NYC Bike Equity

A network-based GIS analysis on the effects of station placement on employment accessibility

ADELINE SETIAWAN | DEC 2025

ABSTRACT

How does network topology shape bike-share equity in New York City? This study applies network-based GIS analysis to evaluate how CitiBike station placement affects job accessibility across 2,324 NYC census tracts. Using street network routing rather than traditional Euclidean buffers, we reveal some blindspots in conventional accessibility measures. We found that 6.8% of tracts have no routable path to any CitiBike Stations, while 31.6% lie beyond practical cycling distance (> 30 min). Our network topology approach exposes a stark spatial divide. Peripheral communities in Eastern Queens, Southeast Brooklyn, and Staten Island face both high rates of unemployment and structural disconnection from bike-share infrastructure. Spatial analysis further reveals strong clustering of job accessibility (Moran's I = 0.31), with high-access areas concentrated in Manhattan and western Brooklyn / Queen's. While nearby tracts share similar accessibility levels, suggesting potential spillover benefits, network barriers prevent these benefits from reaching disconnected zones. LISA cluster analysis confirms low-low accessibility clusters precisely where unemployment is highest, revealing systematic spatial mismatches. Most critically, overlay analysis shows that communities with greatest need—high unemployment, low car ownership, poor job access—are topologically excluded from bike-share networks. These findings demonstrate that infrastructure placement follows existing spatial advantage rather than addressing transportation gaps. For GIS practitioners and planners, results highlight the importance of network-based over buffer-based accessibility measures and suggest that equity-oriented bike-share expansion requires strategic network connectivity improvements, not simply adding stations to underserved areas.

1. BACKGROUND

In large metropolitan cities such as New York City (NYC), access to employment is a spatial problem. Labour market outcomes are often analysed through individual characteristics or macroeconomic conditions. In lived realities however, residents' ability to reach job opportunities depends as much on transportation infrastructure and urban form. Spatial mismatches between where one lives and where jobs are located can limit employment opportunities. This is especially pronounced for low-income and transport constrained populations.

The average American's commute to work in 2025 is about 9.69 miles[1]. The mean one-way travel time was 27.2 minutes, up from 26.8 minutes in 2023 [2] Every day, up to 2 million NYC residents take the extensive rail network to work. Given that the extensive rail system has more than doubled in the last decade [3], transportation remains a substantial expense for many New Yorkers. Yet the greatest share of transportation services are located in neighbourhoods with above average housing costs.

Workers below 200% of the poverty line not only rely more on public transport, they also travel longer distances. This impacts opportunities, family time, adding additional cost an emotional burden of commuting to work [4]. Inequitable access to transportation perpetuates socioeconomic divides [5].

Most studies of bike-share equity focus on usage, station siting, or ridership demographics. Fewer examine the relationship between bike-share networks and employment accessibility at the neighbourhood level. Fewer still use network-based measures rather than Euclidean distance.

This study examines the association between bike-share access and employment accessibility across New York City census tracts. It emphasises network connectivity, structural absence of access, and spatial dependence. Key neighbourhood confounders are explicitly controlled NYC DOT data shows cycling as the fastest-growing transport mode. About 55,000 bike to work [4], with major increases since bike-share systems have expanded rapidly across U.S. cities in the past decade, including New York City's CitiBike program.

Although bike-share is frequently framed as a recreational mode, its potential role as a form of employment-access infrastructure remains under-examined. Bike-share benefits are not spatially uniform. In gentrified neighbourhoods near the Manhattan Central Business District, 2.3% of commuters travel by bicycle. In the Bronx, the share is 0.4% [6]. Station placement, network connectivity, and neighbourhood socioeconomic context shape both usage and potential accessibility gains.Station placement, network connectivity, and neighbourhood socioeconomic contexts all play a role in the usage of bike-sharing networks.

This paper asks: ***Does bike-share infrastructure mitigate or reproduce spatial inequality in employment accessibility, and how does network topology predetermine who can benefit from transportation investments?***

Rather than focusing solely on employment status, this study looks at employment accessibility through the lens of bike equity. This is defined as the number of jobs residents can reach within a fixed travel time (30 min) on a bicycle. We use a network-based spatial framework to analyse the relationship between bike-share access and employment accessibility.

Using New York City census tracts as the unit of analysis, the study integrates American Community Survey (ACS) data, Longitudinal Employer–Household Dynamics (LODES) job data, Citi Bike station locations, and a bikeable street network [1,2,7].

The analysis estimates the association between multiple dimensions of bike-share access and employment accessibility. Models control for median household income, population density, and spatial autocorrelation. By combining network-based accessibility measures with spatial regression, this paper contributes to debates on transportation equity and the role of micro-mobility in shaping access to labor markets [3,4,5].

—

[1] https://www.census.gov/topics/employment/commuting/guidance/acs-1yr.html

[2[https://www.census.gov/topics/employment/commuting/guidance/acs-1yr.html. Note this assumes uniform speed, ignores elevation, traffic, infrastructure quality

[3] https://www.osc.ny.gov/press/releases/2024/09/draft-dinapoli-nyc-metro-transportation-costs-grew-slower-other-major-cities-still-rose-29-percent

[4] The working poor living in the three largest metropolitan areas of the Northeast (Boston, New York, and Philadelphia) and Washington, DC, have a higher cost burden of both commuting and housing than the national median for the working poor. https://inequality.stanford.edu/sites/default/files/media/_media/pdf/key_issues/transportation_policy.pdf

[5] Housing & Transportation Cost Trade-offs and Burdens of Working Households in 28 Metros, Haas, Makarewicz, Benedit

[6] https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/walking-driving-and-cycling/?id=2415#display=summary

[7] https://www.nyc.gov/assets/planning/download/pdf/planning-level/housing-economy/nyc-ins-and-out-of-commuting.pdf

2. CONCEPTUAL FRAMEWORK

2.1 Employment Access as a Spatial Outcome

Employment accessibility is teated as a spatially structured outcome. It is shaped not only by labour supply and demand, but by the spatial interaction between workers and locations of job opportunities. From a transportation geography perspective, employment access reflects the sets of jobs which are reachable within acceptable time and cost constraints, given available modes of travel. In dense urban environments, even small changes in transit options can cause friction, substantially altering the size of one's opportunities.

Bike-share systems influence employment access. First, they reduce access costs to transit stations and job centers by improving first- and last-mile connectivity. Second, accessibility is not solely a function of proximity. Some census tracts are structurally disconnected from the bike share network, prompting questions on gradual distance decay and complete absence of station access. By operating on street networks rather than fixed routes, bike-share offers flexible point-to-point mobility that may be especially valuable for shift workers and non-standard work schedules.

Third, neighbourhood socioeconomic characteristics shape both transportation infrastructure placement and baseline job accessibility. Median household income is included to account for economic geography and historical investment patterns. Population density is included to capture urban form and concentration of economic activity. These variables are treated as confounders rather than mechanisms of interest.

Crucially, employment accesibility is expected to exhibit spatial dependence. Adjacent census tracts share transportation infrastructure, labor markets, and land-use patterns. As a result, accessibility outcomes may be spatially clustered beyond what is explained by observed covariates.
Taken together, this framework treats bike-share infrastructure as a component of a broader spatial system. Its association with employment accessibility is evaluated conditional on neighborhood context and spatial structure, rather than as an isolated or universally beneficial intervention.

2.2 Hypotheses

Based on this framework, the study advances three primary hypotheses:

H1: Bike-share access is positively associated with employment accessibility.

Increasing access to CitiBike stations expands reachable jobs. We measure job accessibility as network based (not Euclidean), and absence of stations by station density. Access is measured within a 5 mile radius by proximity to bike stations or a 30-minute bike ride [2]. Job accessibility is hypothesised to be substantially lower in tracts which lack any reachable stations within the bike network.

This relationship is estimated in multivariate regression models with controls for confounding factors, and correlated with both station placement / absence, and baseline job accessibility (Table X).

H2: Bike and employment accessibility varies with neighbourhood socioeconomic context.

The relationship between bike-share access and job access is expected to differ across neighbourhoods with distinct socioeconomic characteristics. We include median household income and population density as controls to account for systematic differences. These could range from urban form, labour market proximity, and infrastructure investment that may otherwise confound the estimated effects of bike-share access. **This hypothesis is examined with robustness checks rather than as a primary causal claim (Table Y).**

H3: Employment accessibility exhibits spatial dependence across tracts.

