

Supplementary Materials

Near-Complete Service Breadth, Uneven Local Travel: Mapping Provision–Travel Alignment Across Three Eastern Chinese Cities

Decun Wu and He Yang

Contents and conventions

This document contains the full estimates summarized in the main text. Table S1 decomposes the robustness specifications by city; Table S2 combines the pooled interaction models with the group-specific gradients; Table S3 presents descriptive rank-divergence and preferred-model residual spatial diagnostics; Table S4 re-derives the provision–travel classification under alternative cut points and network-based accessibility; Table S5 shows the mode composition of classified short trips; Table S6 compares the city-level null results across variation available for estimation and outcomes; Table S7 consolidates the quantity-sensitivity and typology-stability audit; Table S8 presents leave-one-city-out cross-validation of the gradient-boosting model; Table S9 summarizes the park/scenic composition audit; Table S10 presents the source-taxonomy, hierarchy-tag and strict-taxonomy sensitivity analyses; Table S11 presents the focused park/scenic functional-composition and catchment sensitivities; and Figure S1 shows the Shapley additive explanations (SHAP) attribution profiles summarized in the main text.

Unless stated otherwise, all models estimate

$$Y_{iz} = \beta \text{Breadth}_{iz} + \mathbf{X}'_{iz}\boldsymbol{\gamma} + \mu_z + \varepsilon_{iz},$$

where $\mathbf{X}$ contains log population, mean resident age, the logarithm of average revenue per user (ARPU), the local-registration share, log distance to the nearest bus stop, a metro proximity indicator, the jobs–housing balance score and log distance to the primary center, and μ_z are district fixed effects. Standard errors are clustered by district. Wild cluster bootstrap p -values (p_b) use Rademacher weights with the null imposed and CR1 studentization (Cameron, Gelbach and Miller, 2008). The focal pooled and joint tests use 9999 seeded draws, while the headline single-city M5/M6 tests enumerate all 2^G sign patterns. The number of district clusters is 30 pooled, 9 in Shenzhen, 11 in Nanjing and 10 in Xuzhou, except where a sample restriction reduces it.

Table S1: Robustness estimates for the service-breadth–short-trip association, pooled and by city.

Specification	Pooled			Shenzhen			Nanjing			Xuzhou		
	$\hat{\beta}$	SE	p_b	$\hat{\beta}$	SE	p_b	$\hat{\beta}$	SE	p_b	$\hat{\beta}$	SE	p_b
Outcome: share ≤ 1 km	0.197	(0.088)	0.046	0.441	(0.304)	0.152	0.497	(0.075)	0.015	−0.025	(0.076)	0.750
Outcome: share ≤ 3 km	0.138	(0.165)	0.420	0.380	(0.591)	0.551	0.558	(0.189)	0.011	−0.190	(0.183)	0.320
Outcome: share > 5 km	0.181	(0.192)	0.389	0.179	(0.517)	0.750	−0.303	(0.229)	0.251	0.474	(0.221)	0.111
Supply radius 800 m	0.170	(0.106)	0.131	0.498	(0.396)	0.254	0.464	(0.120)	0.009	−0.128	(0.114)	0.287
Supply radius 1200 m	0.389	(0.114)	0.006	0.728	(0.594)	0.234	0.645	(0.150)	0.017	0.165	(0.130)	0.246
Cells with ≥ 500 trips	0.531	(0.161)	0.005	0.829	(0.464)	0.098	0.955	(0.240)	0.019	−0.000	(0.197)	1.000
Within 30 km of the primary center	0.598	(0.260)	0.053	1.383	(0.356)	0.055	0.756	(0.256)	0.066	−0.047	(0.277)	0.938
Trip-weighted WLS	0.960	(0.299)	0.004	0.020	(0.431)	0.961	1.669	(0.351)	0.012	0.250	(0.410)	0.553

Note: All models include district fixed effects and full covariates; SEs are district-clustered and p_b is a null-imposed wild cluster bootstrap p -value. Every pooled row uses 9999 seeded draws; every single-city row uses exact enumeration of all available sign patterns (normally 512/2048/1024; the 30-km restriction leaves 512/512/64). Exact two-sided tails use a scale-aware numerical tolerance so that sign-symmetric boundary draws are counted consistently.

