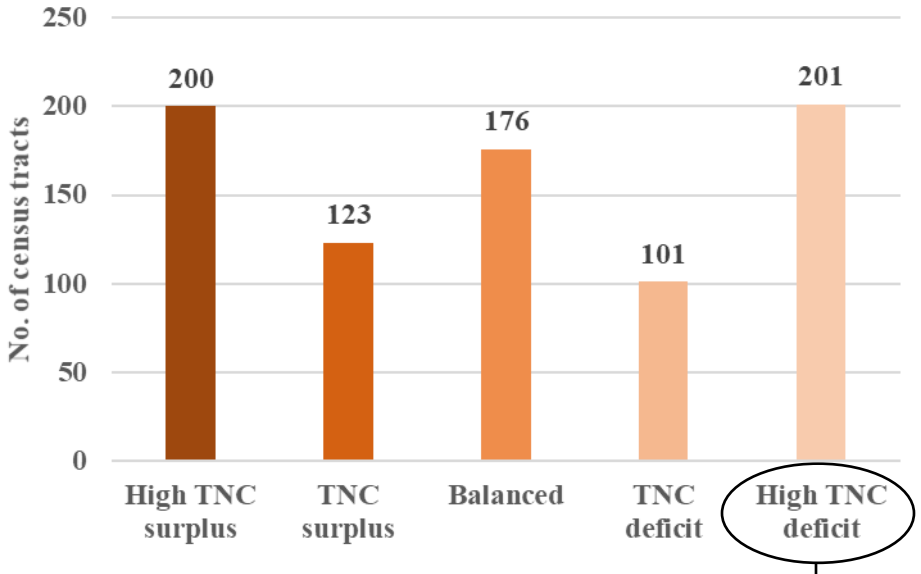


Overall transportation demand

Rank of overall transportation demand – Rank of TNC demand = 1

Data Preparation & Dependent Variable

TNC demand

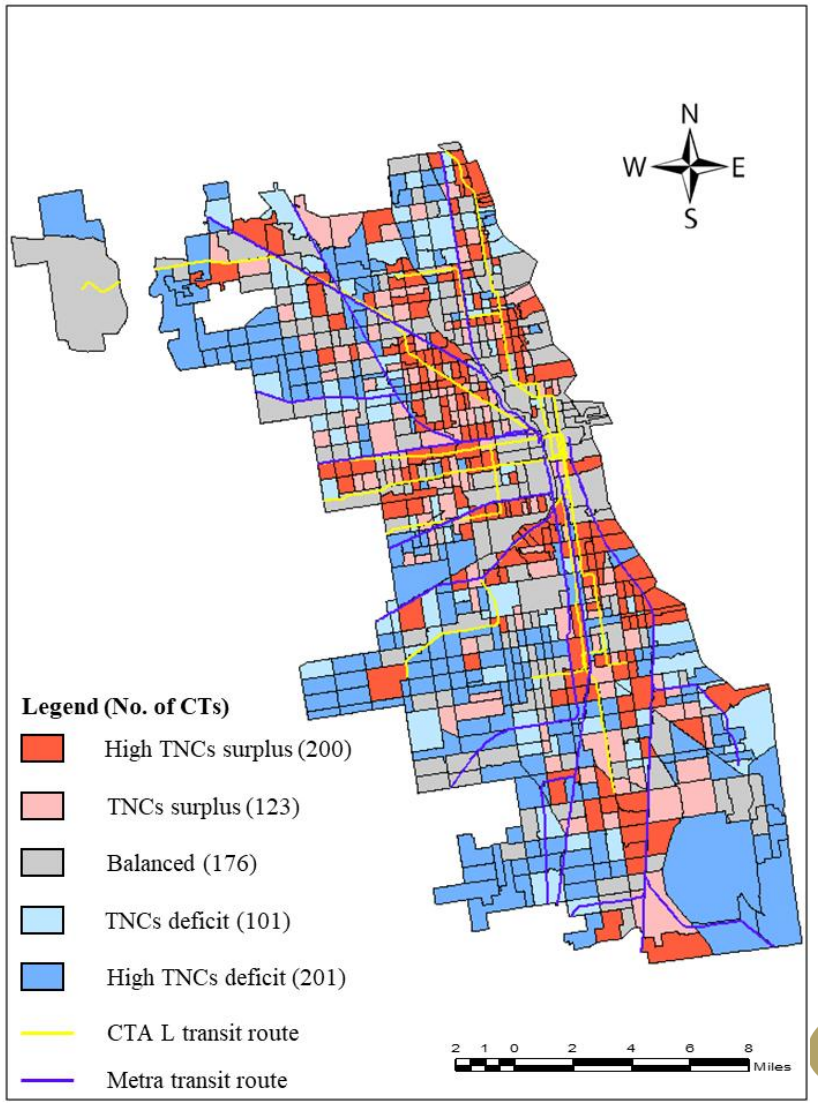


Overall transportation demand

Rank of overall transportation demand – Rank of TNC demand ≥ 2

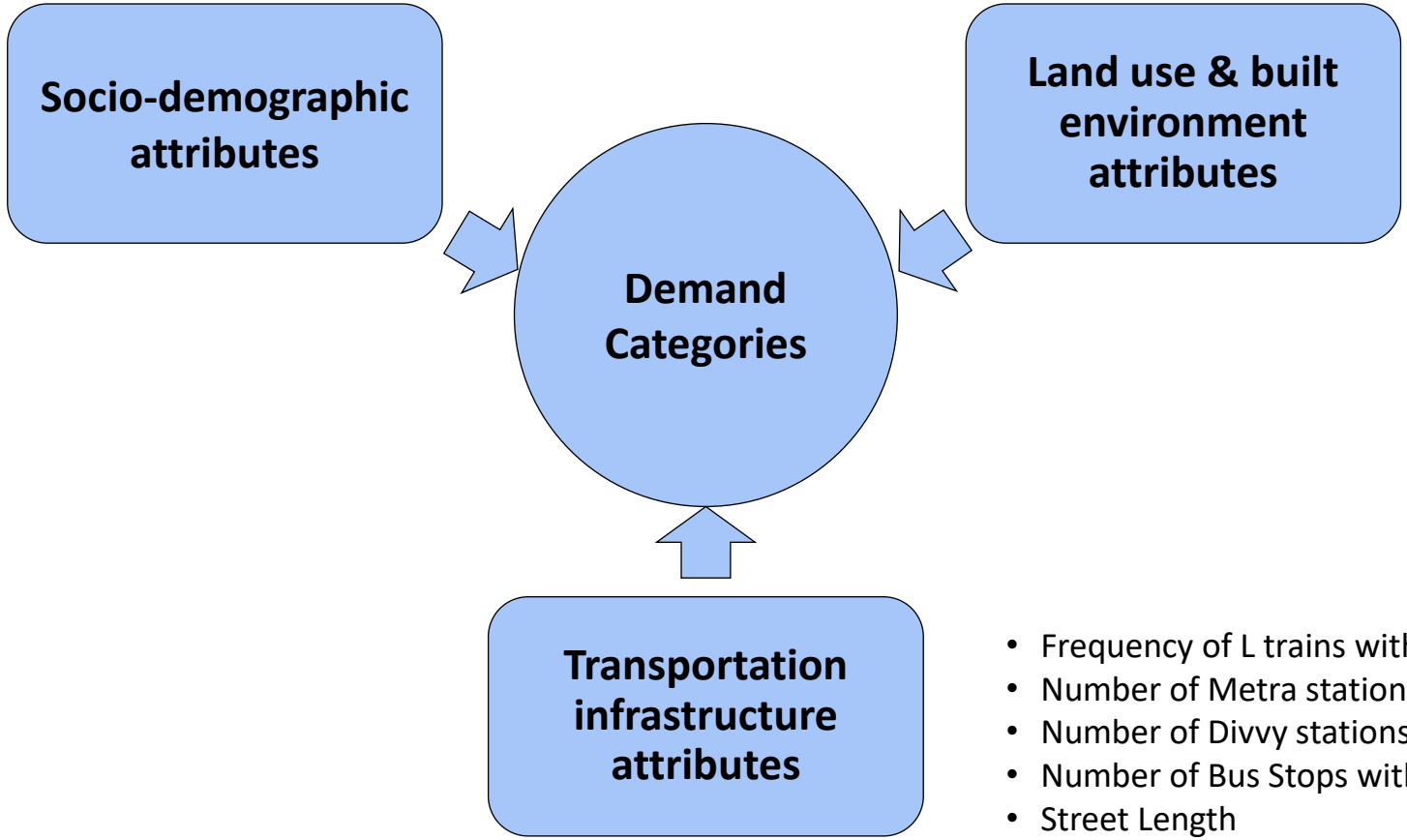
Data Preparation & Dependent Variable

- Distribution of ordinal dependent variable
- TNC deficit region – outer region of urban area
- Balanced region – downtown region
- High TNC surplus & TNC surplus region – located along transit routes



Data Preparation & Independent Variable

- Proportion of employment
- Proportion of Caucasian population
