A REVENUE MODEL OF SELF-STORAGE: TRAINING AND PREDICTION

Estimation on Confidential Stores and semi-final scoring of
Space Shop stores

A transportable revenue model can be trained on one branded portfolio and scored on another. In the semi-final product, Tier A LassoCV is the point estimate, OLS is the benchmark, and Tier B extends coverage when VALUE is missing.

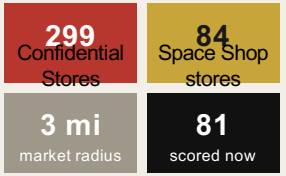
Daniel J. Aronoff, MIT
Sarah Newman, Michigan State University
Mark Serrahn, Landon Companies

April 2026

THE WORKFLOW

A semi-final prediction system, not yet a full external validation

This project estimates a transportable revenue model on a confidential data set of 299 self-storage facilities and then predicts sales at Space Shop stores with the chosen model. The missing final step is to compare those scores to realized Space Shop sales [*]



ESTIMATE

Fit the model on Confidential Stores using market characteristics around each store and log(SPSF + 1) as the target.

SCORE

Apply the saved transforms, training scaler, and harmonized lease-up indicator to Space Shop markets; predict SPSF and total sales when SqFt exists.

VALIDATE

When realized Space Shop sales arrive, compare predictions to outcomes and then decide whether to tighten or re-specify the model.

SEMI-FINAL PRODUCT

Augmented Tier A OLS is now the primary point estimate, Tier A LassoCV is the near-tied benchmark, and Tier B remains the no-VALUE fallback. Governance is explicit: high-risk or high-disagreement stores are flagged for manual review.

[*] A potentially confounding factor is that the confidential store training data is pre-COVD and the Space Shop predictions are post-COVID.

The Space Shop predictions will be inconsistent in COVID induced a shift in model parameters

This may necessitate re-estimating the model with Space Shop sales and a COVID shift term (or separate estimates for the pre and post COVID groups with a comparison of parameter values

A more definitive exercise would be to estimate on a fixed panel of stores, pre and post COVID

DATA, MARKET AREA, AND OUTCOME

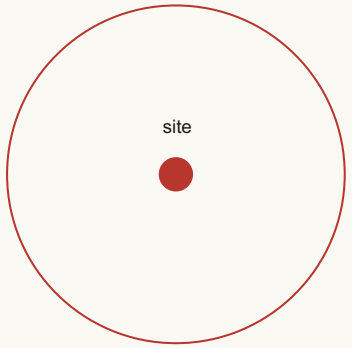
Training on Confidential Stores; scoring on Space Shop; market defined as a 3-mile radius

DATA DESIGN

Training sample: 299 Confidential Stores in 2018. Scoring sample: 84 Space Shop stores from a raw file merged to Yardi and a matched vintage supplement. Exact address joins: 84 / 84 matched for Yardi and 84 / 84 matched for vintage. Transport set includes population, density, income, age, housing stock, movers, VALUE, storage share, signed-log population growth, and opened less than 2 years.

TARGET VARIABLE

The model predicts revenue intensity rather than raw sales: y = log(SPSF + 1). This compresses the upper tail and reduces the influence of a few extreme stores.



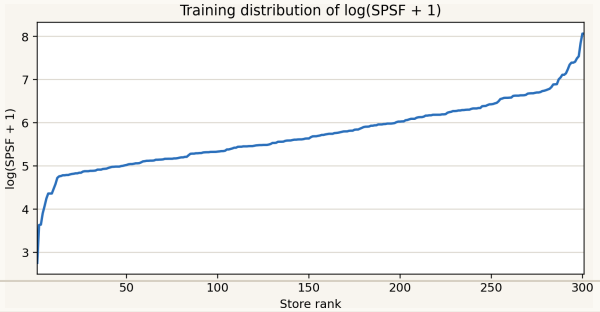
3-mile market area

exact join	19 SqFt invalid
22 missing x	13 final x vars

PRACTICAL OUTPUT

Every score is stored as predicted SPSF. Total sales are produced only when valid square footage exists; otherwise the sales field remains blank by design.

log(SPSF + 1) in the training sample



CHOSEN SCORING MODEL

Augmented Tier A OLS is the primary point estimate; Tier A LassoCV remains the near-tied benchmark

The refreshed scoring product has three layers: a primary model, a benchmark model, and a coverage-extension fallback.