Table S2: Service-breadth interactions by city and neighborhood context, with demographic-group-specific gradients.

Term	$\hat{\beta}$	SE	p
<i>Panel A: breadth $\times$ city (reference: Xuzhou)</i>			
Service breadth (reference group)	0.271	(0.142)	0.056
$\times$ Nanjing	0.154	(0.249)	0.536
$\times$ Shenzhen	−0.295	(0.377)	0.433
<i>Panel B: breadth $\times$ ARPU tercile (reference: low)</i>			
Service breadth (reference group)	0.409	(0.175)	0.019
$\times$ middle ARPU tercile	−0.141	(0.166)	0.396
$\times$ high ARPU tercile	−0.218	(0.205)	0.287
<i>Panel C: breadth $\times$ age tercile (reference: young)</i>			
Service breadth (reference group)	0.849	(0.150)	<0.001
$\times$ middle age tercile	−0.512	(0.177)	0.004
$\times$ old age tercile	−0.836	(0.213)	<0.001

Group	Two-city pooled		Shenzhen		Xuzhou	
	$\hat{\beta}$	p	$\hat{\beta}$	p	$\hat{\beta}$	p
<i>Panel D: age-band gradients</i>						
Ages 19–34	1.135	0.038	0.537	0.346	0.475	0.371
Ages 35–59	0.539	0.137	0.380	0.394	−0.396	0.370
Ages 60+	0.909	0.040	0.649	0.158	−0.314	0.487
<i>Panel E: ARPU-band gradients</i>						
Low ARPU	1.042	0.014	0.526	0.353	−0.089	0.850
Middle ARPU	0.881	0.042	0.531	0.308	−0.225	0.671
High ARPU	0.695	0.168	0.344	0.498	−0.008	0.989

Note: Panels A–C are pooled interaction models ($N = 11,027$) with district fixed effects, full covariates and common covariate slopes. Panels D–E are trip-weighted, group-specific models with district fixed effects and full covariates (19 clusters pooled; 9–10 by city). SEs are district-clustered and p -values use the cluster-robust normal approximation.

Table S3: Descriptive rank-divergence and preferred-model residual spatial diagnostics.*Panel A: descriptive rank divergence under alternative neighbor definitions*

k	City	Moran's I	z	p	HH (%)	LL (%)
4	Shenzhen	0.517	34.3	0.001	12.5	9.8
4	Nanjing	0.563	51.8	0.001	13.0	10.6
4	Xuzhou	0.532	56.3	0.001	14.2	9.6
8 [†]	Shenzhen	0.453	43.0	0.001	16.5	12.9
8 [†]	Nanjing	0.475	63.2	0.001	16.2	14.5
8 [†]	Xuzhou	0.390	57.7	0.001	15.7	12.5
12	Shenzhen	0.401	45.6	0.001	18.6	16.1
12	Nanjing	0.401	63.9	0.001	17.0	16.5
12	Xuzhou	0.291	54.3	0.001	15.8	13.6

Panel B: global Moran's I for preferred-model residuals

City	$k = 4$		$k = 8$		$k = 12$	
	Moran's I	p_{perm}	Moran's I	p_{perm}	Moran's I	p_{perm}
Shenzhen	0.395	0.001	0.328	0.001	0.280	0.001
Nanjing	0.527	0.001	0.462	0.001	0.411	0.001
Xuzhou	0.578	0.001	0.508	0.001	0.449	0.001

Panel C: BH-FDR local-cluster shares for descriptive rank divergence ($k = 8$)

City	BH-FDR HH (%)	BH-FDR LL (%)
Shenzhen	9.15	6.58
Nanjing	9.61	9.04
Xuzhou	10.36	6.27

