

What Drives COVID-19 Tokyo Metro Ridership Changes?

Xi Lin^{*1}, Giuseppe Vizzari², Claudio Feliciani¹, Daichi Yanagisawa¹, and Katsuhiro Nishinari¹

¹School of Engineering, The University of Tokyo

²Department of Informatics, Systems and Communication, University of Milano-Bicocca

Abstract This study explains station-level annual ridership change in Tokyo (2020–2023) relative to 2019 using ordinary least squares (OLS) and geographically weighted regression (GWR). OLS results show that higher pre-pandemic ridership is linked to larger COVID-19-related declines, while road density supports recovery. GWR further reveals strong spatial heterogeneity, with the effects of variables varying across the network and over time. Policy implications suggest prioritizing resilience planning for high demand hubs and investing in mixed-use, accessible, well-connected station areas.

Keywords COVID-19; Metro ridership; Built environment; Network centrality; OLS; GWR

Introduction

The COVID-19 pandemic caused an unprecedented shock to metro ridership, and recovery has been uneven across cities, especially in Tokyo, where daily mobility is highly metro-dependent. Many research on metro ridership change still focuses on short periods and a limited set of factors, leaving unclear how different variables jointly shape station-level declines and recovery under rare but severe disruptions like COVID-19. To address these gaps, we analyze station-level ridership change in Tokyo (2020–2023 vs 2019) by jointly modeling three dimensions: pre-pandemic demand, built environment, and network centrality.

Variables and Method

Table 1 defines all variables and reports their descriptive statistics (using 2020 data as an example).

Table 1: Descriptive statistics of variables.

Variable	Description	Mean	SD
Pre-pandemic demand			
2019 annual ridership (PV)	Ridership of Tokyo metro in 2019 (thousands).	36613	45698
Built environment (Density/Intensity)			
POI total (POI)	Number of POIs within the 1 km buffer.	7256.96	6108.77
Production and Service (PS)	Number of Production and Service within buffer.	2427.32	1759.72
Consumption and Leisure (CL)	Number of Consumption and Leisure within buffer.	4136.24	3925.67
Mobility and Outdoor (MO)	Number of Mobility and Outdoor within buffer.	693.39	526.06
Built environment (Diversity)			
Degree of land use mix (LUM)	Degree of mixed land use in the catchment area.	0.8023	0.0387
Built environment (Destination accessibility)			
Distance to Tokyo station (D-CBD)	Distance from stations to Tokyo Station (km).	5.720	3.812
Built environment (Design)			
Road density (RD)	Road density within the 1 km buffer (km/km ²).	31.14	5.50
Average street segment length (SL)	Mean length of road segments within the buffer (m).	103.24	23.031
Centrality			
Strength (ST)	Strength degree centrality of station.	0.207	0.165
Betweenness (BE)	Betweenness centrality of station.	0.139	0.150
Eigenvector (EV)	Eigenvector centrality of station.	0.062	0.141
PageRank (PR)	PageRank centrality of station.	0.208	0.149

Figure 1 summarizes the framework: station-level annual ridership change is modeled using three groups of variables: pre-pandemic demand, 4D built-environment measures (one from each dimension), and centrality (one of four measures). We first screen pairwise correlations to remove collinear variables, then estimate candidate OLS regression specifications and select the final global model using AICc and adjusted R^2 . We test spatial autocorrelation using Moran's I and P, and then apply GWR, comparing fit with OLS and mapping local coefficients to assess spatial heterogeneity.

Results

In comparing all feasible OLS specifications under the framework in Figure 1, we selected a final global OLS model including PV, PS, RD, LUM, D-CBD, and BE. The OLS result is shown in Table 2.

*Email of the corresponding author: lin-xi@g.ecc.u-tokyo.ac.jp

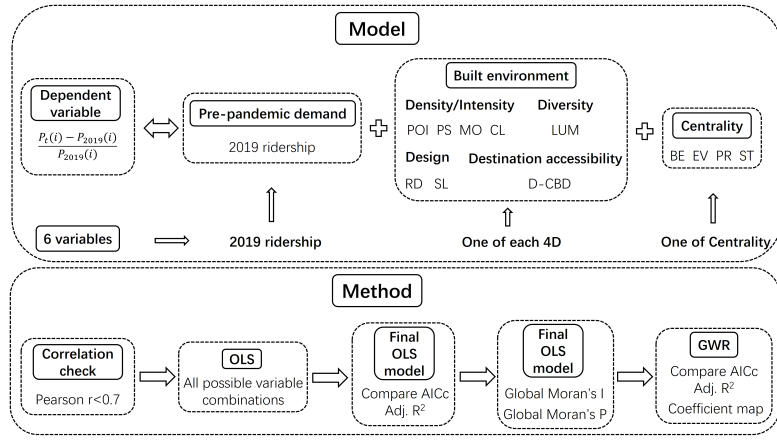


Figure 1: Methodological framework.