OLS scoring equation

$$\hat{y}_i^{OLS} = \beta_0 + \Sigma_j \beta_j g_j(x_{ij})$$

Tier A OLS is now estimated on the transportable predictor set including Opened less than 2 years, and it is the narrow cross-validation winner.

LASSO scoring equation

$$z_{ij} = [g_j(x_{ij}) - \mu_j^{train}] / \sigma_j^{train}, \quad \hat{y}_i^{L} = \beta_0 + \Sigma_j \beta_j z_{ij}$$

Because LASSO coefficients live on standardized transformed predictors, Space Shop stores must be scored with the saved Confidential Stores means and standard deviations.

Raw Space Shop market data → Common-variable mapping → Saved transforms and scaler → Tier A with VALUE + Opened <2 years / Tier B no VALUE

FINAL OUTPUT

Predict smearing-adjusted SPSF for every scoreable store. Convert SPSF to total sales only when SqFt is present and positive. Add transport-risk flags and an OLS-vs-LASSO disagreement check.

PRIMARY
Tier A OLS augmented

BENCHMARK
Tier A LassoCV augmented

FALLBACK
Tier B LassoCV no VALUE

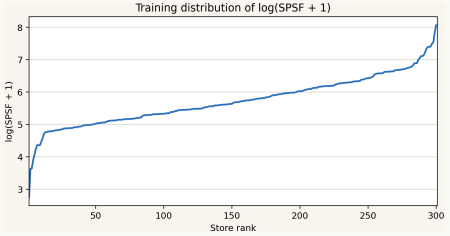
WHICH VARIABLES CARRY THE SIGNAL?

The refreshed transport model emphasizes wealth, housing stock, turnover, and lease-up status

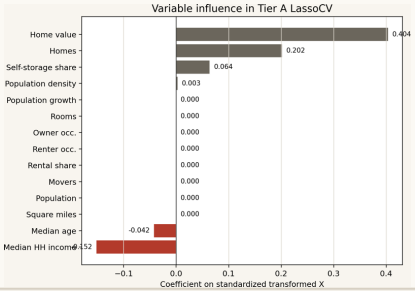
The scoring model is deliberately constrained to variables that both portfolios can define consistently. Within that transportable set, the main economic signals line up with plausible self-storage demand channels, and the new lease-up term sharpens the distinction between newly opened and seasoned stores.

Key transportable variables	sign	possible self-storage demand channel
POP_SQMI	+	denser local markets support deeper storage demand
VALUE	+	higher local housing values proxy for purchasing power and tighter space constraints
MOVERS	+	turnover raises the frequency of storage-using events
ROOMS	–	larger homes reduce the need to rent off-site storage space
MED_AGE	–	older market composition tends to predict lower revenue intensity
ss_share	+	the self-storage composition of the local market carries incremental signal

Training distribution of log(SPSF + 1)



Updated variable influence incl. lease-up

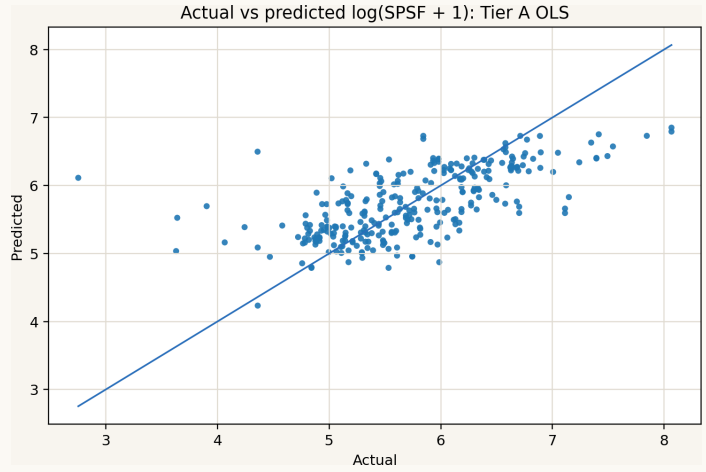


READING THIS SLIDE

The final model is not claiming causal effects. These signs are descriptive guides to what the scoring model uses to rank markets. Opened less than 2 years now enters as a stable negative lease-up indicator in the refreshed specification.