Note: Panel A defines rank divergence as the within-city service-breadth percentile rank minus the short-trip-share rank. Row-standardized k -nearest-neighbor weights use grid centroids and 999 permutations; HH and LL are pointwise-significant high-high and low-low shares. [†] $k = 8$ is the main-text specification. Panel B residuals come from pooled M5 ($N = 11,027$) and diagnose remaining spatial structure; they do not select a model. Panel C applies Benjamini-Hochberg FDR control to all $k = 8$ local permutation p -values within each city; the corresponding raw HH/LL shares appear in Panel A.

Table S4: Provision-travel classes under alternative cut points and accessibility measures.

Class	N	Share (%)	Mean breadth	Mean short-trip share (%)
<i>0.5 percentile-rank cut, Euclidean (focal) ($N = 11,047$)</i>				
Aligned high	3895	35.3	7.63	33.5
High-breadth/low-local-travel	2136	19.3	7.49	14.6
Low-breadth/high-local-travel	1631	14.8	3.65	25.0
Aligned low	3385	30.6	3.20	7.9
<i>Tercile cut, Euclidean ($N = 4803$)</i>				
Aligned high	1826	38.0	7.80	35.0
High-breadth/low-local-travel	534	11.1	7.58	4.3
Low-breadth/high-local-travel	535	11.1	3.03	33.8
Aligned low	1908	39.7	2.61	6.6
<i>0.5 percentile-rank cut, network-based ($N = 11,047$)</i>				
Aligned high	3501	31.7	6.08	32.9
High-breadth/low-local-travel	2021	18.3	5.34	12.0
Low-breadth/high-local-travel	2025	18.3	1.18	27.6
Aligned low	3500	31.7	0.97	9.5

Note: The 0.5 cut defines high as within-city percentile rank ≥ 0.5 and low as rank < 0.5 for each axis, with tied ranks averaged. The tercile cut compares the top and bottom within-city terciles and drops the middle. The network measure replaces the 1-km Euclidean catchment with 1-km shortest paths on the OpenStreetMap pedestrian graph.

Table S5: Mode composition of classified home-based trips ≤ 2 km, averaged across grids (%).

Mode	Two-city pooled	Shenzhen	Xuzhou
Walking	33.58	45.64	26.15
Pedal cycling	24.76	23.07	25.80
E-bike	16.12	11.91	18.71
Cycle/e-bike, unresolved	9.29	9.40	9.22
Car	11.63	5.61	15.34
Metro	0.46	1.13	0.05
Other	4.17	3.24	4.74
<i>All active modes</i>	83.74	—	—
<i>of which e-bike or unresolved</i>	25.40	—	—

Note: Unweighted means of grid-level shares across 5250 grids. Modes use speed–duration–distance signatures (metro flagged directly); strata with < 5 users are suppressed. Active mode includes e-bike and unresolved cycle/e-bike trips.

Table S6: Diagnostics for the city-level null results in Shenzhen and Xuzhou.*Panel A: variation available for estimation and attenuation of the raw gradient*

City	SD of C_i	% at 8	R_C^2	Raw	Reduced	Full	1:2 km
Shenzhen	1.51	77.1	0.307	2.21	1.38	0.61	0.65
Nanjing	2.29	35.6	0.436	2.56	1.55	0.58	0.60
Xuzhou	2.65	21.5	0.482	2.06	0.76	0.04	0.46

Panel B: same grids and month, two different outcomes

City	N	Short-trip share		Local-visit share	
		$\hat{\beta}$	p	$\hat{\beta}$	p
Shenzhen	2021	0.610	0.206	0.384	0.269
Xuzhou	5144	0.042	0.737	0.205	<0.001

Note: In Panel A, R_C^2 is from regressing service breadth on full covariates and district fixed effects; Raw is bivariate, Reduced retains the other covariates and district fixed effects but excludes population and center distance, Full is the baseline model, and 1:2 km is the ratio of ≤ 1 -km to ≤ 2 -km trip shares. Panel B uses the same grids and month for both outcomes, full covariates and district fixed effects; its p -values use the district-clustered normal approximation. These comparisons describe variation available for estimation and outcome sensitivity; they do not establish why the city coefficients differ.