Employment accessibility is expected to be spatially concentrated across certain census tracts. We use diagnostic tests and spatial regression models to test for spatial autocorrelation and account for this dependence. Spatial error and spatial lag specifications are used to prevent biasing coefficient estimates and inference. We also use it assess the robustness of estimated relationship.

Together, these hypotheses frame bike-share accessibility as a spatially contingent part of the NYC transportation infrastructure whose equity impacts depend on neighbourhood context and network placement.

3. Study Area, Data, and Variables

3.1 Study Area and Unit of Analysis

The study area comprises of 2,324 census tracts covering New York City's five boroughs: The Bronx, Brooklyn, Manhattan, Queens, and Staten Island. New York City provides a suitable context for examining bike-share accessibility due to its dense employment geography, extensive bike-share network, and pronounced socioeconomic heterogeneity.

Census tracts provide an appropriate spatial scale for neighbourhood-level analysis while maintaining statistical reliability for American Community Survey (ACS) estimates. This unit also aligns directly with the spatial resolution of ACS and Longitudinal Employer–Household Dynamics (LODES) datasets, commonly used in transportation equity analyses. Geocoded workplace locations and job counts are aggregated to census tracts. These data represent the spatial distribution of employment opportunities rather than individual employment outcomes.

3.2 Network Construction and Access Measures

We construct a routable street network from NYC Street Centerline data, converted to an undirected graph with edge weights representing physical street length in feet. This topology-aware approach captures true cycling distances and identifies disconnected components—critical for revealing structural exclusion.

Network-Based Access Variables:

- **Distance to nearest station** (dist_to_station_ft_net): Shortest-path distance from tract origin to nearest Citi Bike station, NA if no routable path exists
- **Structural exclusion** (no_station_access_net): Binary indicator for tracts with no routable path to any station
- **Capped distance** (dist_to_station_mi_cap): Distance in miles, threshold choice at 5 miles to limit leverage from extreme values

The bike network consists of ~ 79,217 nodes and 122,053 edges. Of the **2,327 census tracts** in the study area, **157 tracts (6.76%)** have no reachable Citi Bike station on the network.[1]

Tracts with no reachable Citi Bike stations are identified using a binary indicator. This distinction separates tracts that are distant but connected from those that are disconnected from the bike-share network entirely. Disconnection typically reflects geographic barriers, limited street connectivity, or absence of nearby stations rather than data error.

3.3 Data Sources (see Appendix)

3.3.1 Dependent Variable: Employment Accessibility

We measure employment accessibility as the number of jobs within a 30-minute bike ride, or approximately 5 miles along the Citi Bike street network (defined in Section 3.3.2). Distances are capped to limit the influence of extreme outliers, and reflect the median cycling commute thresholds. We incorporate both.

Using LODES Workplace Area Characteristics data, we compute tract-to-tract shortest paths and sum destination employment within the threshold.

3.3.2 Street Network

To derive a bikeable street network, only links suitable for cycling are included from the NYC Centerlines dataset. This network defines feasible bicycle routes and underpins all accessibility calculations (e.g. network distances and travel paths). Accessibility calculations depend on successful snapping of census tract interior point to this network.

To ensure geometric validity, we represent each tract by one interior point[2] using st_point_on_surface( ), avoiding issues with centroids falling outside irregular coastal polygons. The street network is projected in EPSG:2263 (NAD83 / New York Long Island) to ensure distance calculations are performed in planar units appropriate for New York City.

3.3.3 Bike-Share Infrastructure (no station access)

Citi Bike station locations and dock capacity are obtained from the Citi Bike General Bikeshare Feed Specification (GBFS) API. Stations define access points to the bike-share system, and dock capacity defines real-time available bikes and empty docks at each station[3]. The analysis focuses on station presence and network reachability rather than observed ridership or trip counts. Temporal assumptions are made with cross-sectional snapshots[4] (see Section 3.3.4).

---

1. Missing trips: 3 tracts could not snap to network. We treat this as missing rather than isolated (see section 3.4.1)
2. Methodological limitation: one interior point per tract may not represent population distribution within tract
3. Data is missing the demand side. This analysis does not incorporate actual ridership or bike availability
4. We can't establish causality or account for station placement endogeneity because it's cross-sectional

3.3.4 Control Variables

Models control for median household income (ACS), population density, and the percentage of households without vehicle access, to capture socioeconomic variation and transportation constraints that may confound the bike-share–accessibility relationship.

3.4. Spatial Econometric Framework

We expect employment accessibility to exhibit spatial dependence due to shared transportation networks, labour markets, and neighbourhood context across adjacent census tracts. This analysis employs spatial econometric models alongside conventional regression to account for this structure.

3.4.1. Spatial Weights Matrix:

We use Queen contiguity weights (row-standardised) to capture spatial relationships between census tracts. Queen contiguity is appropriate given NYC's irregular tract boundaries which often meet at vertices rather than edges. The resulting weights matrix contains 3 isolates and 5 disconnected subgraphs, reflecting natural geographic barriers such as water bodies. These isolates are treated as missing, and retained using standard zero-policing handling.

Model Specifications:

1. **Ordinary Least Squares (OLS)** used as baseline to estimate conditional relationships between bike-share access and employment accessibility,
2. **Spatial Lag Model (SLM)**: $y=\rho Wy+X\beta+\varepsilon$ to capture accessibility spillovers. Accessibility in one tract may depend on accessibility in neighbouring tracks,
3. **Spatial Error Model (SEM)**: $y=X\beta+u$, where $u=\lambda Wu+\varepsilon$.This model accounts for omitted spatial structure: unobserved spatially correlated factors which influence employment accessibility, but are not explicitly included as covariates.

We used Lagrange Multiplier tests to diagnose spatial dependence and guide model selection. When both spatial lag and spatial error processes are significant, statistics are used to assess the robustness of coefficients.

4. LIMITATIONS

4.1 Temporal Assumptions and Limitations

This analysis adopts a cross-sectional design. Socioeconomic characteristics are measured using ACS (2018 - 2022) data, while employment locations are drawn from LODES 2021. Citi Bike stations data is taken from 2025 snapshot. Citi Bike data and the street network reflect the contemporary configuration at the time of analysis. We acknowledge there's a temporal mismatch.

Job geography is treated as stable relative to transportation infrastructure over this short temporal window. The study does not examine temporal dynamics or changes in employment outcomes.

4.2 Model Limitations

We acknowledge the limitations of the model as beyond the scope of this analysis:

1. No interaction terms: missing test of H2: car constrained neighbourhoods benefit more
2. Spillover interpretation; SLM shows $\rho \approx 0.30$-0.41 but with network measures, SEM is preferred (suggesting omitted spatial structure rather than true diffusion) (see Table 1)
3. Scale dependency: results sensitive to 30-minute threshold choice
4. Missing alternative transit options: No control for subway access, or mixed modes of transit (e.g. subway and CitiBike)

4.3 SUMMARY OF VARIABLES

DEPENDENT VARIABLE: EMPLOYMENT ACCESSIBILITY

(log - transformed) count of jobs accessible within 30 min bike ride. (~ 5 miles network distance)

Due to extreme right-skew (median = 387,788; max = 2,546,141), we log-transform the outcome (i.e. log_jobs_access = log1p(jobs_access_30min))

INDEPENDENT VARIABLE: PRIMARY EXPOSURE (network-based)

Network distance to nearest CitiBike station (max 2 miles)
Operationalised by taking shortest path from tract interior point to station via street network.