ESTIMATION ON CONFIDENTIAL STORES

The chosen model is selected on transport-ready cross-validation, not only on in-sample fit



LEGACY WORKBOOK

OLS: R^2 = 0.5725
Adj. R^2 = 0.5264
RMSE = 0.4927

REFRESHED MODEL CHOICE

R^2 = 0.5078
RMSE = 0.5286
α = 0.075

BOTTOM LINE

With signed-log population growth added, Tier A LassoCV still edges Tier A OLS in cross-validated RMSE (0.5653 versus 0.5713). Population growth is now admissible under the transport mapping, but the selected Tier A LassoCV shrinks its coefficient to zero.

The gap is small. Shrinkage slightly improves predictive discipline while OLS remains close enough to serve as a transparent benchmark.

BOTTOM LINE

Tier B no-VALUE LassoCV is materially weaker, which is why it is treated as coverage extension rather than coequal with Tier A.

TRANSPORTABILITY AND SUPPORT CHECKS

Space Shop is scored only when the variables can plausibly be treated as draws from the Confidential Stores design

The central transport question is not whether Space Shop looks identical to Confidential Stores, but whether the Space Shop markets fall inside a support envelope that makes the learned mapping usable.

1 Common-variable mapping

Every scored Space Shop field must map cleanly into the Confidential Stores design. The model includes a matched Space Shop vintage supplement so Opened less than 2 years can be carried consistently across portfolios.

2. Support Envelope

Each Space Shop store is checked against training support using range and standardized-distance heuristics. High-risk rows are flagged for review.

3. Identical transforms

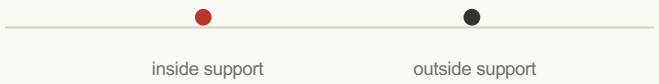
Space Shop predictors are transformed exactly as in training; LASSO additionally uses the saved Confidential Stores means and standard deviations.

4. Conservative scoring rules

No blanket imputation. Population growth is retained under the signed-log transform; fields that still fail a defensible mapping are dropped rather than forced into the score.

Training support

Simple support heuristic

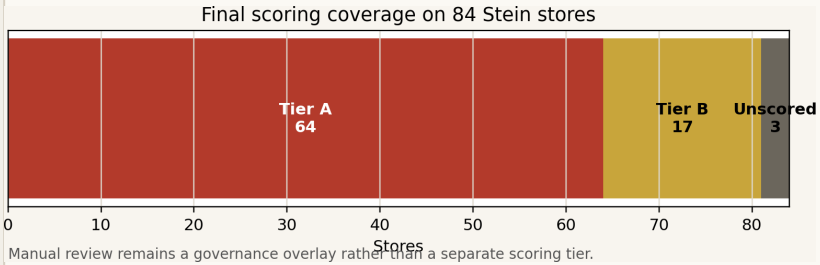


SIGNED-LOG GROWTH UPDATE

The PopChange issue is resolved with $\text{sign}(x) \cdot \ln(|x| + 1)$, which permits negative growth values without breaking transport scoring. In the updated Tier A LassoCV, the new term is admissible but shrunk to zero.

MISSING DATA AND COVERAGE RULES

Coverage is extended deliberately rather than by blanket imputation



Final scoring coverage on 84 Space Shop stores

64	17	20	3
Tier A rec.	Tier B rec.	manual review	unscored

HOW MISSINGNESS IS HANDLED

VALUE is the main missing Space Shop field. Rather than imputing it broadly, the refreshed product keeps a stronger Tier A model with VALUE and a weaker Tier B fallback without VALUE. Coverage remains 81 of 84 stores even after the lease-up term is added.

PREDICTORS

The matched Space Shop vintage supplement is complete: 84 / 84 stores now carry the Opened less than 2 years flag.

SqFt limits

Nineteen rows have missing or nonpositive SqFt. Those stores still receive SPSF predictions, but total sales are intentionally left blank.

LEASE-UP INDICATOR: NEW STORES VERSUS SEASONED STORES

Opened less than 2 years is a transportable predictor and materially improves fit

WHY VINTAGE MATTERS

The vintage term is best interpreted as a lease-up indicator rather than a calendar-year effect. In Confidential Stores, Opened less than 2 years enters with a stable negative coefficient; in the refreshed Tier A OLS it is about −0.649 and highly significant.

HOW THE SPLIT IS TESTED

Portfolio-specific raw thresholds are now harmonized into one common transportable concept: opened_lt_2yrs = 1 if the store opened less than two years before observation, 0 otherwise. The matched Space Shop supplement is clean: 84 / 84 matched, with 59 flagged 1 and 25 flagged 0.

CURRENT DECISION

The augmented model should replace the prior baseline. Tier A OLS augmented is the narrow CV winner at 0.5415 RMSE, with Tier A LassoCV essentially tied at 0.5419. This is a model update, not just a relabeling exercise.