- Proportion of African American population
- Proportion of Latin and Hispanic population
- Proportion of population with age 15-24
- Proportion of population with age ≥ 65
- Proportion of HH with vehicles ≥ 2
- Low Income Indicator

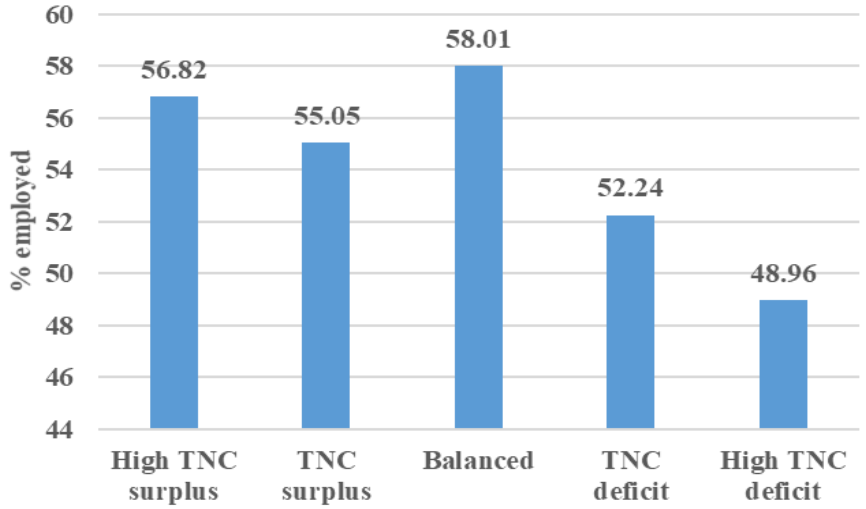


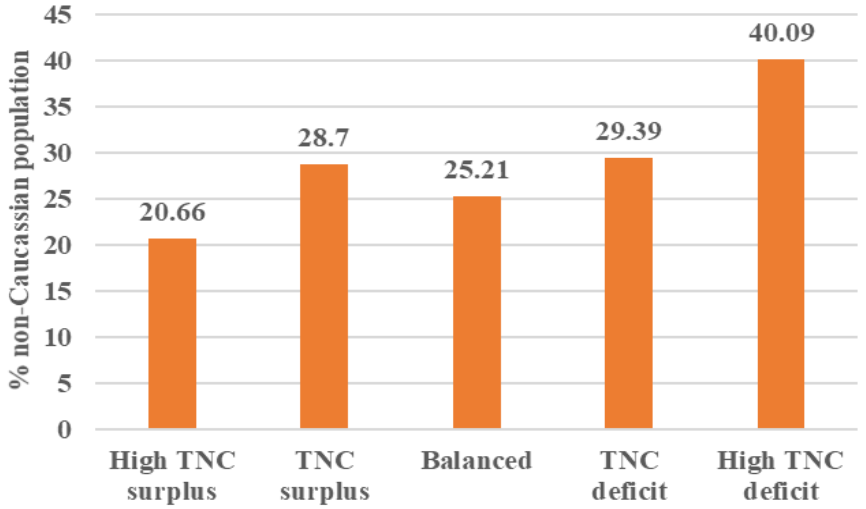
- Residential area
- Commercial area
- Institutional area
- Industrial area
- Number of restaurants
- Land use mix

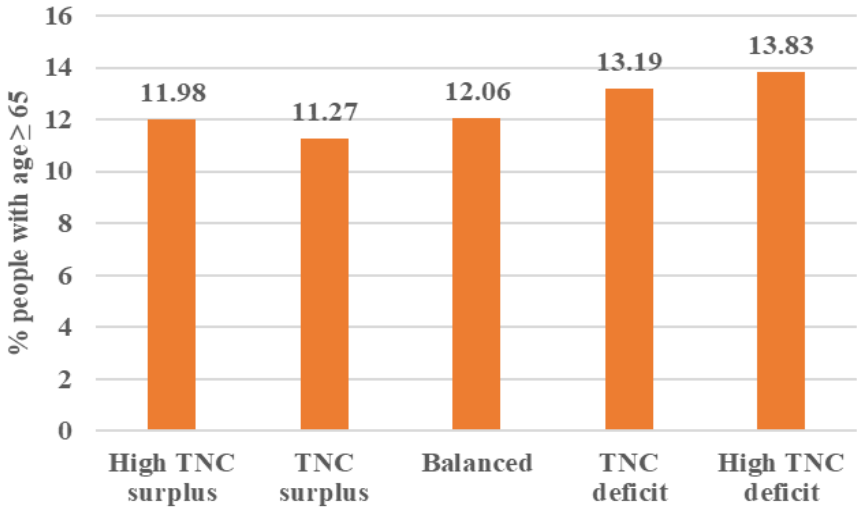
- Frequency of L trains within 1 mile buffer
- Number of Metra stations within 1 mile buffer
- Number of Divvy stations within 1 mile buffer
- Number of Bus Stops within 1 mile buffer
- Street Length
- Length of sidewalk