157 tracts (6.8%) have no routable path;
735 tracts (31.6%) are > 2 miles away

Alternative Exposure (density-based): STATION DENSITY
Citibike stations per square mile within tract. Used in baseline models for comparison

---

CONTROLS

1. Median income - ACS median household income
2. Population density - population per square mile
3. Percentage renter - calculated but not used in final models
4. Unemployment rate - used for mapping / equity analysis only

4.4 VARIABLES AND DATA SOURCES

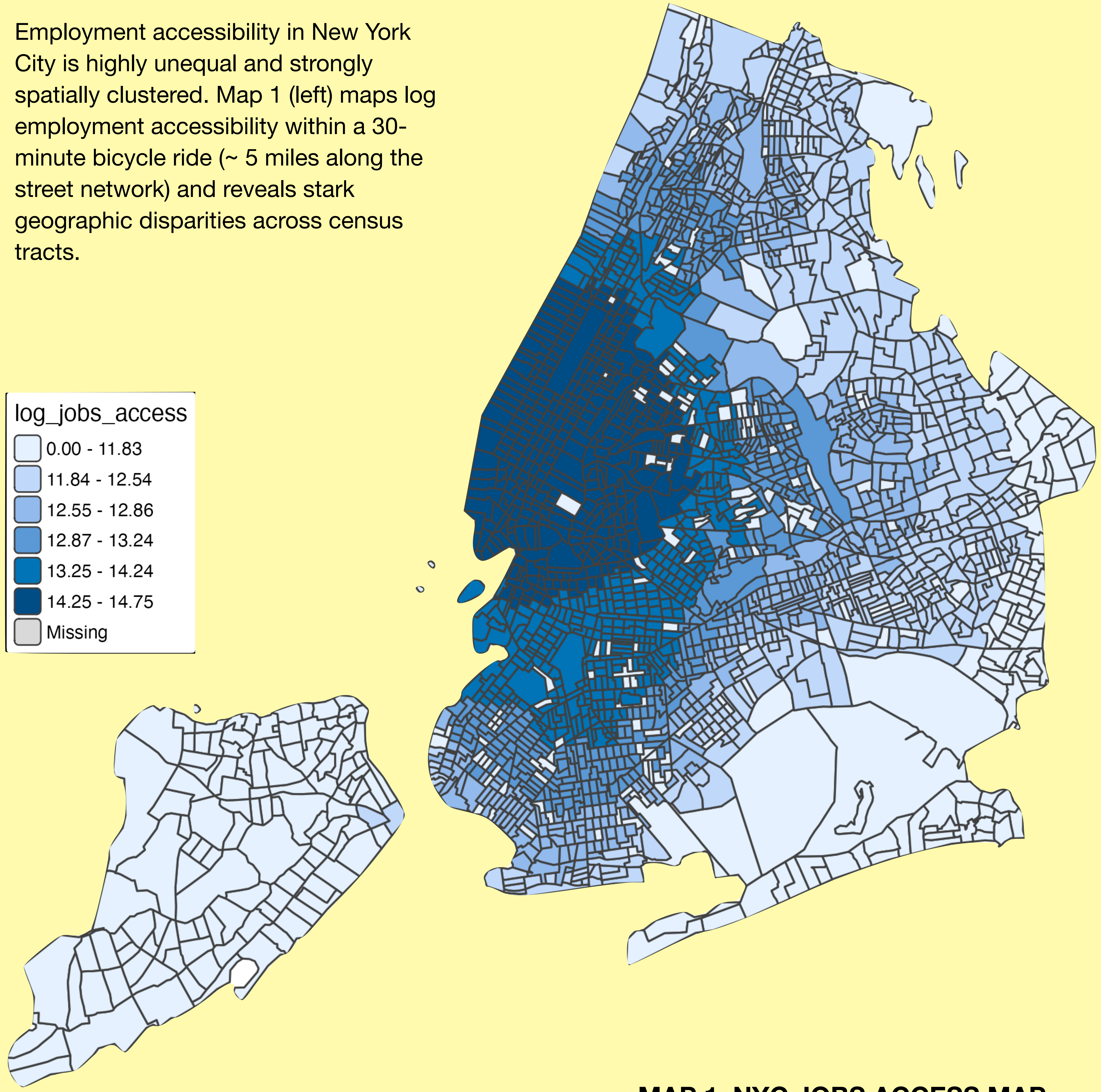
Variable Name	Description	Data Source	Role in Model	Transformation
log_jobs_access	Log of network-based job accessibility index	LODES 2021 + bike network	Dependent variable	Log-transformed
dist_to_station_mi_cap	Network distance to nearest Citi Bike station, capped at 2 miles	Citi Bike GBFS + street network	Key independent variable	Capped, continuous
no_station_access_net	Indicator for no reachable Citi Bike station	Citi Bike GBFS + network	Key independent variable	Binary (0/1)
Median household income	Tract-level median income	ACS 2022	Control (confounder)	Level
Population density	Residents per unit area	ACS 2022	Control (confounder)	Level or log
Unemployment rate	Share of labor force unemployed	ACS 2022 (B23025)	Descriptive / robustness	Level
% households without vehicle	Proxy for car constraint	ACS 2022	Descriptive / robustness	Level
Educational attainment	% bachelor’s degree or higher	ACS 2022	Control / context	Level

TABLE 1

5. RESULTS : DESCRIPTIVE PATTERNS

5.1 Spatial inequality in employment accessibility

Employment accessibility in New York City is highly unequal and strongly spatially clustered. Map 1 (left) maps log employment accessibility within a 30-minute bicycle ride (~ 5 miles along the street network) and reveals stark geographic disparities across census tracts.



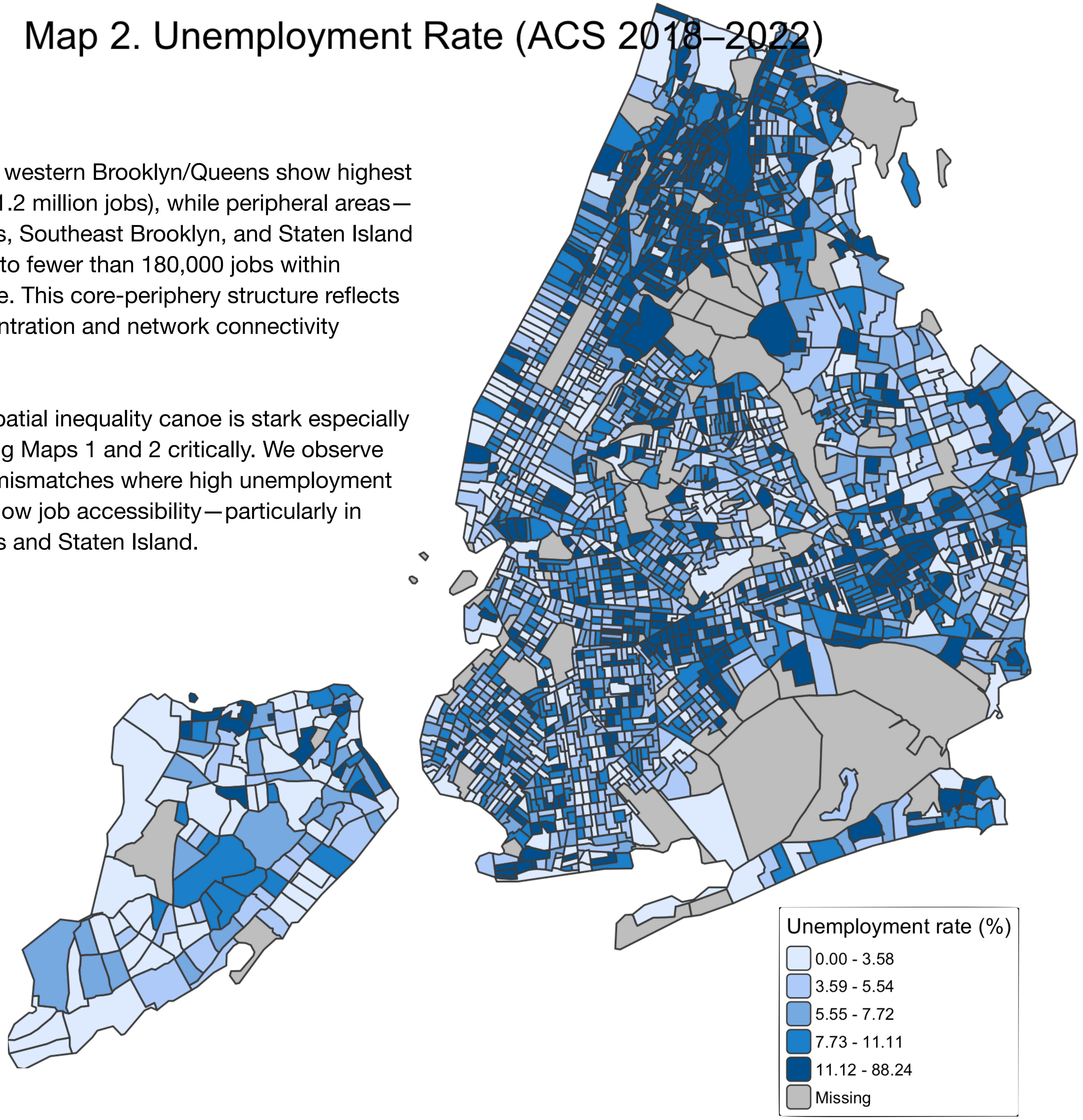
MAP 1: NYC JOBS ACCESS MAP

Source: US Census Bureau, LODES (2021)

Map 2. Unemployment Rate (ACS 2018–2022)

Manhattan and western Brooklyn/Queens show highest accessibility (>1.2 million jobs), while peripheral areas—Eastern Queens, Southeast Brooklyn, and Staten Island—have access to fewer than 180,000 jobs within cycling distance. This core-periphery structure reflects both job concentration and network connectivity constraints.