WHY MORE DATA HELPS

Once realized Space Shop sales arrive, the next step is to validate whether the lease-up term improves out-of-portfolio prediction as much as it improves fit inside the current design.

REFRESHED SPACE SHOP SCORING RESULTS

The lease-up indicator changes predicted levels materially while preserving coverage

43.5%

mean absolute SPSF change baseline vs augmented across 81 comparable rows

INTERPRETATION

The model leaves tier coverage essentially unchanged, but predicted SPSF levels move materially once the lease-up term is added. That is exactly what we would expect if the prior baseline was missing an economically important dimension of newly opened stores.

64	17	81
Tier A	Tier B	comparable

RULE FOR REVIEW

Manual review still turns on transport risk and large cross-model disagreement. The governance layer remains intact even though the preferred primary estimate has changed.

HOW TO READ THE PORTFOLIO OUTPUT

Use the Tier A OLS augmented score as the default estimate, keep Tier A LassoCV augmented nearby as a benchmark, and treat Tier B LassoCV as a fallback only for stores that cannot be scored with VALUE. The 81 overlapping rows move by about 43.5 percent on average relative to a model without store vintage

OPERATIONAL RECOMMENDATION

Use the model as an underwriting aid with explicit governance

PRIMARY ESTIMATE

Use Tier A OLS augmented with the smearing adjustment whenever the store can be scored in Tier A. That is now the main point estimate for SPSF.

BENCHMARK AND REVIEW

Keep Tier A LassoCV augmented alongside the primary estimate. If the OLS–LASSO gap is large or transport risk is high, route the store to manual review.

FALLBACK AND LIMITS

Use Tier B LassoCV only when Tier A cannot score because VALUE is missing. Do not convert SPSF to total sales when SqFt is missing or nonpositive.

BOTTOM LINE

The refreshed product is suitable for portfolio screening, relative ranking, and disciplined first-pass revenue estimation. It is still not a final valuation engine, because realized Space Shop sales have not yet been brought in for external validation.

WHY MORE DATA WILL IMPROVE PRECISION

The next gains come from more stores, more years, and realized outcomes

MORE TRAINING STORES

Expanding the training sample beyond 299 Confidential Stores would tighten coefficients, widen the empirical support envelope, and make support-flag heuristics more informative.

MORE YEARS

Comparing a panel of pre and post COVID stores would enable the identification of specific changes in sales generator between the two regimes at the level of the individual store tracked between the two regimes.

REALIZED SPACE SHOP SALES

Observed Space Shop outcomes would turn the current semi-final product into a true external validation exercise, allowing direct comparison of prediction error across tiers and model families.

PRACTICAL IMPLICATION

The current model can identify some effects of COVID on model parameters, once we hve Space Shop sales. However, there is no ability to track at the individual store level. Tracking pre and post-COVID sales for a group of individual stores in different markets (with heterogeneous market variables) would enable the identification of COVID induces shifts in sales drivers

TAKEAWAYS

A semi-final transport model for self-storage revenue

1. The refreshed model is a governed scoring system: Tier A OLS augmented point estimate, Tier A LassoCV augmented benchmark, and Tier B LassoCV fallback.

2. Opened less than 2 years is a transportable lease-up indicator. It enters with a stable negative sign and materially changes Space Shop SPSF projections while improving cross-validated fit.

3. The next empirical step is direct comparison of the refreshed Space Shop scores to realized Space Shop sales. That is what will turn the current semi-final product into a fully validated external prediction exercise.

CURRENT PORTFOLIO OUTPUT
64 Tier A · 17 Tier B · 20 manual review · 3 unscored

FINAL MESSAGE FOR THE LUNCH
This is best read as a disciplined method for prediction and scoring across portfolios. The key update is that a harmonized lease-up indicator is now transportable and strong enough to change the preferred specification.

