Online Resource 1

Supplementary Information

Retail density and category overlap condition street-level shop survival near large commercial complexes: evidence from Zhengzhou, China

Applied Spatial Analysis and Policy. Author name, affiliation and e-mail address are withheld from this file for double-anonymous peer review, and are given on the title page.

This file reports quantities referred to but not tabulated in the main text, and contains Fig. S1. Section numbers refer to the main text.

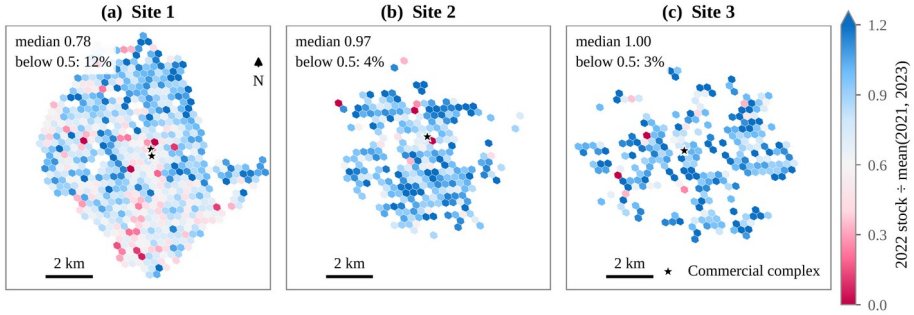


Fig. S1 The 2022 coverage gap. Each cell's shop count in 2022 divided by the mean of 2021 and 2023, by H3 resolution 9 hexagon. The shortfall is concentrated at Site 1, which is why 2022 is excluded rather than interpolated (Section 3.2)

Table S1 Checks on the two conditional estimates. Full numbers and sources are given in the sections that follow.

Check	What it varies	Effect on the estimate	Effect on the inference
Conley truncation	500 m to 3,000 m	None: +0.01162 throughout	p from 0.057 to 0.012
Dropping a site	Two sites instead of three	+0.0098 / +0.0115	Unchanged
Clustering scale	250 m to 2 km cells	None: +1.14 throughout	p from 0.018 to 0.097
Placebo moderator	Site ordinal for density	-0.90	Not significant (p = 0.125)
Chain outlets	16,481 to 22,883 shops	+1.17; reversal 318	Unchanged
Continuous overlap	Ratio for indicator	Same sign, slightly smaller	Unchanged
Category stratification	Three largest categories	Ordering reproduced	Shopping services p = 0.30-0.40
Random forest ordering	No functional form	Spearman 0.929	Exact one-sided p = 0.0011
Three-way interaction	Overlap on the reversal	+0.23	Not distinguishable from zero

S1 Regional accessibility of the three case areas (Section 3.1)

Sites 1 and 3 have near-identical grain, at 9.39 against 9.81 kilometres of street per square kilometre and 23.3 against 23.5 intersections per square kilometre, but differ at the regional

scale. Advancing 5,000 metres along the network from each origin point by Dijkstra's algorithm reaches 520 kilometres of street at Site 1, 327 at Site 3 and 290 at Site 2, so the same distance budget buys Site 1 about 1.6 times as much network as Site 3. Site 1's dense fabric extends continuously outward, whereas Site 3's density is local and does not propagate. Local density is therefore not regional accessibility, and the two are separated in the morphological indicators of Table 1.

S2 Spatial lags in the grid-scale specification (Section 3.3)

A regressor and its own spatial lag are strongly correlated in these data, at 0.74, 0.59 and 0.51 across the three sites. The lagged terms are therefore assessed by a joint Wald test rather than by individual t statistics. None of the three site-level tests rejects. The lags are retained for consistency with the diagnostic that selected the spatial Durbin family, and no lag coefficient is interpreted. Because the estimated member of that family is the spatial lag of X specification, which carries no lagged dependent variable, no decomposition into direct and indirect effects is defined and none is reported.

S3 Robustness of the Site 1 distance gradient (Section 4.1)

The focal distance coefficient at Site 1 is invariant to the Conley truncation distance, holding at +0.01162 across every cut-off tested, while its p value falls as the kernel widens: 0.057 at 500 metres, 0.059 at 750 metres, 0.034 at 1,500 metres and 0.012 at 3,000 metres. The main text reports this range rather than the most favourable value, and reads the gradient as marginal rather than significant under the shortest cut-offs.

Estimating on two sites rather than three leaves the Site 1 coefficient essentially unchanged: +0.0098 (SE 0.0035) with Site 2 dropped, and +0.0115 (SE 0.0008) with Site 3 dropped.