This extreme spatial inequality canoe is stark especially when comparing Maps 1 and 2 critically. We observe severe spatial mismatches where high unemployment coincides with low job accessibility—particularly in Eastern Queens and Staten Island.



Source: US Census Bureau, ACS (2018 - 2022)

5.1.1 Moran's I

When we conducted Global Moran's I for log employment accessibility, spatial autocorrelation is strong. Moran's I for log employment accessibility is approximately 0.31 (permutation $p< 0.001$). This confirms that high- and low-access tracts are spatially clustered rather than randomly distributed.

Local Indicators of Spatial Association (LISA) further confirm this pattern. Map 5 reveals contiguous high-high clusters in central areas with dense employment and low-low clusters in perisperhal neighbourhoods with limited access.

5.1.2 Baseline Proxy Model: Station Density and Spillovers

We started with CitiBike station density, a commonly used proxy for bike share access a the census tract level. We then used an ordinary least squares (OLS) model regressing log employment accessibility on station density and socioeconomic controls. We found that it yields a positive and statistically significant coefficient on density, but explains relatively little of the overall variation in accessibility ($R^2 \approx 0.11$; Table 2).

Diagnostic tests reveal substantial spatial dependence in this specification. Lagrange Multiplier tests strongly favour a spatial lag model over a spatial error alternative. Estimating a spatial lag model (SLM) substantially improves fit (AIC ≈ 9,167 versus 9,389 for OLS) and yields a large, highly significant spatial autoregressive parameter ($\rho \approx 0.41$). This suggests that accessibility outcomes in one tract are strongly associated with outcomes in neighbouring tracts when access is proxied by station density.

However, Maps 3 and 4 illustrate a key limitation of this proxy-based approach. Station density often diverges sharply from experienced access measured along the street network. In many peripheral or infrastructurally fragmented areas, density overstates effective connectivity, while in some centrally located areas it understates it. As a result, the strong spillovers observed in the density-based SLM may reflect unmodeled spatial structure—such as job concentration and land-use patterns—rather than genuine diffusion of bike-share benefits.

5.1.4 Network-based Access and Structural Exclusion

We replaced measures of station density with network-based measures. This dramatically changed both model fit and interpretation. We introduce two variables:

(i) network distance from each tract's interior point to the nearest Citi Bike station, measured along the street network and capped at 2 miles;

(ii) an explicit indicator for tracts with no routable path to any station. The network OLS model specification explains nearly 90 percent of the variation in employment accessibility ($R^2 \approx 0.89$, TABLE 1), representing a step change relative to the proxy-based specification. Distance to the nearest station is strongly negative and highly significant.

Importantly, the indicator for no station access carries a very large negative coefficient, far exceeding the marginal effect of distance. This result indicates that structural exclusion from the bike-share network is not merely an extreme case of being "far," but a qualitatively distinct condition with disproportionately large consequences for employment accessibility. Maps 2 and 3 visualise these mechanisms. Map 2 shows that approximately seven percent of tracts have no routable access to the Citi Bike network at all, while Map 3 reveals a much larger group of tracts that are technically reachable but located beyond a reasonable cycling distance. Together, these results underscore the importance of distinguishing between continuous distance burdens and categorical exclusion when evaluating transportation equity.

While this network-based OLS model provides clear evidence on the magnitude and nature of experienced access, spatial dependence remains a concern. In the next section, we assess whether remaining spatial autocorrelation reflects outcome spillovers or omitted spatial structure by comparing spatial error and spatial lag specifications.

Importantly, the indicator for no station access carries a very large negative coefficient, which far exceeds the marginal effects of distance.

(Map 6 maps the SEM residuals and shows that residual clustering is substantially reduced relative to the baseline model, supporting this interpretation.)

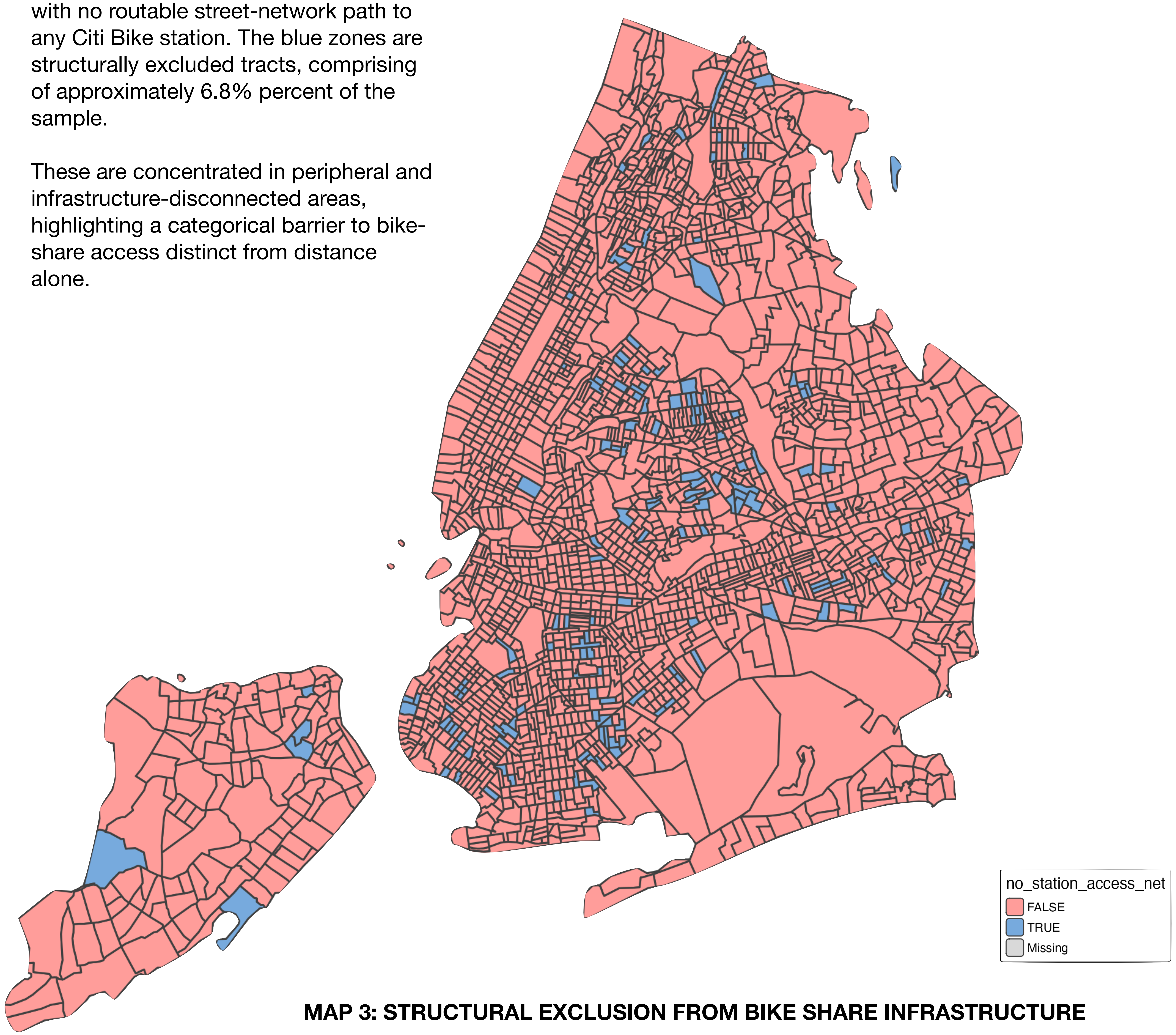
Table 2. Determinants of employment accessibility by bicycle

Ordinary least squares (OLS), spatial lag (SLM), and spatial error (SEM) models relating census-tract employment accessibility to Citi Bike access measures and socioeconomic controls. Density-based models use station density as a proxy for access, while network-based models use street-network distance to the nearest station and an indicator for tracts with no routable station access.