Table S7: Quantity sensitivity, city-specific estimates, and typology stability audit.*Panel A: raw-scale common-cap estimates, pooled and by city*

Score	Weighting	Pooled	Shenzhen	Nanjing	Xuzhou
Breadth / $C^{(1)}$ (baseline)	place-based	0.301** (0.111)	0.610 (0.482)	0.578** (0.140)	0.042 (0.126)
	trip-weighted	0.960*** (0.299)	0.020 (0.431)	1.669** (0.351)	0.250 (0.410)
Common-cap $C^{(3)}$	place-based	0.369** (0.130)	0.692 (0.448)	0.727*** (0.133)	0.027 (0.154)
	trip-weighted	1.158*** (0.266)	0.308 (0.393)	1.818*** (0.355)	0.429 (0.425)
Common-cap $C^{(5)}$	place-based	0.422** (0.139)	0.739 (0.439)	0.817*** (0.129)	0.037 (0.164)
	trip-weighted	1.268*** (0.251)	0.487 (0.393)	1.904*** (0.325)	0.598 (0.416)

Panel B: pooled place-based estimates, standardized supply score

Supply score	$\hat{\beta}$ per SD	Cluster SE	p_b	BH q	R^2
Breadth / $C^{(1)}$	0.759	0.281	0.0179	0.0193	0.6011
Common-cap $C^{(2)}$	0.907	0.333	0.0193	0.0193	0.6014
Common-cap $C^{(3)}$	1.034	0.363	0.0162	0.0193	0.6016
Common-cap $C^{(5)}$	1.211	0.400	0.0108	0.0189	0.6020
Common-cap $C^{(10)}$	1.426	0.454	0.0079	0.0184	0.6024
Category-balanced $C^{(\log 95)}$	2.398	0.569	0.0006	0.0042	0.6062
Log total points of interest (POIs)	1.647	0.493	0.0043	0.0151	0.6036

Note: $C^{(m)} = \sum_k \min(n_{ik}, m)/m$ scores each category by the number of facilities present, capped at m ; breadth is the $m = 1$ case. All integer caps from 1 through 10 were estimated, and Panel B reports representative points from this path. The caps are sensitivity parameters, not service-adequacy thresholds. Models use the full covariate set and district fixed effects ($N = 11,027$). Panel A presents raw-score coefficients with district-clustered SEs; stars use the null-imposed full-refit CR1 wild-cluster bootstrap. Panel B standardizes each score, and its BH column adjusts the seven displayed bootstrap tests as one family. * $p_b < 0.1$, ** $p_b < 0.05$, *** $p_b < 0.01$.

Table S7: Quantity sensitivity, city-specific estimates, and typology stability audit (continued).

Panel C: trip-weighted estimates by city, coefficient per within-city SD [p_b]

	Supply score	Shenzhen	Nanjing	Xuzhou
Breadth / $C^{(1)}$		0.031 [0.961]	3.824 [0.012]	0.662 [0.553]
Category-balanced $C^{(\log 95)}$		2.614 [0.090]	7.174 [0.001]	3.160 [0.004]
Log total POIs		3.356 [0.027]	7.255 [0.001]	2.585 [0.100]

Panel D: typology stability at the 0.5 percentile-rank cut against the breadth baseline