A1

Opened less than 2 years lease-up effect

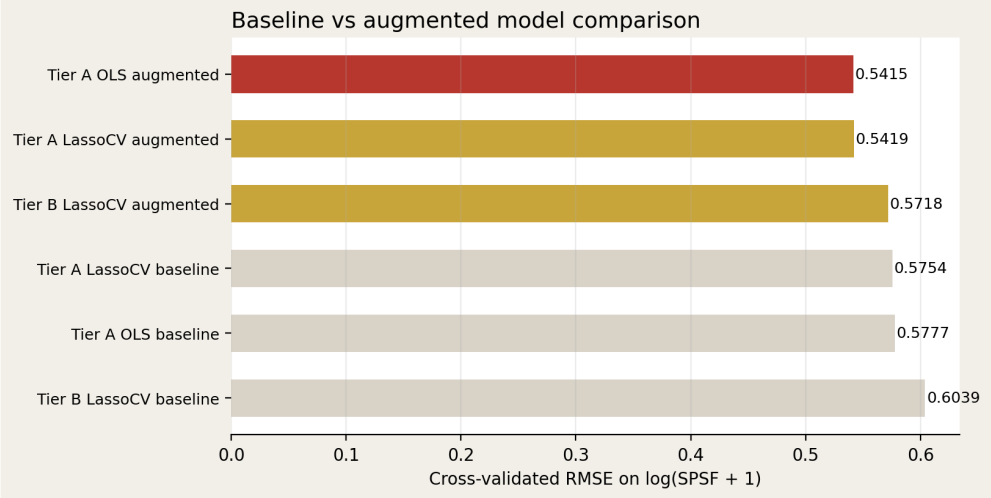
Harmonized lease-up indicator using portfolio-specific thresholds

Indicator used

Opened_lt_2yrs = 1 if store opened less than 2 years before observation

Confidential Stores uses the legacy opening-year flag; Space Shop uses the matched vintage supplement merged 84 / 84 by Match_addr. This makes the lease-up term transportable rather than merely descriptive.

Baseline vs augmented model comparison



WHAT CHANGES IN THE REFRESHED MODEL

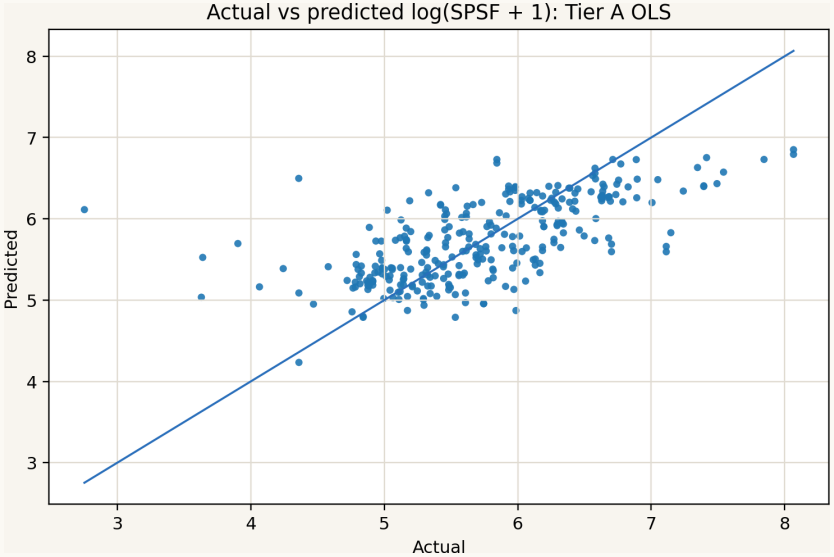
- Tier A OLS augmented is the narrow CV winner at 0.5415.
- Tier A LassoCV augmented is essentially tied at 0.5419.
- Both materially outperform the baseline w/o a lease-up parameter
- Opened less than 2 years is stable and negative in both OLS tiers.

A2

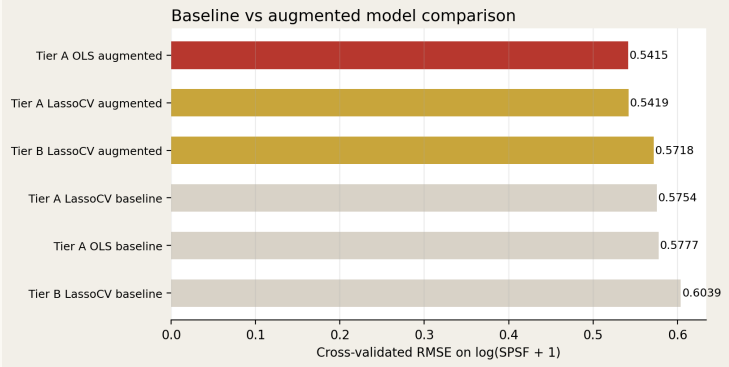
Enlarged performance charts

Legacy fit on the left; refreshed baseline-vs-augmented comparison on the right

Actual vs predicted log(SPSF + 1)