model	term	estimate	std.error	statistic	p.value
OLS (Density)	(Intercept)	11.2422864214111	0.12537342694105	89.6704086001983	0.0
OLS (Density)	station_density	0.0211563357631642	0.0024197864005011	8.74305920505324	4.4199847416879295e-18
OLS (Density)	median_income	5.27004884335296e-06	1.07980948475207e-06	4.88053579614832	1.13401906050582e-06
OLS (Density)	pop_density	1.00627203247805e-05	1.31500017690142e-06	7.65225777268834	2.9355919413223104e-14
SLM (Density)	(Intercept)	6.6562095325979	0.327678621636749	20.3132248889178	0.0
SLM (Density)	station_density	0.0119989782048109	0.0022908866764816	5.23770046244233	1.62589659735701e-07
SLM (Density)	median_income	2.78167422225489e-06	1.01214181977413e-06	2.74830480068066	0.0059904292325057
SLM (Density)	pop_density	6.0933110787822e-06	1.24270424087528e-06	4.90326730879301	9.42555806959788e-07
OLS (Network)	(Intercept)	13.3134993647835	0.058983330160398	225.716305413394	0.0
OLS (Network)	dist_to_station_mi_cap	-0.838882537291652	0.0213508266339901	-39.290400867017	3.59687896248051e-256
OLS (Network)	no_station_access_net	-6.46231341851622	0.064868531067652	-99.6217011878473	0.0
OLS (Network)	median_income	4.06411729541407e-06	3.81785776218012e-07	10.645020188215	7.75171752190328e-26
OLS (Network)	pop_density	2.69350704312843e-06	4.98360436512563e-07	5.40473690483359	7.19311725113294e-08
SEM (Network)	(Intercept)	10.6492269573964	0.192796265861342	55.2356494552401	0.0
SEM (Network)	dist_to_station_mi_cap	-0.200957059506988	0.0334217040816118	-6.01277119252434	1.8237837906554e-09
SEM (Network)	no_station_access_net	-7.15606242031675	0.0470897785941931	-151.966363698283	0.0
SEM (Network)	median_income	-5.85194788223646e-07	2.93020545030312e-07	-1.99711180034529	0.0458130393320351
SEM (Network)	pop_density	-2.58592496188899e-07	3.12104592739582e-07	-0.828544347646518	0.40736229051844
SLM (Network)	(Intercept)	9.53438383219376	0.14411144025012	66.1597983869002	0.0
SLM (Network)	dist_to_station_mi_cap	-0.607588947142419	0.0199731993352081	-30.4202114516217	0.0
SLM (Network)	no_station_access_net	-6.47751455436526	0.0559310596238712	-115.812476965852	0.0
SLM (Network)	median_income	2.29340007584275e-06	3.29837864776604e-07	6.95311339526179	3.5731417824535998e-12
SLM (Network)	pop_density	1.10390389680761e-06	4.29000243120701e-07	2.57320109838965	0.0100762641102192

This binary map indicates census tracts with no routable street-network path to any Citi Bike station. The blue zones are structurally excluded tracts, comprising of approximately 6.8% percent of the sample.

These are concentrated in peripheral and infrastructure-disconnected areas, highlighting a categorical barrier to bike-share access distinct from distance alone.



MAP 3: STRUCTURAL EXCLUSION FROM BIKE SHARE INFRASTRUCTURE

5.2 NETWORK ACCESS DISTRIBUTION

5.2.1 Baseline proxy model: station density and apparent spillovers

We first estimate a baseline model in which Citi Bike access is proxied by station density within each tract. Ordinary least squares (OLS) estimates show a positive and statistically significant association between station density and employment accessibility. But the model explains little of the overall variation ($R^2 \approx 0.11$).

When we conducted Lagrange Multiplier diagnostics, the results strongly favour a spatial lag specification over a spatial error alternative. The corresponding spatial lag model (SLM) yields a large and highly significant spatial autoregressive parameter ($\rho \approx 0.41$) and improves model fit substantially (AIC ≈ 9,167 versus 9,389 for OLS). This pattern suggests strong outcome spillovers when access is measured using station density.

However, station density overstates access in some neighbourhoods while understating it in others (see Table 4: Model Comparison and Effect Decomposition).

This is especially pronounced in areas where stations are present but poorly connected by the street network. This raises concerns that the apparent spillovers in the density model may reflect unmodeled spatial structure rather than true diffusion of bike-share benefits.

Assumption: Station density is treated as an exogenous proxy for access, despite not accounting for network connectivity or routability.

5.2.2. STRUCTURAL EXCLUSION

We notice the 157 tracts (6.8%) with no routable path to any Citi Bike stations. These are not merely distant but topologically disconnected. An additional 735 tracts (31.6%) lie beyond 2 miles network distance. This represents a sizeable practical exclusion given typical cycling ranges.

Among reachable tracts, the median distance is 0.50 miles. But the distribution has a heavy right tail extending to 14 miles, indicating severe access disparities even within the connected network.

5.2.3. Marginal Effects and Spillovers Analysis

Table 3: SLM vs SEM Network Comparison

Model Type	Variable	Direct Effect	Indirect (Spillover)	Total Effect	% Spillover
SLM-Density (ρ=0.41)					
	Station density	0.0124	0.0079	0.0203	39%
	Median income	0.0000029	0.0000018	0.0000047	38%
	Pop density	0.0000063	0.0000040	0.0000103	39%
SLM-Network (ρ=0.30)					
	Distance to station (per mile)	-0.619***	-0.252***	-0.871***	29%
	No station access (binary)	-6.595***	-2.688***	-9.283***	29%
	Median income	0.0000023	0.0000010	0.0000033	30%
	Pop density	0.0000011	0.0000005	0.0000016	31%
SEM-Network (λ=0.98)					
	Distance to station (per mile)	-0.201***	—	-0.201***	—
	No station access (binary)	-7.156***	—	-7.156***	—
	Median income	-0.0000006*	—	-0.0000006	—
	Pop density	-0.0000003 (ns)	—	-0.0000003	—

Reading the two models, we can infer that both models tell different stories. SEM is preferred for statistical inference, while spatial lag models are retained for substantive interpretation of spillovers.

Looking at SLM, we notice that spillovers matter.
- A resident living 2 miles from a station, loses access to ~75% of jobs (e^-1.74 ≈ 0.18)
- A resident with no station access loses access to 99.96% of jobs
- 30% of these effects spill over to neighbours—improving one tract helps adjacent tracts
- This suggests strategic cluster placement could generate cascading benefits