Sample	Alternative score	Agreement (%)	κ	High-breadth/ low-local-travel J	Low-breadth/ high-local-travel J
Pooled	$C^{(3)}$	91.23	0.878	0.765	0.770
Pooled	$C^{(5)}$	88.73	0.844	0.706	0.715
Pooled	$C^{(\log 95)}$	89.67	0.857	0.705	0.757
Pooled	Log total POIs	87.23	0.823	0.651	0.697
Shenzhen	$C^{(3)}$	81.06	0.736	0.662	0.493
Shenzhen	$C^{(5)}$	73.19	0.637	0.539	0.394
Shenzhen	$C^{(\log 95)}$	72.16	0.621	0.466	0.422
Shenzhen	Log total POIs	70.77	0.602	0.441	0.394

Panel E: joint test of the two city-by-score interaction terms

Estimator	Supply score	CR1 Wald p	Joint p_b	BH q
M5 place-based	Breadth	0.563	0.849	0.849
	$C^{(3)}$	0.445	0.749	0.849
	$C^{(5)}$	0.385	0.666	0.849
	$C^{(10)}$	0.241	0.493	0.849
	$C^{(\log 95)}$	0.094	0.226	0.849
M6 trip-weighted	Log total POIs	0.288	0.541	0.849
	Breadth	< 0.001	0.163	0.196
	$C^{(3)}$	0.001	0.142	0.196
	$C^{(5)}$	0.005	0.163	0.196
	$C^{(10)}$	0.011	0.163	0.196
	$C^{(\log 95)}$	0.014	0.130	0.196
	Log total POIs	0.122	0.339	0.339

Note: $N = 11,027$ for regressions and 11,047 for typology comparisons. All regressions contain the main covariates and district fixed effects. p_b is from a null-imposed, full-refit CR1 wild-cluster procedure: one-coefficient rows use bootstrap- t , while Panel E uses a two-degree-of-freedom bootstrap Wald statistic. Pooled and joint models use 9999 seeded Rademacher draws, and city models enumerate all sign patterns (9–11 district clusters). Panel C standardizes each score within city, so its magnitudes are not estimates on a common raw-score scale. The common-cap and $C^{(\log 95)}$ measures are construct-sensitivity scores rather than canonical or normative planning indices. The former places each category on a common capped 0–1 scale; the latter applies log transformation, category-specific normalization, upper capping and equal weighting. $C^{(\log 95)}$ uses empirical category-specific 95th-percentile caps on positive log counts for scaling, not as service-sufficiency thresholds. J is the Jaccard overlap of class membership. In Shenzhen, 77.1% of grids have breadth eight, so a continuous count score breaks a large breadth tie and materially changes identity. No common cross-category grade, capacity or quality field is observed; the limited source-taxonomy hierarchy tags are examined separately in Table S10. Panel E uses pooled common covariate slopes and district fixed effects; its joint test evaluates whether the two city-interaction coefficients depart from zero.

Table S8: Leave-one-city-out cross-validation of the gradient-boosting model.

Held-out city	N_{test}	Held-out R^2 (no city dummies)	Held-out R^2 (with city dummies)
Shenzhen	2022	−0.252	−0.431
Nanjing	3872	0.388	0.463
Xuzhou	5153	0.274	−0.174
Shuffled five-fold (pooled, no dummies)			0.650
Shuffled five-fold (pooled, with dummies)			0.681

Note: XGBoost with the main-text hyperparameters, trained on two cities and evaluated on the third (grouped by city); the no-dummy shuffled five-fold row is the same-feature within-distribution comparator for the primary LOCO rows, while the dummy specification reproduces the pooled auxiliary model. The dummy-column LOCO results are retained only as an encoding diagnostic because the held-out city’s category is absent from training; they are not primary transfer estimates. A negative held-out R^2 means the transferred model predicts worse than the held-out city’s own mean.

Table S9: Park/scenic composition audit of the low-breadth/high-local-travel group.