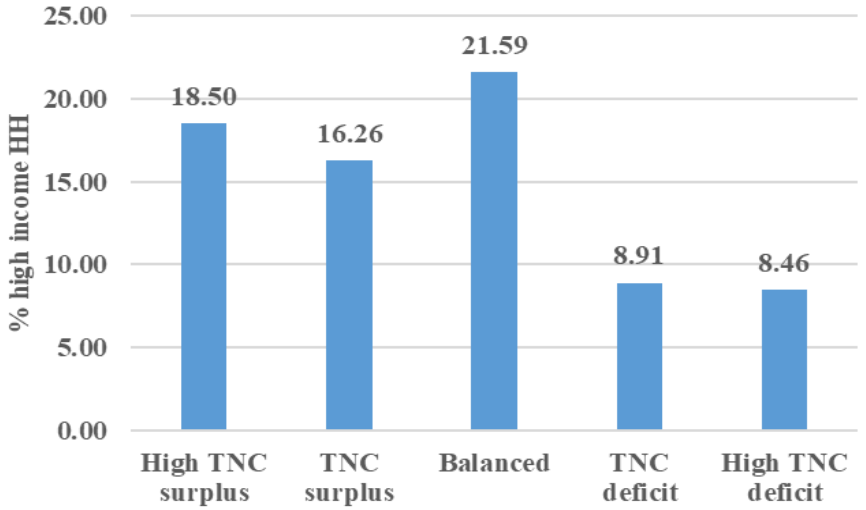
Georgia Tech

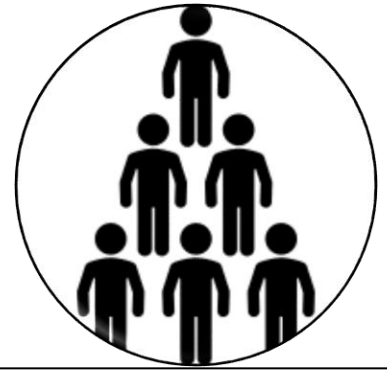
TNC Metric Analysis







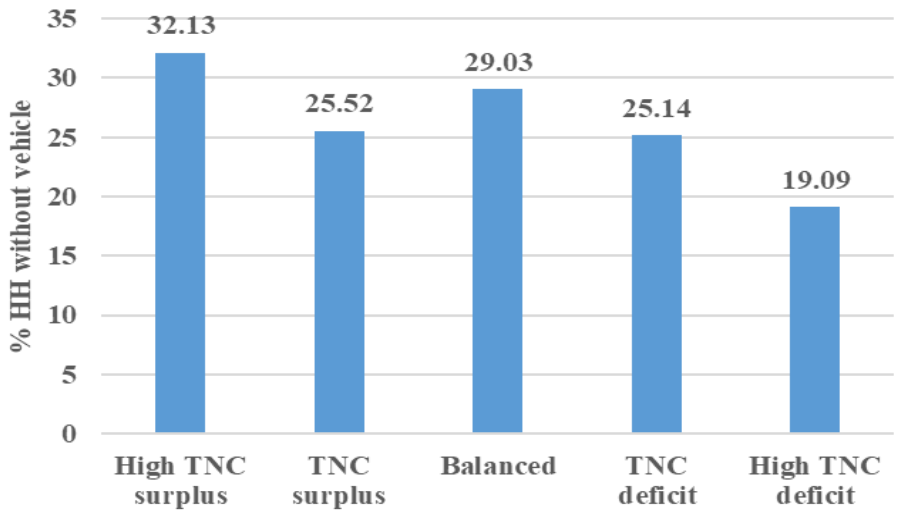


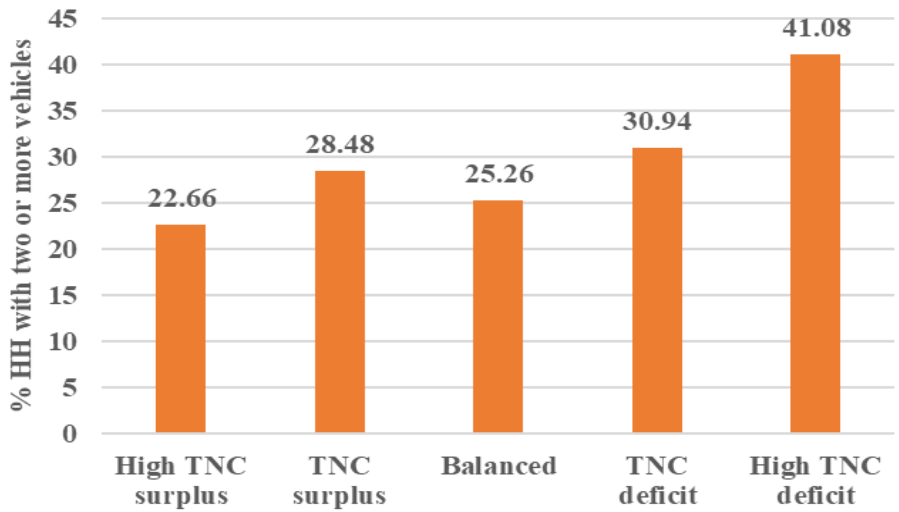


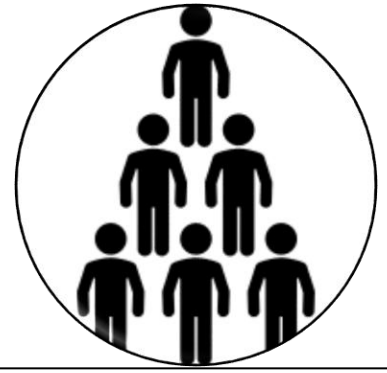
Socio-demographic Attributes

Georgia Tech

TNC Metric Analysis

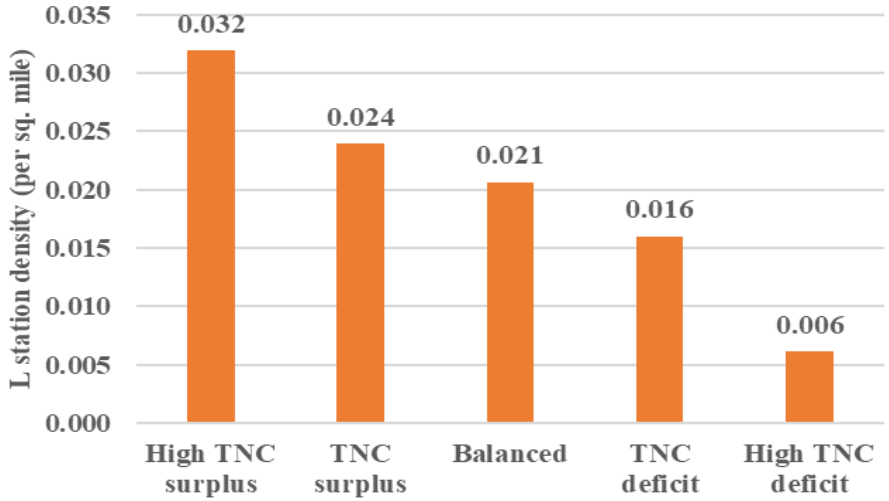


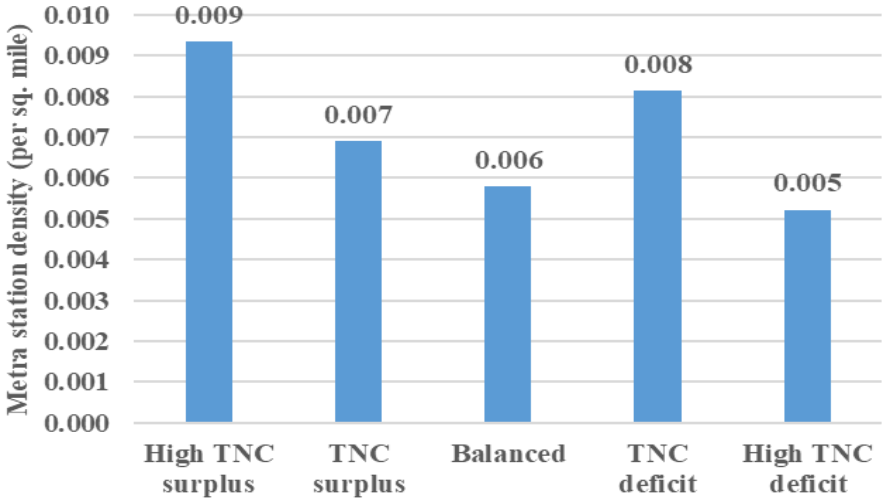


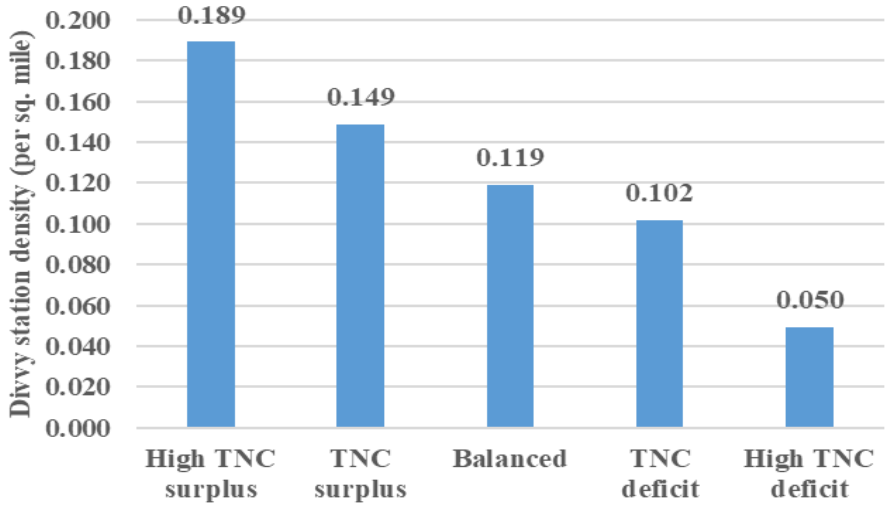


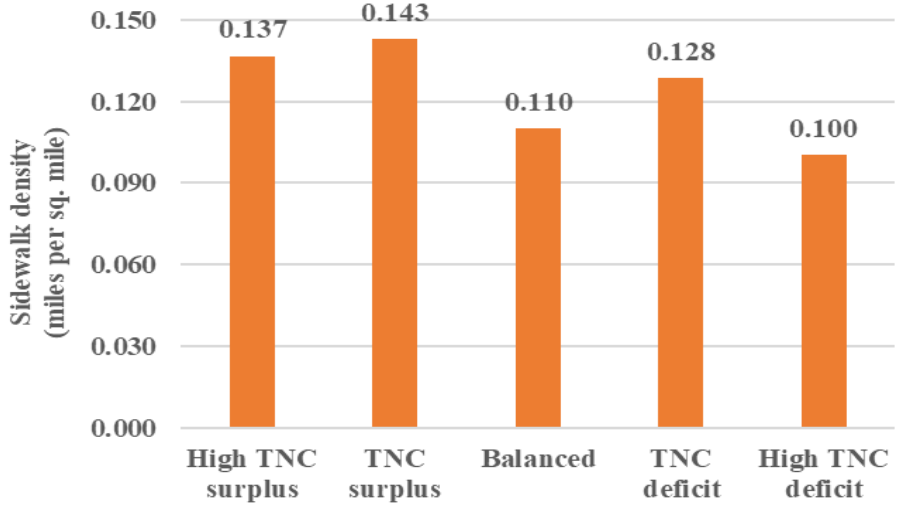
Socio-demographic Attributes

TNC Metric Analysis









Transportation Infrastructure Attributes

Georgia Tech

Econometric Methodology

Ordinal Demand Category

$$y_q = k, \qquad if\ \psi_{k-1} < y_q^* < \psi_k$$

Econometric Methodology

Latent Continuous Variable

$$y_q^* = (\alpha + \gamma_q) x_q + \xi_q$$

Ordinal Demand Category

$$y_q = k, \quad if\ \psi_{k-1} < y_q^* < \psi_k$$

- ξ_q — Logistic distributed error term
- x_q — Vector of explanatory variables
- γ_q — Unobserved effect
- α — Coefficient

Econometric Methodology

Latent Continuous Variable

$$y_q^* = (\alpha + \gamma_q)x_q + \xi_q$$

- ξ_q: Logistic distributed error term
- x_q: Vector of explanatory variables
- γ_q: Unobserved effect
- α: Coefficient