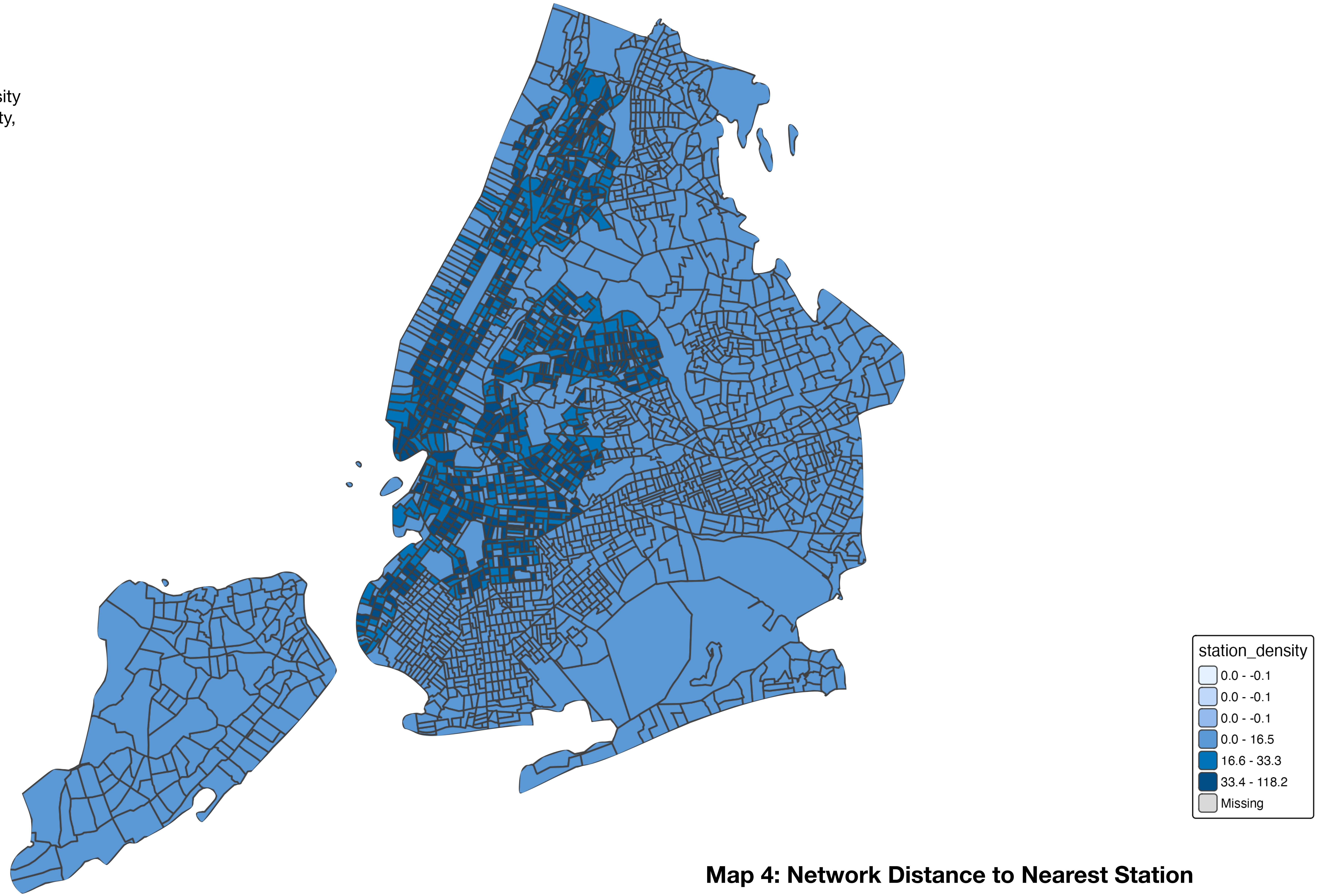
Looking at SEM, we notice that geography matters more.
- The λ=0.98 is extraordinarily high, suggesting perfect spatial error correlation
- This could mean that unobserved spatial factors (bridges, job centers, subway lines) drive most variations
- CitiBike effects shrink dramatically (-0.84→-0.20 for distance) when spatial structure is properly modeled
- This suggests that "spillovers" in SLM were likely spurious—neighbouring tracts are similar due to shared geography, not because bike access diffuses

Table 4: Model Comparison and Effect Decomposition

model	N	R2	AIC	rho	lambda
OLS (Density)	2201	0.104587203996268	9388.58849951877	nan	nan
SLM (Density)	2201	nan	9166.92729277713	0.410218846785426	nan
OLS (Network)	2201	0.886559839058814	4843.3019091377	nan	nan
SEM (Network)	2201	nan	2464.80597371105	nan	0.979375102918646
SLM (Network)	2201	nan	4204.91120076962	0.302207419611556	nan

Map 4: Citi Bike station density

Station density by census tract, used as a baseline proxy for bike-share access. Comparison with Map 3 illustrates that density often misrepresents experienced connectivity, particularly in areas where stations exist but are poorly connected by the street network.



Map 4: Network Distance to Nearest Station

Source: CitiBike Stations (2025)

5.3 Spatial dependence: Spatial error vs spatial lag

Spatial diagnostics for the network model differ sharply from those of the density model. Robust LM tests strongly favour a spatial error specification, while evidence for a spatial lag is substantially weaker. The spatial error model (SEM) dramatically improves model fit (AIC ≈ 2,465) and yields an extremely large and significant spatial error parameter ($\lambda \approx$ 0.98).

This pattern suggests that, once experienced access is modeled directly, remaining spatial dependence primarily reflects omitted spatial structure. This could be due to the underlying geography of jobs, land use patterns, and historical developments, rather than diffusion of bike-share benefits.

5.3.1. Spillovers among connected neighborhoods

Although the SEM is preferred for inference, estimating a spatial lag model for the network specification provides insight into spillover effects. The spatial autoregressive parameter remains positive and significant ($\rho \approx 0.30$), and impact decomposition reveals meaningful indirect effects. For both network distance and structural exclusion, approximately 30–40 percent of the total effect operates through neighbouring tracts.

Crucially, these spillovers are conditional on connectivity. Tracts that are structurally excluded from the network do not benefit from neighbouring access improvements, limiting the extent to which bike-share expansion can diffuse benefits without addressing connectivity gaps directly.

Map 7 underscores this point by contrasting standardised station density with standardised network distance. In many peripheral neighbourhoods, density-based measures substantially overstate access relative to experienced network connectivity, masking both exclusion and constrained spillovers.

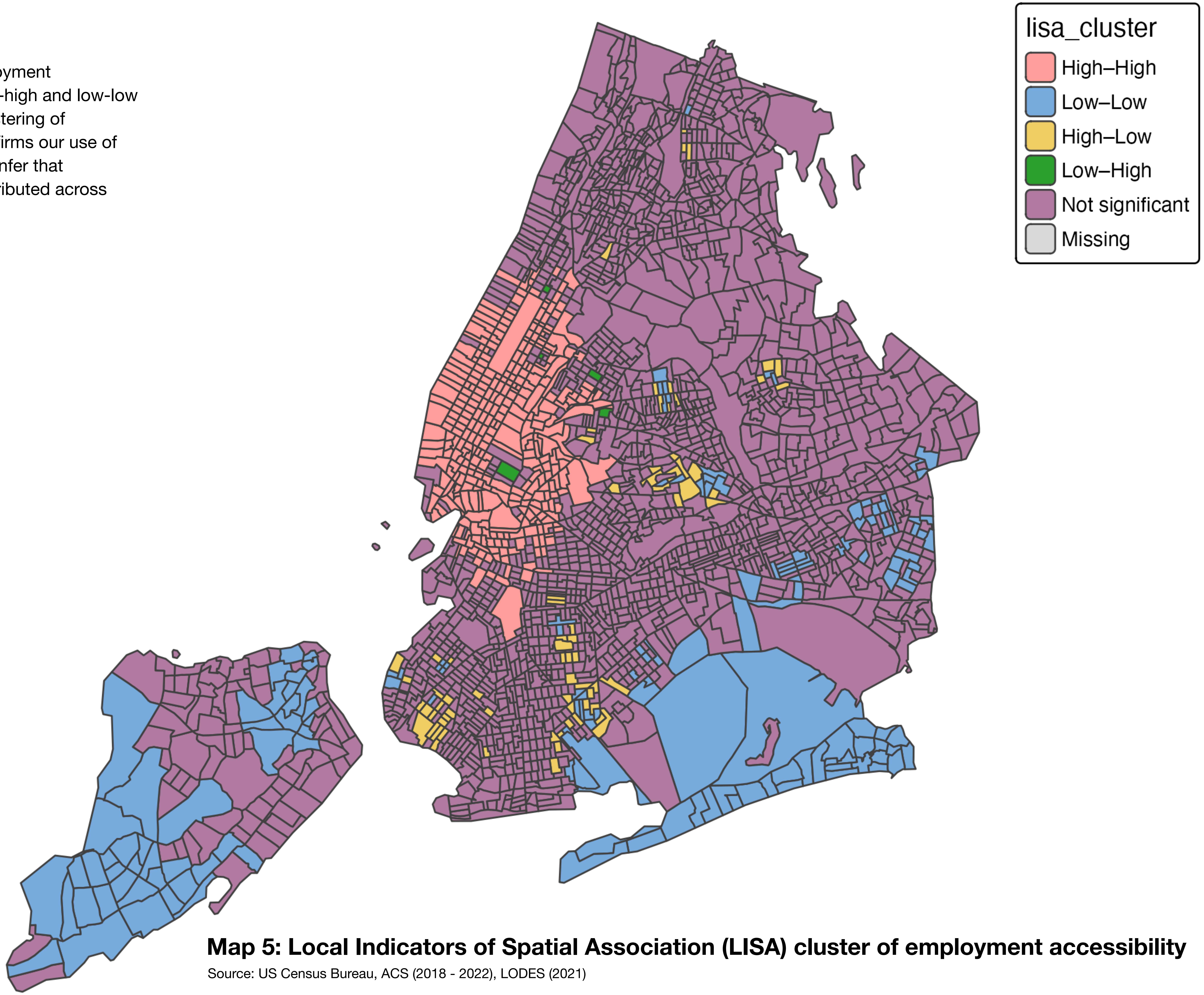
5.3.2 Summary of findings

Three core results emerge. First, employment accessibility in New York City is deeply unequal and spatially clustered. Second, density-based measures of bike-share access conflate experienced access with spatial structure, producing misleading evidence of spillovers. Third, network-based measures reveal that structural exclusion is a dominant mechanism shaping who benefits from bike-share systems. Spillovers do exist, but they accrue primarily among neighbourhoods that are already connected to the network.