S4 Identifying variation and inference in the density interaction (Section 4.2)

The residualised variance of the distance-by-density interaction term falls from 1.455 under M2 and 1.441 under M3 to 0.789 under M4, a loss of 46 per cent, because site-specific distance slopes absorb the between-area contrast in full. A variance decomposition attributes 37.1 per cent of the variation in baseline retail density to differences between sites and 62.9 per cent to differences within them.

The point estimate does not respond to the level of clustering. At cell sizes of 250 metres, 500 metres, one kilometre and two kilometres the interaction holds at +1.14 percentage points per kilometre per log unit of density, while the number of clusters falls from 1,337 to 508, 172 and 56 and the p value moves from 0.018 to 0.054, 0.054 and 0.097. At two kilometres the cluster

count is below the number at which cluster-robust asymptotics can be relied upon (Cameron & Miller, 2015), which is why the main text estimates are clustered at 500 metres.

A placebo interaction replacing baseline retail density with the site ordinal gives -0.90 percentage points (t = -1.54, p = 0.125). With three sites this test has limited power and is not read as proof that density and site order are separable.

Including chain outlets raises the sample from 16,481 to 22,883 shops and leaves the interaction at +1.17 percentage points (t = 1.91), with the reversal at 318 units within 500 metres.

S5 The overlap measure (Section 4.3)

The overlap ratios rest on very unequal numbers of in-complex records: 932, 586, 289, 69, 21, 16, 1 and 0 across the eight categories. The ratio of 0.00 for vehicle repair rests on zero observations and has an exact upper bound of 0.138.

Replacing the binary low-overlap indicator with the continuous overlap ratio gives the same direction and a slightly smaller coefficient, which is what measurement error in a regressor predicts. The indicator version requires only that a shop's category be classified correctly, while the continuous version requires the ratio itself to be accurate, so the indicator is used in the main specification and the continuous estimate is read as conservative.

Stratifying the three largest categories separately by overlap tier reproduces the ordering shown in Fig. 6. Under standard errors clustered on 500 metre cells, shopping services alone is not individually significant, at p between 0.30 and 0.40 across specifications. This is consistent with the position taken in the main text, that eight categories support an ordering and a two-group contrast but not eight separate inferences.

S6 The random-forest ordering check (Section 4.3)

A random forest was fitted to survival with no reference to the overlap ratio, using 300 trees, a minimum leaf size of 25, the square root of the feature count at each split, class weights balanced by subsample, and a fixed random seed. Ranking the eight categories by the slope of the distance-attribution relationship recovered from the fitted surface by SHAP attribution, and comparing that ranking with the ranking by overlap ratio, gives a Spearman correlation of 0.929. Enumerating all 40,320 orderings of eight categories, 45 reach a correlation at least this high, an exact one-sided p of 0.0011.

The forest is used for ordering alone. A SHAP attribution slope is measured in attribution units per kilometre rather than in percentage points of survival, and its apparent precision reflects the smoothing of the fitted surface rather than sampling variation in the data, so no attribution magnitude is quoted in the main text or here.

S7 Whether overlap moves the reversal density (Section 4.3)

The three-way interaction between street-network distance, baseline retail density and the low-overlap indicator is +0.23 percentage points per kilometre with a t statistic of 0.20, and is indistinguishable from zero.

Solving each of the eight categories separately for its own reversal density produces a set of values that is not ordered by overlap. The Spearman correlation between the two orderings is -0.31, and 9,310 of the 40,320 possible orderings reach a correlation at least this negative, an exact p of 0.23. An implied reversal density is a ratio of two coefficients, and in a category of a few hundred shops the denominator is not estimated precisely enough for the ratio to carry information: sports and leisure, with 501 shops, returns a threshold of about 3,700 units within 500 metres, above any density observed anywhere in the study area.

A three-way interaction identified from eight categories, three of which hold fewer than a thousand shops, has little power. The absence of evidence that the two conditions interact is not evidence that they do not.

S8 The near-zone threshold model (Section 4.1)

The model replaces the linear street-network distance term at Site 1 with an indicator for lying within a radius of γ kilometres of a complex. γ is chosen by grid search over the observed range of distances to minimise the residual sum of squares. The selected radius is 1.30 network kilometres, inside which cell-level survival is 7.32 percentage points lower than outside it; the zone contains 49 of Site 1's 632 cells and 13.0 per cent of its baseline shops. The F statistic of the selected model against the model without a threshold is 46.6. Because γ was chosen by search rather than fixed in advance, that statistic is not a valid test for the presence of a threshold and is reported here only for completeness; the radius is described in the main text as where the effect is concentrated and not as a tested boundary. It is estimated at one site in one city and is not a quantity to be transferred.