<i>Panel A: low-breadth/high-local-travel group versus the aligned-low group</i>				
Measure	Low-breadth/ high-local-travel	Aligned-low	Difference [95% CI]	Cluster boot. p ; BH q
Any park/scenic POI (%)	20.36	20.03	0.33 [−3.01, 3.79]	0.865; 0.865
Park/scenic count	0.605	0.641	−0.037 [−0.287, 0.217]	0.730; 0.865
Share among positive-POI grids (%)	4.14	4.92	−0.79 [−1.95, 0.21]	0.126; 0.252
Park/scenic-majority grids (%)	0.61	1.60	−0.98 [−1.48, −0.53]	< 0.001; 0.002

<i>Panel B: adjusted association with membership in the low-breadth/high-local-travel group among low-breadth grids</i>				
Focal observable	$\hat{\beta}$ (pp)	Cluster 95% CI	p_b ; BH q	N
Any park/scenic POI	0.26	[−3.55, 4.06]	0.898; 0.898	4997
$\log(1 + \text{park/scenic count})$	0.61	[−2.09, 3.30]	0.669; 0.898	4997
Share, per 10 percentage points	0.27	[−0.93, 1.48]	0.680; 0.898	4542
Park/scenic-majority grid	−11.56	[−19.12, −4.00]	0.010; 0.041	4997

Note: Baseline class sizes are 1631 low-breadth/high-local-travel and 3385 aligned-low grids. Panel A uses a city-stratified district-cluster bootstrap (3999 draws); BH adjusts the four pooled comparisons. Panel B controls for log total POIs, breadth, the main covariates and district fixed effects, with cluster intervals and a null-imposed, full-refit wild-cluster bootstrap (9999 draws); BH adjusts the four pooled wild-bootstrap tests. The share model excludes zero-POI grids. The majority result has only 63 positive cases and is a sparse diagnostic, not evidence that parks reduce travel. Xuzhou has higher descriptive presence/count in the low-breadth/high-local-travel group, but no higher share or majority prevalence. Across fourteen protocol-specified combinations of seven supply scores and cuts at the 0.5 percentile-rank threshold or the extreme terciles, the low-breadth/high-local-travel minus aligned-low share difference remains negative (−0.85 to −0.24 pp), as does the majority difference (−1.78 to −0.86 pp). The data contain no park area, popularity or visitor counts and no trip-to-POI destination link; the table tests counted composition only.

Table S10: Source-taxonomy hierarchy-tag and strict-taxonomy sensitivity.

Panel A: selected focal-score estimates under hierarchy adjustment						
Estimator	Focal score/specification	$\hat{\beta}$ (pp per SD)	CR1 SE	95% CI	p_b	
M5	$C^{(\log 95)}$, baseline	2.398	0.569	[1.282, 3.514]	0.0006	
M5	$C^{(\log 95)}$ + both hierarchy tags	2.260	0.568	[1.148, 3.373]	0.0013	
M6	$C^{(\log 95)}$ + both hierarchy tags	4.917	1.234	[2.499, 7.336]	0.0001	
M5	Breadth + both hierarchy tags	0.767	0.280	[0.219, 1.315]	0.0168	

Panel B: individual tag terms with domain-count adjustment and joint tests						
Estimator	Source-taxonomy presence tag	$\hat{\beta}$ (pp)	CR1 SE	95% CI	p_b	Holm p
M5	Grade III-A hospital	4.476	0.836	[2.838, 6.114]	0.0002	0.0004
M5	Rated scenic site	-0.026	1.750	[-3.456, 3.403]	0.9881	0.9881
M6	Grade III-A hospital	2.229	1.388	[-0.491, 4.949]	0.2093	0.2093
M6	Rated scenic site	4.131	1.926	[0.355, 7.907]	0.0674	0.1348
M5	Both tags jointly beyond breadth	Wald = 32.238; $q = 2$			0.0010	—
M6	Both tags jointly beyond breadth	Wald = 12.262; $q = 2$			0.0137	—
M5	Both tags jointly beyond $C^{(\log 95)}$	Wald = 21.557; $q = 2$			0.0044	—
M6	Both tags jointly beyond $C^{(\log 95)}$	Wald = 7.989; $q = 2$			0.0829	—