Ordinal Demand Category

$$y_q = k, \quad if\ \psi_{k-1} < y_q^* < \psi_k$$

Threshold Equation

$$\psi_{qk} = \psi_{k-1} + exp((\beta_k + \theta_{kq})z_{qk})$$

- z_{qk}: Vector of explanatory variables
- θ_{kq}: Unobserved effect
- β_k: Coefficient
- ψ_{k-1}: K-1 threshold

Econometric Methodology

Latent Continuous Variable

$$y_q^* = (\alpha + \gamma_q)x_q + \xi_q$$

- ξ_q: Logistic distributed error term
- x_q: Vector of explanatory variables
- γ_q: Unobserved effect
- α: Coefficient

Ordinal Demand Category

$$y_q = k, \quad if\ \psi_{k-1} < y_q^* < \psi_k$$

Threshold Equation

$$\psi_{qk} = \psi_{k-1} + exp((\beta_k + \theta_{kq})z_{qk})$$

- z_{qk}: Vector of explanatory variables
- θ_{kq}: Unobserved effect
- β_k: Coefficient
- ψ_{k-1}: K-1 threshold

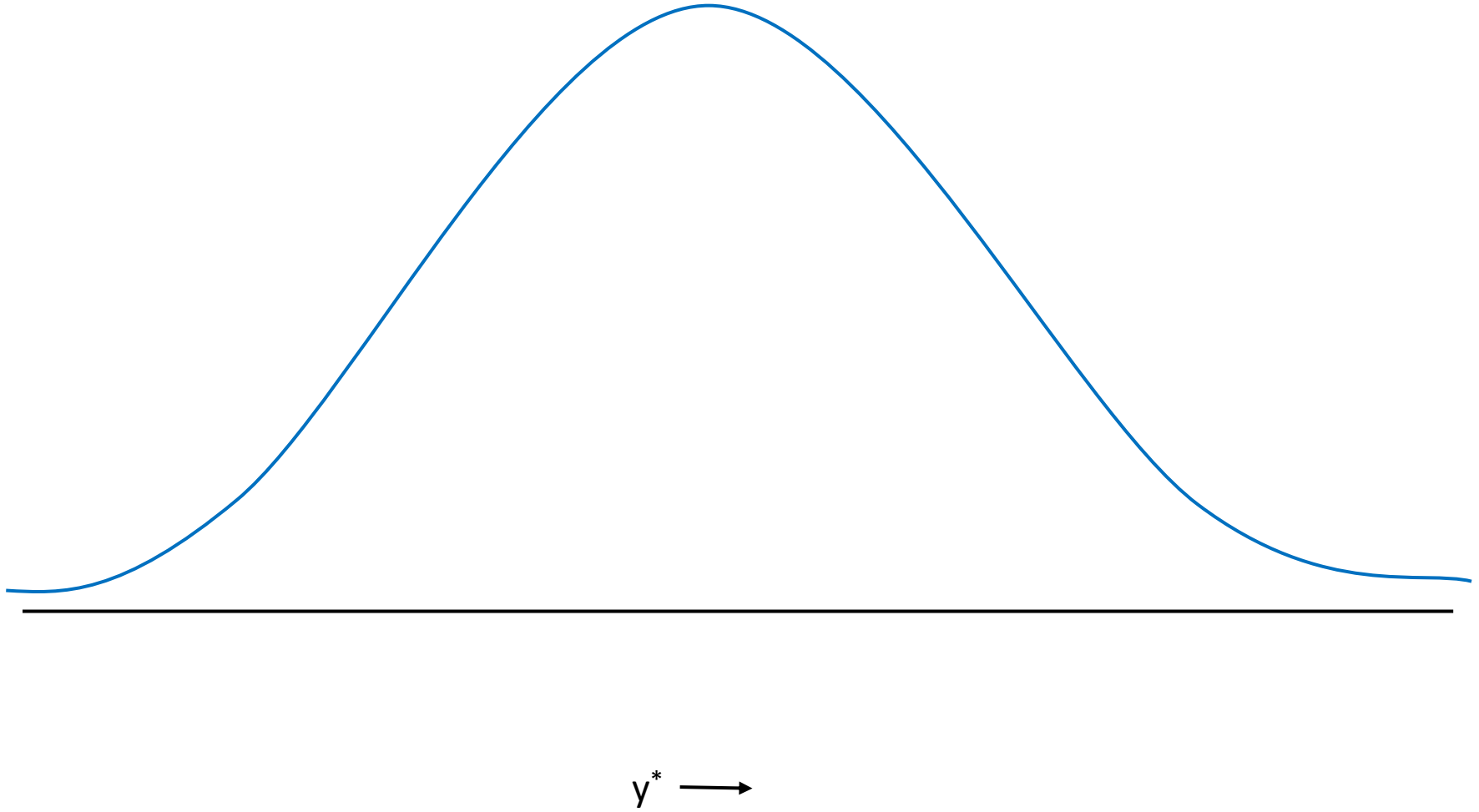
Probability Equation

$$Pr(y_q = k) = \Lambda(\psi_{qk} - (\alpha + \gamma_q)x_q) - \Lambda(\psi_{q,k-1} - (\alpha + \gamma_q)x_q)$$