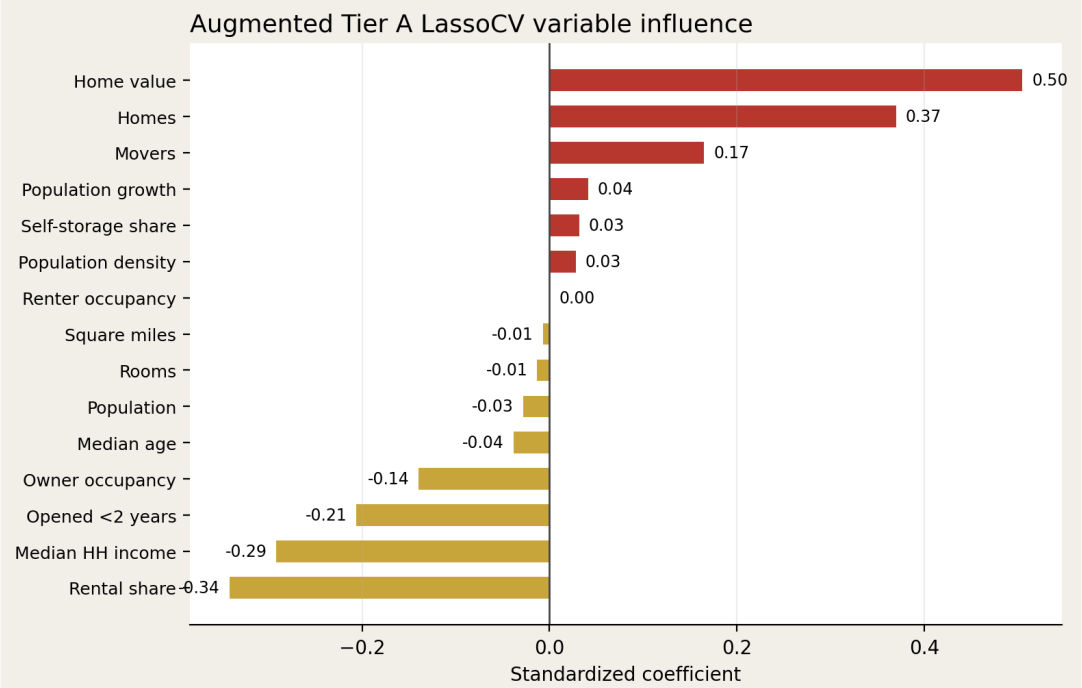
Refreshed CV RMSE comparison



A3

Variable influence in augmented Tier A LassoCV

Standardized coefficients after adding Opened less than 2 years



How to read the chart

- Home value, homes, and movers remain positive contributors in the augmented near-tie benchmark.
- Rental share, median household income, and Opened less than 2 years pull predicted SPSF down.
- The lease-up term is now selected rather than excluded, which is why the refreshed deck changes materially.

These are predictive weights on standardized transformed X, not causal effects.

A4

Variables considered in the updated transport spec

Transport-ready candidate variables after adding Opened less than 2 years

This appendix lists the transport-ready variables considered in the refreshed score and whether they survive with a non-zero coefficient in the augmented Tier A LassoCV.

Variable	Transform	Tier A	Selected
Population	log(1+x)	Yes	No
Population density	log(1+x)	Yes	Yes
Square miles	log(1+x)	Yes	No
Median HH income	log(1+x)	Yes	Yes
Median age	log(1+x)	Yes	Yes
Homes	log(1+x)	Yes	Yes
Rental share	log(1+x)	Yes	No

Variable	Transform	Tier A	Selected
Movers	log(1+x)	Yes	No
Owner occupancy	log(1+x)	Yes	No
Renter occupancy	log(1+x)	Yes	No
Rooms	log(1+x)	Yes	No
Home value	log(1+x)	Yes	Yes
Self-storage share	raw ratio	Yes	Yes
Population growth	sign(x)·ln(|x|+1)	Yes	Yes

Vintage / opening-year indicators were reviewed as a transportable lease-up concept, Opened less than 2 years. Fields excluded from the main transport-ready set were only those that could not be mapped consistently or scored honestly across portfolios.