Panel C: exposure support on the regression sample						
Tag	Exposed cells	Exposed (%)	Districts with variation	Nanjing	Shenzhen	Xuzhou
Grade III-A hospital	388	3.52	22/30	144	154	90
Rated scenic site	165	1.50	23/30	80	36	49

Panel D: exploratory taxonomy-cleaning diagnostics						
Diagnostic	Sample/estimator	Estimate	CR1 SE	p_b	Agreement; κ	
Strict essential-service breadth	M5	1.362 pp/SD	0.348	0.0004	—	
Strict essential-service breadth	M6	3.545 pp/SD	0.715	0.0001	—	
Middle-taxonomy richness	M5	1.657 pp/SD	0.506	0.0042	—	
Middle-taxonomy richness	M6	4.595 pp/SD	1.105	0.0001	—	
Strict-breadth typology	Pooled	—	—	—	88.6%; 0.842	
Strict-breadth typology	Shenzhen	—	—	—	75.3%; 0.663	

Note: The source workbooks' fine-category field was not used in the focal breadth score. "Grade III-A hospital" means that at least one POI within 1 km carries the source fine-category label translated as "Grade III-A hospital"; "rated scenic site" means a national/provincial scenic or World Heritage source label. These are exact English translations after trimming the source string and retaining the token before its first vertical bar. They are presence indicators for database-tagged POIs, not independent institution counts or a common quality scale. Re-aggregation exactly reproduces all eight focal category counts in all 11,047 cells. No records with an exact normalized-name match fall within the 50/100-m clustering thresholds. At 100 and 500 m, respectively, deliberately aggressive name-blind spatial clustering changes hospital exposure in 6 and 17 cells and scenic-site exposure in 0 and 7 cells; the M5 adjusted-depth and tag conclusions remain unchanged. The clustering exercise bounds the sensitivity of tag exposure without constructing an institution crosswalk. A 5/10-km spatial heteroskedasticity-and-autocorrelation-consistent (spatial-HAC) check gives SEs of 0.437/0.502 for the adjusted depth term and 0.854/0.864 for the hospital tag; these are spatial-HAC diagnostics, while the reported p_b remains the primary small-cluster inference. Panels A–C and the Panel D regressions use 11,027 complete cases and 30 district clusters; the strict-typology rows use all 11,047 cells. Continuous scores are standardized on the 11,027 complete cases, while the $q_{.95}$ caps are defined on all 11,047 cells. Panel A uses the full controls and district fixed effects. In Panel B, each individual hospital or scenic row additionally adjusts for its corresponding log strict- health or log non-park-scenic POI count. The protocol joint rows test both tags beyond focal breadth; the two additional joint rows test them beyond $C^{(\log 95)}$ as a construct sensitivity.

The 95% CIs are normal-reference intervals based on the CR1 SEs; inferential judgements use p_b . The latter uses the null-imposed, full-refit CR1 wild-cluster procedure with 9999 draws over 30 districts. Holm adjusts the two tag tests within M5 and, separately, within M6. The strict taxonomy removes animal care and medical sales from health, entertainment from sport, non-public training/ media from culture and non-core financial/life entries. Panel D is exploratory, and its p_b values are not multiplicity-adjusted. Beds, places, area, staffing, hours, price, ratings, service volume and experienced quality remain unavailable.

Table S11: Post-hoc park/scenic functional-composition and catchment sensitivity.*Panel A: simultaneous conditional-composition model in the baseline low-breadth sample*

Term/test	$\hat{\beta}$ (pp/log unit)	CR1 SE	CR1 95% CI	p_b	Holm p
$\log(1 + \text{local-park/plaza records})$	1.643	0.871	$[-0.065, 3.351]$	0.145	0.290
$\log(1 + \text{other park/scenic records})$	-0.351	0.537	$[-1.403, 0.701]$	0.565	0.565
Both terms jointly		Wald = 3.945; $q = 2$		0.248	—

