Online Appendix

Who Gets the House? Related-Party Transfers and the Cash–Mortgage Price Gap in New York City

Pablo Loschi

OA1 Replication map and classification definitions

I reproduce the house analysis from the frozen house inputs in replication archive v3.8 (`data/house_pairs_enriched.csv.gz`, `data/parties.json`, `results/related_party_flags.csv`, and `results/deed_match_audit.csv`). Dataset version 1.2.0 (Loschi, 2026) releases the underlying linked sale extracts (which contain 105,452 four-borough condominium sales under broader extraction rules), whereas the paper's corrected condominium panel (86,381 sales and 6,836 repeat pairs) comes from rebuilding the legacy single-parcel linkage rules over unit lots 1001–6999 while retaining archived financing labels on unchanged keys. Because dataset 1.2.0 does not bundle that derived legacy condo panel, this source-and-checks package includes `condo/condo_pairs_corrected_with_flags.csv.gz`, `condo/condo_bootstrap_corrected.csv.gz`, and `summary_figure/HOUSE_CONDO_CORRECTION.json` so an outside replicator can re-estimate the 0.83 and −0.37 condominium contrasts and check all 999 bootstrap draws offline.

Use the v3.8 inputs with the current package's scripts and paths in `house_paths.py`. The older `code/reproduce.py` also runs superseded review stages.

- `scripts/company_split.py`, `scripts/holding_period_split.py`, `scripts/doubled_pairs.py`, and `scripts/run_matched_deletion.py` estimate their respective comparisons from v3.8. The associated generators format the saved estimates for the manuscript.
- `million_split/million_split.py` estimates the property-value splits and `million_split/render.py` formats the results. `million_split/PRESALE_ASSESSMENT_MANIFEST.json` pins the Class 1 roll vintages and file hashes. `million_split/extract_presale_assessment.py` checks timing and keys against the bundled extract. The 12 non-Class-1 values enter from the saved extract.

- `summary_figure/make_figure.py` assembles Figure 1 from saved results and `summary_figure/make_mechanism_figure.py` assembles Figure 2 from the saved mechanism outputs. `scripts/house_condo_tail.py` refits the property-type sensitivity below. `scripts/extract_mechanism_attributes.py`, `scripts/extract_condo_ltv.py`, `scripts/family_type_split.py`, `scripts/ltv_gradient.py`, `scripts/fha_condo_mechanism.py`, `scripts/render_mechanism_checks.py`, and `scripts/render_fha_condo_mechanism.py` produce the building-class, LTV, direction-asymmetry, and house-versus-condominium LTV checks in Section OA7 and Tables OA6–OA8.

- `human_scoring/score_deidentified_validation.py` reproduces the public human-review counts and uncertainty from the de-identified case sheets. The sampled-endpoint regression refits require the restricted lookup that connects reviewed cases to transactions. The separate analytical panel and valuation files carry public parcel identifiers and transaction dates.

- `master_verify.py` checks archived file hashes, validation scores, selected estimates, saved draws and generated results. The three build commands in the README compile the PDFs separately.

Tables 1 and 2 and Figure 3 come from v3.8 outputs; Figure 2 comes from `summary_figure/make_mechanism_figure.py`; Table 3 comes from `code/robust_estimators.py`; Table 4 uses v3.8 houses and the bundled corrected condominium panel; Tables OA1–OA4 contain the inversion, co-operative assignment, credit, and market-conditions exhibits; Table OA5 contains the separate-band property-value fits; Tables OA6–OA8 contain the building-class, loan-to-value, direction-asymmetry, and condominium-versus-house LTV checks.

The house and condominium loan windows use document dates with a recording-date fallback; co-operative windows use personal-property recording dates. Party categories are reproducible regular-expression classifications, with unknown records kept separate from individuals.

OA1.1 Holding-period and upper-tail check

I use `scripts/holding_period_split.py` and `scripts/doubled_pairs.py`. They read the house replication inputs without changing them through `house_paths.py` (using `HOUSE_REPLICATION_DIR` or relative discovery). The split uses the first 199 draws of the common parcel bootstrap and refits the published nuisance design in each draw. The consecutive holding-period bins are $[3, 4)$, $[4, 5.5)$ and $[5.5, 99)$ years. The all-pair contrasts and percentile intervals are 8.45 [4.32, 12.11], 7.54 [4.30, 12.07] and 10.62 [7.20, 14.33]. Without same-surname pairs they are 4.17 [-0.16, 7.44], 2.09 [-1.06, 5.79] and 6.87 [3.98, 9.89].

The shortest-minus-longest interval is [-6.53, 2.42] for all pairs and [-7.06, 1.07] without same-surname pairs. These comparisons do not isolate renovation effects or show that contrasts are equal across holding periods. The upper-tail comparison contains 163 pairs that more than doubled among 1,044 cash-to-financed pairs. `scripts/generate_holding_numbers.py` reads `scripts/holding_period_b199.json` and `scripts/doubled_pairs.json` and writes `generated/holding_numbers.tex`; its adjacent manifest records the JSON hashes and rounding rules.

OA1.2 Property type, holding period and growth tails

I apply the surname screen to the house and condominium samples under common borough-quarter controls. I keep pairs held at least five years and trim annualized log growth below the 2.5th or above the 97.5th percentile separately within each property type. This leaves 3,587 house and 3,272 condominium pairs. The house estimate is 4.78 log points [3.17, 6.63] and the condominium estimate is -1.03 [-1.64, -0.36].