Table 4: Interpretation of Models and Effects

Finding	What It Means	Policy Implication
Distance Effect (SLM)	Each additional mile to nearest CitiBike station reduces log job access by 0.87 (approximately 58% fewer jobs)	Station placement matters enormously
Structural Exclusion (SLM)	Tracts with zero network access have 9.28 lower log job access (approximately 99.96% fewer accessible jobs)	Disconnected areas face catastrophic disadvantage
Spillover Share: 29-30%	Nearly 1/3 of CitiBike's effect operates through neighboring tracts	Benefits cascade spatially BUT only if connected
SEM Lambda = 0.98	Massive spatial error correlation	Most "effect" is omitted spatial structure, not true causal impact

The Local Moran's I for log employment accessibility shows how the high-high and low-low clusters reveal strong spatial clustering of accessibility outcomes. This confirms our use of spatial models. We can strongly infer that accessibility is not randomly distributed across NYC.

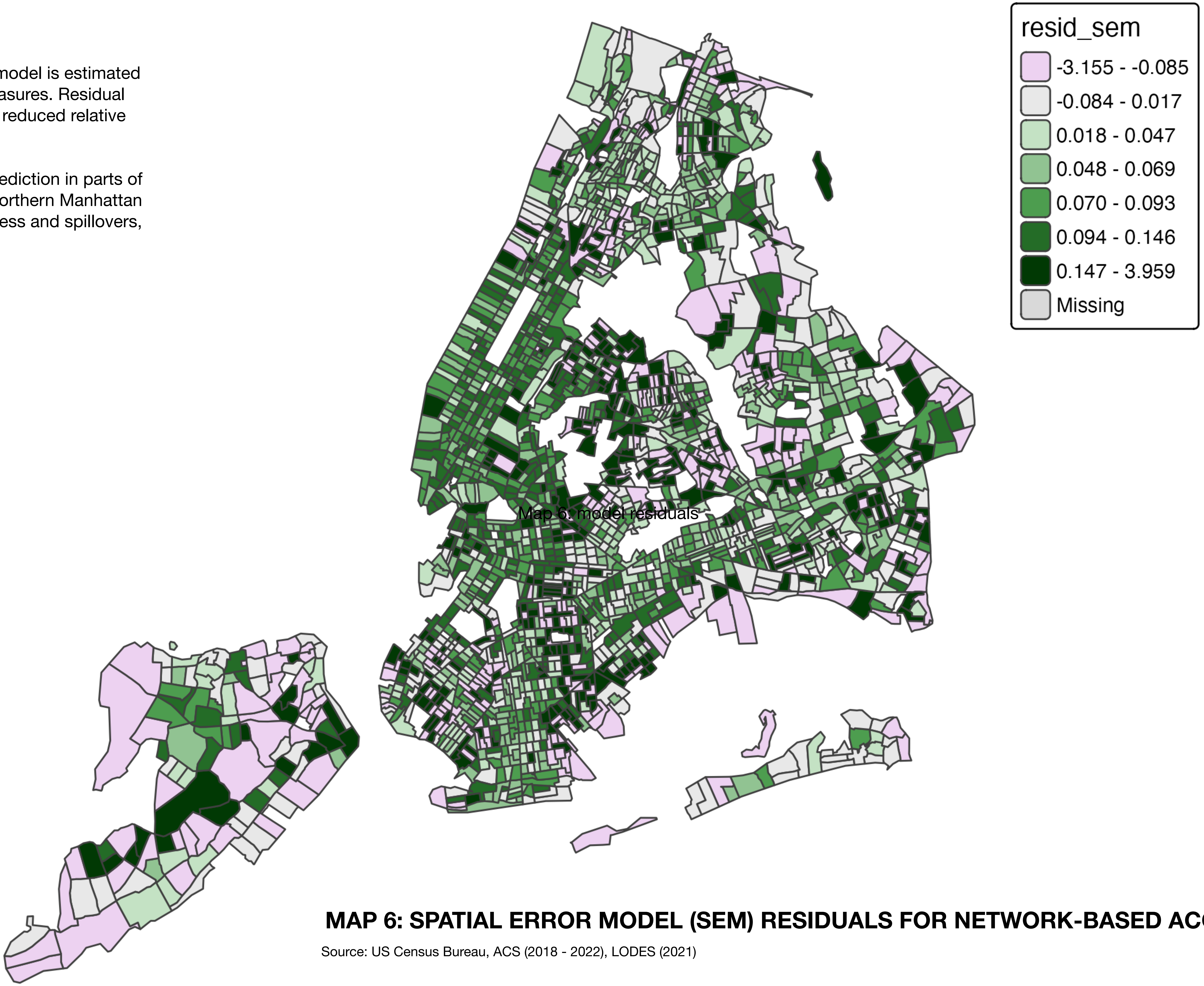


Map 5: Local Indicators of Spatial Association (LISA) cluster of employment accessibility

Source: US Census Bureau, ACS (2018 - 2022), LODES (2021)

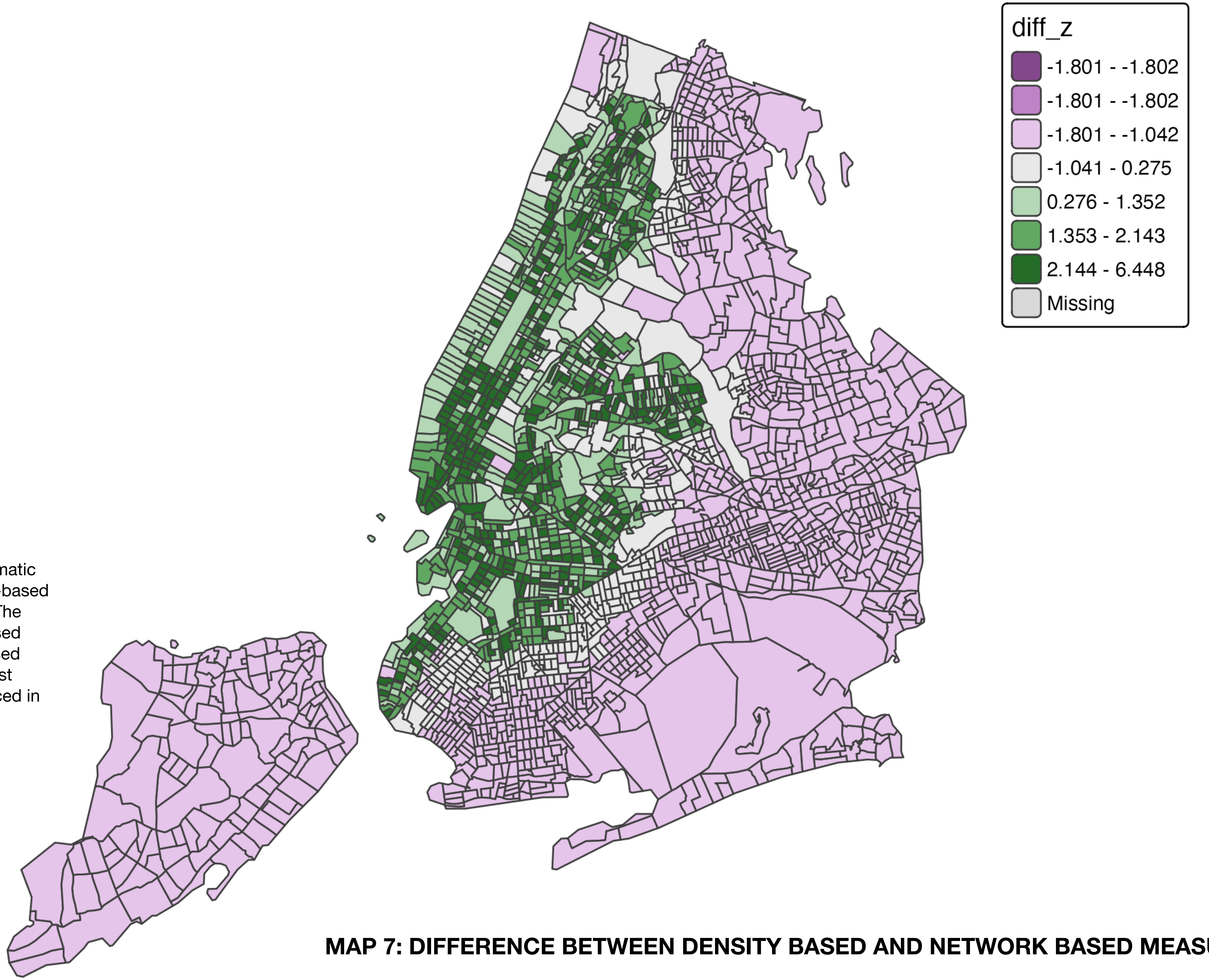
Residuals from the spatial error model is estimated using network based access measures. Residual spatial clustering is substantially reduced relative to baseline models.