A5

Space Shop locations and projected SPSF

Appendix table of recommended smearing-adjusted SPSF projections (1/4)

Space Shop location	Opened <2 yrs	Tier	Projected SPSF
10 Blackland Rd NW, Atlanta, Georgia, 30342	Yes	Tier A	229.3
100 Cadillac Pkwy, Dallas, Georgia, 30157	Yes	Tier A	103.3
1000 E Dublin Granville Rd, Columbus, Ohio, 43229	Yes	Tier B	226.5
1009 Oak Rd SW, Lilburn, Georgia, 30047	Yes	Tier A	151.7
1011 E Price Blvd, North Port, Florida, 34288	Yes	Tier A	68.2
10193 County Line Rd, Spring Hill, Florida, 34608	Yes	Tier A	152.9
10211 Highway 278, Covington, Georgia, 30014	No	Tier A	176.0
105 Matanzas Woods Pkwy, Palm Coast, Florida, 32137	Yes	Tier A	118.2
110 Highland Ave, Largo, Florida, 33770	Yes	Tier B	319.9
112 Tyrone Rd, Fayetteville, Georgia, 30214	Yes	Tier A	132.9
1141 E Memorial Blvd, Lakeland, Florida, 33801	Yes	Tier A	132.0
115 McKinna Pl, Brunswick, Georgia, 31520	No	Tier A	78.9
1189 Grayson Hwy, Lawrenceville, Georgia, 30045	Yes	Tier A	170.2
1250 Custer Ave SE, Atlanta, Georgia, 30316	Yes	Tier NA	—
12721 Bonita Beach Rd SE, Bonita Springs, Florida, 34135	Yes	Tier A	101.8
1363 Dogwood Dr SW, Conyers, Georgia, 30012	Yes	Tier A	49.6
14571 Snow Rd, Brook Park, Ohio, 44142	Yes	Tier NA	—
148 Mason Ave, Daytona Beach, Florida, 32117	Yes	Tier A	157.9
151 14th St NW, Atlanta, Georgia, 30318	Yes	Tier B	326.5
1575 Marketplace Dr, Rochester, New York, 14623	No	Tier A	126.9
1610 Glenside Dr, Richmond, Virginia, 23226	Yes	Tier B	234.0

A6

Space Shop locations and projected SPSF

Appendix table of recommended smearing-adjusted SPSF projections (2/4)

Space Shop location	Opened <2 yrs	Tier	Projected SPSF
1799 W Atlantic Blvd, Pompano Beach, Florida, 33069	Yes	Tier B	302.8
1813 Skyline Blvd, Cape Coral, Florida, 33991	Yes	Tier A	214.1
1819 James L Redman Pkwy, Plant City, Florida, 33563	Yes	Tier A	109.2
1850 Conyers Station Rd NE, Conyers, Georgia, 30013	Yes	Tier A	69.3
1868 Woodruff Rd, Greenville, South Carolina, 29607	No	Tier A	248.9
2002 Dresden Dr, Atlanta, Georgia, 30341	No	Tier B	540.3
2008 Olive St, Lakeland, Florida, 33815	No	Tier A	139.5
2040 Timothy Rd, Athens, Georgia, 30606	Yes	Tier A	44.5
22 Beth Stacey Blvd, Lehigh Acres, Florida, 33936	Yes	Tier A	70.2
2274 Northlake Center Dr, Tucker, Georgia, 30084	Yes	Tier B	230.9
2315 Highway 540a E, Lakeland, Florida, 33813	Yes	Tier A	148.9
2326 W Memorial Blvd, Lakeland, Florida, 33815	No	Tier A	131.7
23815 S Dixie Hwy, Homestead, Florida, 33032	Yes	Tier A	221.0
2520 Spring Rd SE, Smyrna, Georgia, 30080	Yes	Tier B	237.7
2590 Panola Rd, Lithonia, Georgia, 30058	No	Tier B	283.4
2600 Lincoln Way East, Massillon, Ohio, 44646	No	Tier A	177.2
2695 Buford Hwy, Duluth, Georgia, 30096	Yes	Tier A	67.6
27048 Lorain Rd, North Olmsted, Ohio, 44070	No	Tier A	206.7
2760 Hickory Tree Rd, Saint Cloud, Florida, 34772	Yes	Tier A	167.3
2872 N Pontiac Trl, Commerce Township, Michigan, 48390	Yes	Tier A	139.0
2925 W Frank Gordy Pkwy, Marietta, Georgia, 30066	No	Tier A	175.1

A7

Space Shop locations and projected SPSF

Appendix table of recommended smearing-adjusted SPSF projections (3/4)