Table 2: OLS Results.

Y (year vs 2019)	PV	PS	LUM	D-CBD	RD	BE
2020	-0.0147**	-0.0509***	-0.1617	0.0021	0.0038***	0.0339
2021	-0.0201***	-0.0310**	0.3129*	0.0038*	0.0032***	0.0495
2022	-0.0198***	-0.0175	0.5304***	0.0021	0.0030***	0.0588*
2023	-0.0235***	-0.0109	0.4906**	0.0007	0.0029***	0.0427

Notes. *, **, *** indicate $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

OLS residuals exhibit significant positive spatial autocorrelation in all years (Moran's $I = 0.090-0.225$, $p \leq 0.002$), indicating clustered errors and motivating the use of GWR to capture spatial heterogeneity.

We estimated station-level GWR coefficients for 2020-2023, revealing pronounced spatial heterogeneity and Figure 2 presents results of the local GWR coefficients for selected variables and years.

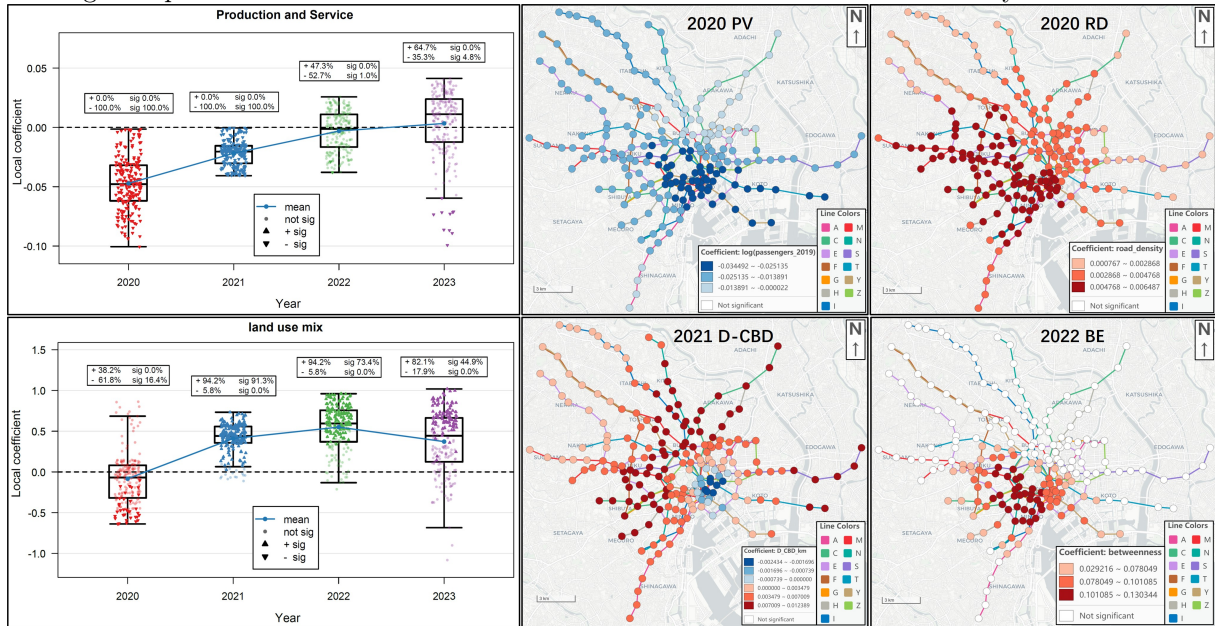


Figure 2: Results of the local GWR coefficients.

Pre-pandemic ridership shows a consistently negative and significant effect: stations that were busier before COVID-19 experienced larger pandemic-related ridership declines, with the strongest losses concentrated in the urban core and the southeast corridor. Road density is consistently positive and statistically significant, indicating that stations in denser street networks retained more ridership during the pandemic, likely because fine grained streets improve walk access and support short local trips, with the strongest positive effects extending from the city center toward the southwest. Production and service intensity is negative in 2020 and 2021 and weakens thereafter, suggesting that work oriented areas were hit hardest early on and then stabilized. By contrast, land use mix becomes positive from 2021 onward, indicating faster recovery in more mixed use station areas as local and discretionary activity returns. Two effects appear only in specific years. Distance to Tokyo Station is mostly positive in 2021, indicating smaller losses farther from the center, while a small cluster around Tokyo Station shows the opposite pattern, consistent with a localized core reversal. Betweenness centrality is mainly positive in 2022, especially from the center toward the southwest, suggesting that key transfer stations regained demand more strongly as network-wide travel resumed. Overall, these spatially varying relationships support two priorities: resilience planning for the highest demand hubs and investment in accessible, well connected, mixed use station areas to support recovery across the network.