This reveals systematic under-prediction in parts of Queens and over-prediction in Northern Manhattan after accounting for network access and spillovers, suggesting additional barriers.



MAP 6: SPATIAL ERROR MODEL (SEM) RESIDUALS FOR NETWORK-BASED ACCESSIBILITY

Source: US Census Bureau, ACS (2018 - 2022), LODES (2021)



This map highlights the systematic discrepancies between proxy-based and network - based values. The difference between standardised station density and standardised network distance to the nearest station, is especially pronounced in periphery neighbourhoods.

MAP 7: DIFFERENCE BETWEEN DENSITY BASED AND NETWORK BASED MEASURES

5. Synthesis: Network Topology, Spatial Spillovers, and the Geography of Bike-Share Equity

This analysis questions the impact of bike-share infrastructure on employment accessibility in New York City. Our models show how employment accessibility is fundamentally conditioned by network topology and spatial structure, challenging conventional proximity-based approaches to transportation equity analysis.

Our findings demonstrate three critical insights for understanding micro-mobility's role in urban labor market access.

First, network-based measures expose hidden inequities that traditional density metrics mask. While station density explains only 11% of variation in job accessibility, network-based measures capture 89%. This revealing 6.8% of NYC census tracts are topologically disconnected from the CitiBike system entirely. An additional 31.6% is located beyond practical cycling distance. This study shows how access to bike-share systems cannot be understood solely in terms of station presence or density. When access is measured as experienced connectivity along the street network, large disparities emerge that are obscured by proxy-based measures. Most notably, a non-trivial share of census tracts in New York City are structurally excluded from the Citi Bike network, lacking any routable path to a station. For these tracts, marginal changes in distance or density are irrelevant: exclusion constitutes a categorical barrier rather than a continuous disadvantage.

Secondly, we find that spatial spillovers are conditional but not universal. Spatial lag models indicate substantial indirect effects (29-30% of total impact), suggesting that bike-share improvements in one tract diffuse across and benefit neighbouring areas — but only among areas already connected to the system.

Structurally excluded tracts face a catastrophic accessibility penalty (7.2 log-units, equating to 99.9% fewer accessible jobs) and cannot capture spillover benefits even if stations are added to adjacent areas. This creates a "spillover shadow"—zones where spatial multiplier effects cannot reach due to topological barriers. This challenges optimistic narratives which emphasise network effects as a mechanism for equity improvements.

Third, the contrast between station density and network-based access reflects methodological issue in transportation equity research. The extraordinarily high spatial error correlation (λ=0.98) in our SEM specification indicates that unobserved spatial factors still dominate the accessibility landscape. Granted, our models used Queen contiguity weights and were temporally cross-sectional. However, density measures implicitly assume proximity translates into usability, overlooking barriers such as discontinuous street networks, limited transit networks, or infrastructural segmentation.

When these structural factors are properly controlled, the marginal effect of bike-share distance shrinks by 76% (from -0.84 to -0.20), suggesting that CitiBike stations may follow advantage rather than create it.

The analysis focused on potential accessibility rather than realised behaviours (i.e. accessibility should interpreted as an opportunity measure rather than a prediction of commuting behaviour). Given the limitations, these findings challenge optimistic narratives about bike-share's transformative potential and indicates that infrastructure placement may be endogenous to existing accessibility patterns.

6. IMPLICATIONS

Our network-topology approach shows how GIS methods could capture mobility constraints in urban settings. The difference between density- and network-based models for transportation equity underscore the need to consider routing constraints beyond distance - based proximity.

The structural exclusion zones identified were shown to be disconnected from both direct access and spatial spillovers, suggesting counter-intuitively that conventional infrastructure expansion may depend rather than alleviate spatial inequality.

We posit that simply adding stations to under-served areas without addressing network connectivity, such as protected bike lanes, bridge access, safe routes, may yield minimal improvements to accessibility. Of course this rides on the limitations of the models we built, primarily on NYC street centrelines with assumptions that New York residents who bike take only biking routes.

Regardless, our results indicate that 157 census tracts containing approximately 6.8% of New York City's population currently exist outside of the reach of existing bike-share benefits. This is true even just taking into account station proximity in Euclidean space.

For policymakers looking to pursue Bikeshare as an equity intervention, our findings note that micro-mobilty infrastructure cannot overcome macro-level spatial disconnection. The communities who most need improved job access are in peripheral areas with high unemployment and low car ownership. They are also precisely those excluded from bike-share networks and their spillover benefits. Effective bike-sharing focused on transportation equity must prioritise network connectivity as much, if not more than station count and placements.

Broader transportation options notwithstanding, the focus on linking disconnected communities to the broader transportation network must also rethink optimising coverage within already-connected areas. Without strategic intervention, bike-share systems risk becoming amenities which reinforce rather than redistribute accessibility advantages.

APPENDIX:

	CONTENT TABLE OF MAPS	
1	Jobs Access Map (log)	Inequality + Clustering
2	No Station Access (net)	Structural Exclusion
3	Distance to Station (miles)	Experienced Access
4	Station Density	Proxy
5	LISA Log	Inequality + Clustering
6	SEM residuals (network model)	Omitted Structure vs Diffusion
7	z (Station Density) - z (Distance to Station)	Problematising Density

APPENDIX:

Jobs accessible within 30-minute bike ride (≈5 miles)

Statistic	No. Of Jobs
Min	0
1st Quartile	219,170
Median	387,788
Mean	683,825
3rd Quartile	787,288
Max	2,546,141
Missing tracts	3

Log-transformed DV

(log_jobs_access = log1p(jobs_access_30min))

Statistic	No. Of Jobs
Min	0.00
1st Quartile	12.30
Median	12.87
Mean	12.50
3rd Quartile	13.58
Max	14.75
Missing tracts	3

APPENDIX: Network and Accessibility Diagnostics

Category	Metric	Value	Interpretation
Study Geography	Census tracts (total)	2,327	Full NYC tract universe used in analysis
Bike Network	Number of nodes	79,217	Bikeable street intersections and endpoints
	Number of edges	122,053	Bikeable street segments
Network Connectivity	Tracts with no reachable station	157	Structurally disconnected from Citi Bike network
	Share with no access	6.76%	Reflects geographic and infrastructure gaps
Capping Diagnostics	Reachable tracts beyond 2 mi	735	Top-coded to reduce leverage
Spatial Weights	Weight type	Queen contiguity (row-standardised)	Appropriate for irregular tract geometry
	Isolates	3	Tracts with no contiguous neighbours
	Subgraphs	5	Due to water bodies and geography

APPENDIX:

Distance to Nearest Station (Uncapped)	Minimum	0.0 ft	Tracts containing a station
	1st quartile	509.8 ft (0.10 mi)	High-access neighbourhoods
	Median	2,646.4 ft (0.50 mi)	Typical tract
	Mean	10,021.3 ft (1.90 mi)	Right-skewed distribution
	3rd quartile	15,165.0 ft (2.87 mi)	Poorly served but connected
	Maximum	73,706.6 ft (13.96 mi)	Extreme outliers

Distance to Nearest Station (Capped at 2 mi)	Minimum	0.00 mi	Station-present tracts
	1st quartile	0.10 mi	Close proximity
	Median	0.81 mi	Typical accessibility
	Mean	1.01 mi	Stabilised by capping
	3rd quartile	2.00 mi	Cap threshold
	Maximum	2.00 mi	Top-coded