Space Shop location	Opened <2 yrs	Tier	Projected SPSF
2971 Windy Hill Rd SE, Marietta, Georgia, 30067	No	Tier B	431.1
3000 Delk Rd SE, Marietta, Georgia, 30067	Yes	Tier B	218.6
3045 Hickory Grove Rd NW, Acworth, Georgia, 30101	Yes	Tier A	134.1
3323 17th St, Sarasota, Florida, 34235	Yes	Tier A	111.1
375 Fischer Rd, Sharpsburg, Georgia, 30277	Yes	Tier A	125.0
3751 Buford Hwy NE, Brookhaven, Georgia, 30329	No	Tier B	583.5
3765 Jefferson Rd, Athens, Georgia, 30607	Yes	Tier A	38.4
3860 Centerville Rosebud Rd, Snellville, Georgia, 30039	Yes	Tier A	68.0
3907 Lower Fayetteville Rd, Sharpsburg, Georgia, 30277	Yes	Tier A	88.2
400 Carpenter Dr, Atlanta, Georgia, 30328	Yes	Tier B	270.6
401 Pottery Factory Dr, Commerce, Georgia, 30529	Yes	Tier A	57.6
402 N Jackson Loop, Talmo, Georgia, 30575	No	Tier A	97.9
4163 Falcon Pkwy, Flowery Branch, Georgia, 30542	Yes	Tier A	97.0
426 Exchange Blvd, Bethlehem, Georgia, 30620	No	Tier A	182.7
4370 N State Road 7, Lauderdale Lakes, Florida, 33319	Yes	Tier A	195.7
4379 Wade Green Rd NW, Kennesaw, Georgia, 30144	Yes	Tier A	63.2
4399 John Tyler Hwy, Williamsburg, Virginia, 23185	Yes	Tier A	171.1
4470 Chamblee Dunwoody Rd, Atlanta, Georgia, 30338	Yes	Tier B	271.6
4510 Calhoun Memorial Hwy, Easley, South Carolina, 29640	Yes	Tier A	66.4
4822 Whipple Ave NW, Canton, Ohio, 44718	No	Tier A	134.3
5305 Peachtree Blvd, Atlanta, Georgia, 30341	No	Tier B	460.6

A8

Space Shop locations and projected SPSF

Appendix table of recommended smearing-adjusted SPSF projections (4/4)

Space Shop location	Opened <2 yrs	Tier	Projected SPSF
5844 US Highway 98 N, Lakeland, Florida, 33809	Yes	Tier A	78.4
635 Poplar Springs Rd, Riverdale, Georgia, 30274	No	Tier B	267.1
64 N Brookwood Ave, Hamilton, Ohio, 45013	No	Tier B	333.8
6401 McEver Rd, Flowery Branch, Georgia, 30542	Yes	Tier A	129.1
6410 Mechanicsville Tpke, Mechanicsville, Virginia, 23111	Yes	Tier A	105.2
6568 Malone Rd, Douglasville, Georgia, 30134	Yes	Tier A	93.9
6785 Thomasville Rd, Tallahassee, Florida, 32312	No	Tier A	135.4
682 Tanners Bridge Rd, Bethlehem, Georgia, 30620	Yes	Tier A	124.0
7060 Midnight Pass, Graniteville, South Carolina, 29829	Yes	Tier A	41.2
7105 Old National Hwy, Riverdale, Georgia, 30296	No	Tier A	197.7
7352 Friendship Springs Blvd, Buford, Georgia, 30519	Yes	Tier A	178.8
769 Nickajack Rd SW, Mableton, Georgia, 30126	Yes	Tier A	133.0
7816 N Military Hwy, Norfolk, Virginia, 23518	No	Tier NA	—
7868 Reynolds Rd, Mentor, Ohio, 44060	Yes	Tier A	53.5
807 Bookman Rd, Elgin, South Carolina, 29045	Yes	Tier A	153.8
810 Starratt Rd, Jacksonville, Florida, 32218	Yes	Tier A	122.5
813 Jonesboro Rd, Mcdonough, Georgia, 30253	Yes	Tier A	107.9
970 Hays Mill Rd, Suite A, Carrollton, Georgia, 30117	No	Tier A	109.9
Mill St, Batavia, Illinois, 60510	No	Tier A	122.9
Plaza 92 Dr, Douglasville, Georgia, 30134	No	Tier A	99.5
Ridgewalk Pkwy, Woodstock, Georgia, 30188	Yes	Tier A	131.8

Primary estimate shown = augmented recommended smearing-adjusted SPSF.