Panel B: radius-specific restriction to grids with zero counted park/scenic POIs

Catchment	Zero in baseline low-breadth/high-local-travel	Fixed low-breadth/ high-local-travel N ; y (%)	Fixed aligned-low N ; y (%)	Reclassified low-breadth/high-local-travel / aligned-low N
800 m (boundary)	1389 (85.16%)	1389; 23.59	2889; 6.99	1214 / 2009
1000 m (baseline reference)	1299 (79.64%)	1299; 23.34	2707; 6.70	961 / 1589
1200 m (boundary)	1126 (69.04%)	1126; 22.53	2443; 6.24	918 / 1408
2000 m (outcome-aligned)	583 (35.74%)	583; 19.90	1527; 5.06	508 / 647

Note: These modules are post-hoc exploratory sensitivities, not confirmatory tests. Panel A uses 4997 complete cases in 30 district clusters. Both terms enter simultaneously, conditional on seven-category non-park breadth, log non-park POI volume, the main covariates and district fixed effects. p_b is the null-imposed CR1 wild-cluster bootstrap with 9999 draws; Holm adjustment covers the two individual terms. Panel B retains the baseline labels in the fixed-class columns and re-ranks outcome and seven-category park-free breadth only in the final column. At 2 km, the 583 zero-park grids in the baseline low-breadth/high-local-travel group comprise 479 in Xuzhou, 102 in Nanjing and 2 in Shenzhen. As in Table S9, these data provide facility counts rather than destination-linked use measures. The subtype and catchment checks therefore cannot identify a park-attraction mechanism.

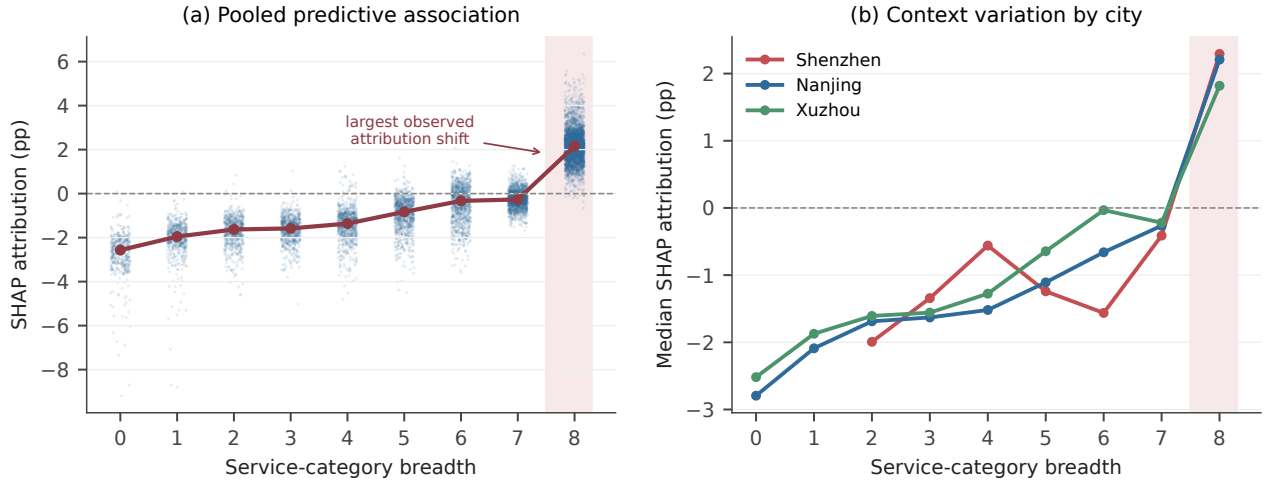


Figure S1: SHAP attribution profiles for service breadth. Panel (a) shows cell-level attributions and medians; panel (b) shows city-specific median profiles. Shading marks the 7-to-8-category transition. Values are percentage-point model attributions from the pooled gradient-boosting fit, not marginal effects. Table S8 shows the pooled model's cross-city predictive performance.

Reference

Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics* **2008**, *90*, 414–427.