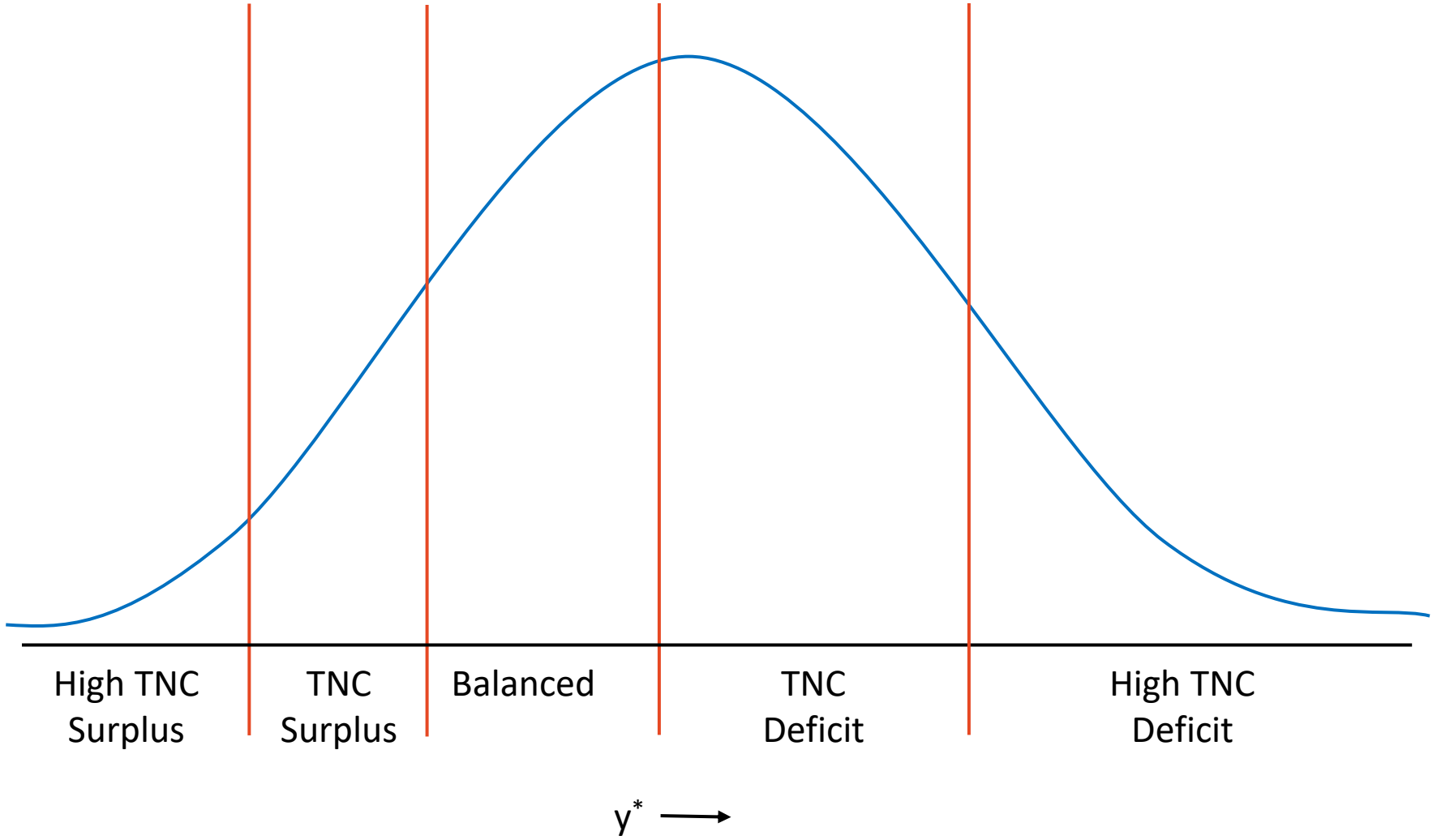
- Λ: Standard Logistic Cumulative Distribution

Georgia Tech

Econometric Methodology



Econometric Methodology



Econometric Methodology

Model Estimation

$$L_q = \int_{\Omega} \prod_{k=1}^{K} [Pr(y_q = k)]^{d_{qk}} \, d\Omega$$

d_{qk}: 1 if census tract q has a demand of level k and 0 otherwise

$$LL(\Omega) = \sum_{q} \ln L_q \, (\alpha, \psi, \Omega)$$

Measure of Fit

Model	Measures	
	Log-likelihood (LL)	Bayesian Information Criterion (BIC)
Ordered Logit (OL)	-1125.6	2418.3
Generalized Ordered Logit (GOL)	-1081.7	2397.4
Mixed Generalized Ordered Logit	-1080.4	2408.2

Estimation Result (Generalized Ordered Logit)

Variable	Propensity	Threshold between TNC surplus and Balanced	Threshold between Balanced and TNC deficit	Threshold between TNC deficit and High TNC deficit
Socio-demographic attributes				
Proportion of population with age between 15 and 24	--	--	▼	--
Proportion of population with age >65	▲	▼	--	--
Proportion of population with age >65 * Low income	▼	--	--	--
Proportion of Latin and Hispanic population	▲	--	--	▼
Proportion of African American population	--	--	--	▼
Proportion of physically challenged population	▼	--	--	--
Low-income census tracts	▲	--	--	--
Proportion of HH with 2 or more vehicles	▲	▲	--	--
Proportion of HH with 2 or more vehicles *CBD indicator	▲	--	--	--
Proportion of employment	▼	--	--	--

Georgia Tech

Estimation Result (Generalized Ordered Logit)

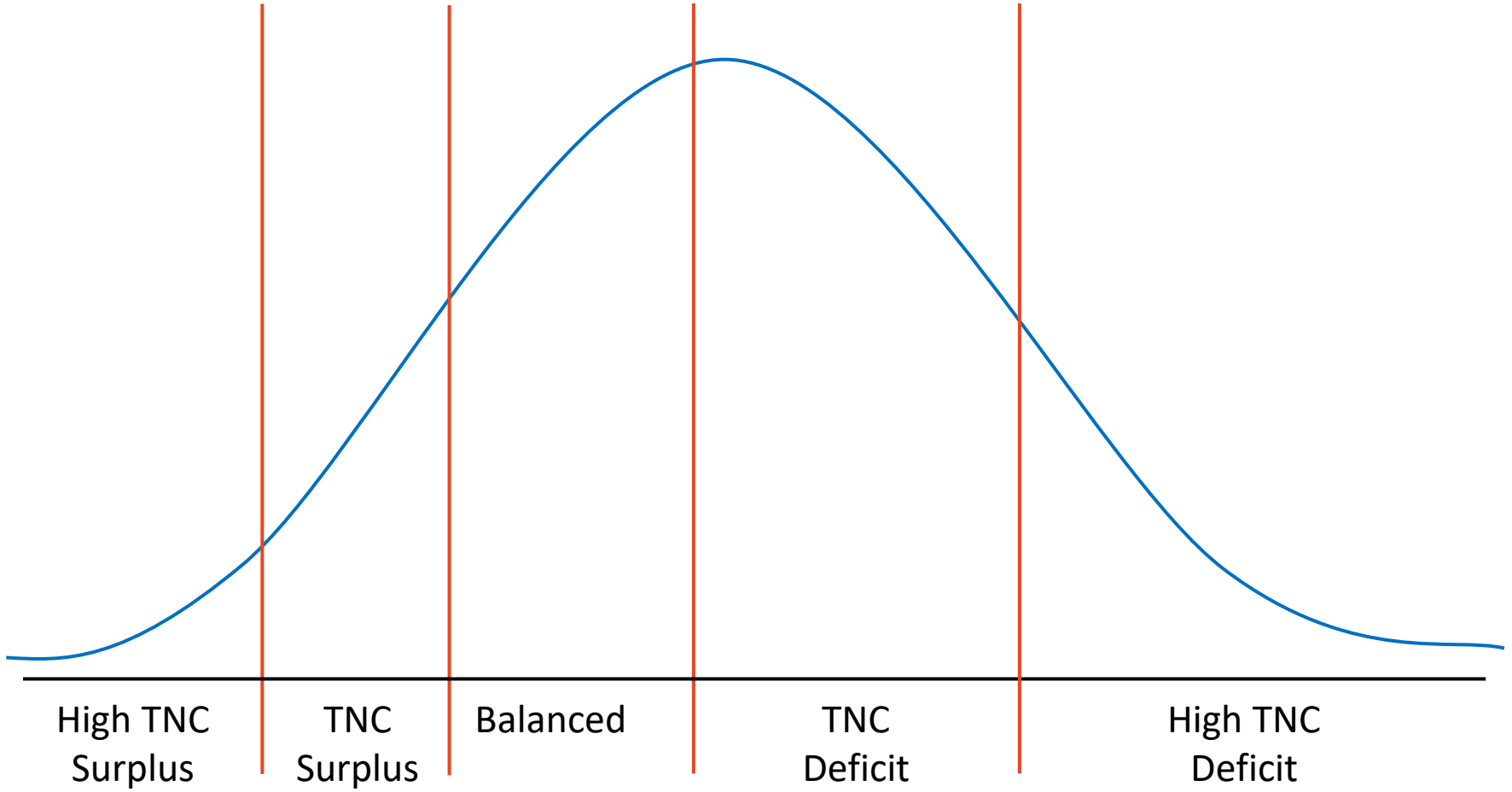
Variable	Propensity	Threshold between TNC surplus and Balanced	Threshold between Balanced and TNC deficit	Threshold between TNC deficit and High TNC deficit
Land use and built environment attributes				
Network distance from CBD	▲	--	▼	--
Proportion of residential area	▼	--	--	--
Proportion of residential area * Far south	▲	--	--	--
Proportion of commercial area	--	--	--	▲
Number of restaurants	▲	--	--	--
Length of sidewalk	▼	--	▲	--
Transportation infrastructure attributes				
Frequency of L trains within 1 mile buffer	▼	--	--	--
Frequency of L trains within 1 mile buffer* Far south	▼	--	--	--
Number of Bus stops within 1 mile buffer	▲	--	--	--
Number of Divvy stations within 1 mile buffer	▲	--	▲	--

Georgia Tech

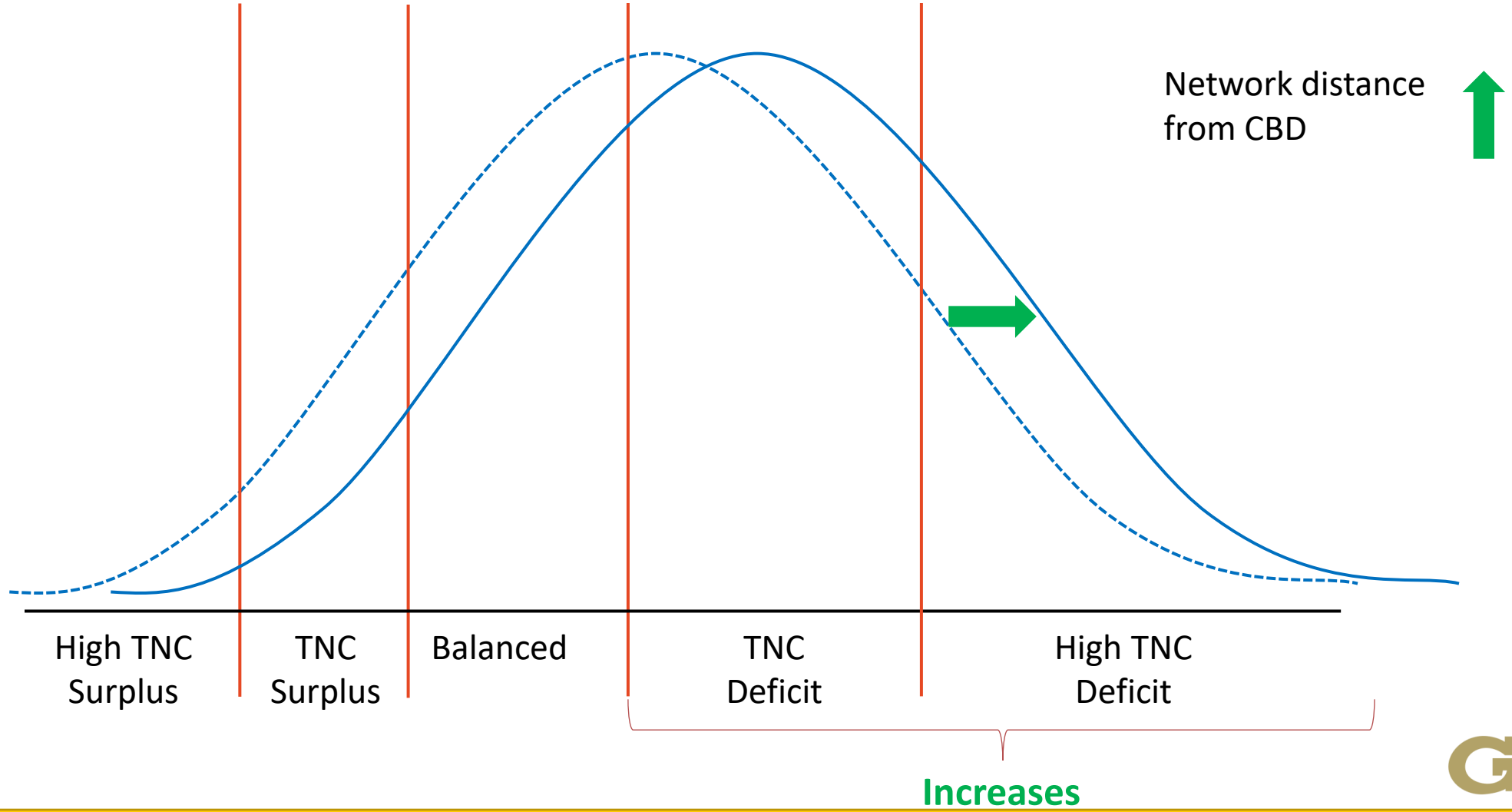
Estimation Result (Generalized Ordered Logit)

Variable	Propensity	Threshold between TNC surplus and Balanced	Threshold between Balanced and TNC deficit	Threshold between TNC deficit and High TNC deficit
Spatial attributes				
Central	▲	--	--	--
North	▲	--	--	--
West	▲	▲	--	--
Far south	▼	--	--	--

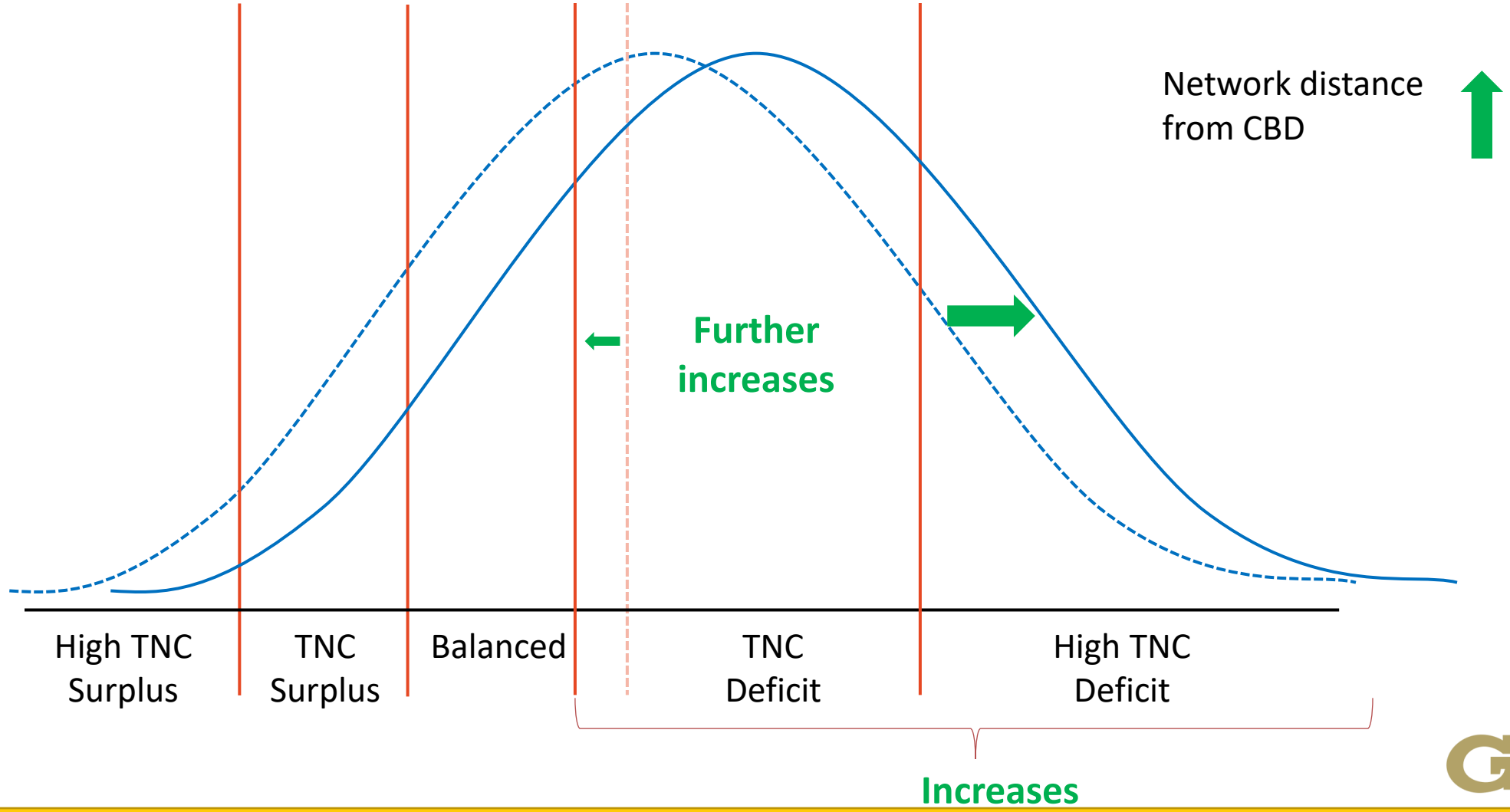
Estimation Result (Generalized Ordered Logit)



Estimation Result (Generalized Ordered Logit)



Estimation Result (Generalized Ordered Logit)



Elasticity Analysis

Variable	High TNC surplus	TNC surplus	Balanced	TNC deficit	High TNC deficit
Socio-demographic attributes					
Proportion of population with age between 15 and 24	0.000	0.000	-1.540	0.694	1.040
Proportion of population with age ≥ 65	-1.908	-3.701	0.245	1.201	3.437
Proportion of Latin and Hispanic population	-0.922	-0.693	-0.051	-1.885	2.378
Proportion of African American population	0.000	0.000	0.000	-1.382	0.703
Proportion of physically challenged population	0.726	0.331	-0.219	-0.340	-0.574
Low income census tracts	-5.054	-3.115	0.087	1.836	6.082
Proportion of household with two or more vehicles	-9.192	-1.685	0.120	2.324	9.104
Employment density	6.509	1.956	-1.487	-2.904	-5.020

Elasticity Analysis

Variable	High TNC surplus	TNC surplus	Balanced	TNC deficit	High TNC deficit
Land use and Built Environment attributes					
Network distance from CBD	-11.665	-6.501	-8.567	5.246	21.021
Proportion of residential area	2.316	0.668	-0.618	-1.449	-1.477
Proportion of commercial area	0.000	0.000	0.000	2.498	-1.271
Number of restaurants	-1.471	-0.341	0.556	0.868	0.764
Length of sidewalk	3.194	1.686	4.228	-3.097	-6.554
Transportation infrastructure attributes					
Frequency of L trains within 1 mile buffer	2.231	0.345	-0.774	-1.228	-1.158
Number of Bus stops within 1 mile buffer	-6.336	-2.256	1.329	3.072	5.089
Number of Divvy stations within 1 mile buffer	-1.044	-0.407	1.613	-0.249	-0.016

Elasticity Analysis

Variable	High TNC surplus	TNC surplus	Balanced	TNC deficit	High TNC deficit
Spatial attributes					
Central	-5.287	-3.204	-0.033	1.720	6.542
North	-2.856	-1.557	0.468	0.853	3.023
West	-3.540	2.020	0.123	0.695	1.856
Far south	16.331	-1.547	-5.996	-6.141	-7.060

Summary

Current study

- Evaluating TNCs in the context of overall transportation demand.
- Compute a novel metric to identify the potential imbalance between overall transportation demand and TNC demand.
- Framework developed can be applied for any urban region to identify spatial pockets underserved by TNCs.

Findings

- Based on LL and BIC, GOL model offers superior data fit.
- Univariate analysis of TNC metric shows spatial pockets that are underserved.
- Among socio-demographic variables, the results indicate a clear demarcation in TNC usage based on age and ethnicity composition.

Limitation

- The model developed can be further enhanced in the presence of more disaggregate spatial data.

Georgia Tech

Thank You