The difference is 5.81 log points [4.00, 7.77]. The percentile intervals use 999 draws from the union of parcel keys. I use the same draw for both property types, refitting the borough-quarter controls and trimming cutoffs within each draw. This comparison addresses the role of shorter holds and extreme growth. Differences in the sample, transaction composition and financing measurement still limit a property-type interpretation.

OA2 Execution-risk inversion

Consider a stylised seller comparing a certain cash price P_c with a financed offer P_f. Let q be the probability that the financed transaction fails and let dP_c be the loss, relative to the cash alternative, conditional on failure. If continuation following failure is worth $(1-d)P_c$, indifference implies

$$P_c = (1-q)P_f + q(1-d)P_c, \qquad \frac{P_f - P_c}{P_c} = \frac{qd}{1-q}. \tag{1}$$

For an observed log contrast π, treating that entire contrast as the model premium requires

$$d^*(q;\pi) = (e^{\pi} - 1)\frac{1-q}{q}. \tag{2}$$

This inversion makes the required loss explicit while leaving q as an assumption. I do not substitute national mortgage denial rates for q, because a denied application differs from the failure of an accepted purchase contract.

Table OA1: Relisting-loss scenarios implied by the observed contrasts

Failure probability q	Houses: $100d^*$	Condos: $100d^*$
0.1	87.2	7.5
0.2	38.8	3.3
0.3	22.6	2.0

Scenario calculations from $d^* = (\exp(\pi)-1)(1-q)/q$. Houses use the full geographic joint estimate; condos use the borough-quarter joint estimate. They have different covariate sets and populations. Values of q are illustrative assumptions; contract-failure probabilities are unobserved. No common relisting technology is imposed as an empirical fact.

At a common failure probability, the unscreened house contrast implies a loss more than ten times the loss implied by the unscreened condominium contrast. At $q = 0.1$ the house result requires a loss of about 87 percent of the price, or 63 percent using Huber. The surname screen reduces those figures to 45 and 37 percent. The unscreened condominium contrast implies about 8 percent, and the negative screened condominium contrast implies no positive loss. A seller can relist the property after a failed closing. The records do not show the chance of failure or the value of the seller's alternatives.

Two qualifications apply. First, equation (2) evaluates the counterfactual in which the entire observed contrast represents a causal financing premium, which my design does not establish, and so shows what execution risk would have to explain under that assumption. Second, failure probabilities could differ between the two property types, and to reconcile the least-squares contrasts with a common loss, the failure odds would have to be about twelve times higher for houses than for condominiums. The mechanism in Han and Hong (2024), where market-specific relisting prospects interact with closing risk, is the natural candidate for such a difference, and testing it would require contract-level data.

OA3 Co-operative assignment and adjacency audit

OA3.1 What the comparison can establish

Co-operative share financing is a possible institutional comparison, but the recorded loan measure differs from the house and condominium measure. Timing-unique filing assignments leave 1,765 repeat pairs with an interval including zero. I use this comparison as a measurement sensitivity check; it cannot serve as a control group that identifies a mortgage-tax effect.

OA3.2 Why the co-operative comparison requires an audit

A building-level UCC filing need not be a particular buyer's purchase loan. Several sales and filings can fall close together, some sales lack a recovered unit identifier, and filings can concern other credit. When I count possible competitors for a filing, the co-operative extract therefore includes every observed price and incomplete-unit record as well as the estimation sample.

The broadest rule, the expanded count rule, counts eligible sales within $d \pm 45$ days of a sale at date d and filings within $[d - 60, d + 135]$. A zero filing count is coded cash, a filing count at least as large as the sale count is coded financed, and the rest are ambiguous. The expansion matters. Of 60,496 purchases labelled financed, 4,385 have no filing in the base $[-15, +90]$-day window, and counting all competing sales also changes assignments.

The alternative rules require exactly one candidate filing and exactly one observed sale eligible to claim it, with narrow, base and wide windows of $[-15, +30]$, $[-15, +90]$ and $[-15, +180]$ days. Documents linked to several parcels are excluded. These conditions stop one filing from being used twice but say nothing about its debtor, collateral unit or purpose, and zero candidates means only that no record was matched.

Adjacency uses all observed known-unit sales before endpoints with ambiguous financing or out-of-band prices are dropped. Forming pairs within the estimation sample would bridge an intervening unit sale in 271 cases, so every row of Table OA2 forms adjacency on all known-unit sales. Base timing uniqueness leaves 1,765 pairs in 1,082 buildings, and its interval includes zero, as do the narrow and wide alternatives. That is consistent with a small co-operative contrast in selected, differently measured samples, and it is silent on whether the mortgage recording tax causes the house gap.

Table OA2: Co-operative assignment sensitivity

Assignment rule	Pairs	N_{cf}	N_{fc}	Joint	95% CI
Expanded count rule	5,839	247	602	-1.18	[-3.11, 0.75]
All competing sales	5,261	229	576	-0.89	[-2.77, 1.00]
Unique timing: narrow	2,840	441	537	-0.41	[-2.00, 1.18]
Unique timing: base	1,765	219	252	-1.42	[-3.70, 0.86]
Unique timing: wide	1,265	108	143	0.09	[-2.99, 3.17]

Log points and building-clustered CR0 intervals. Every row forms adjacency using all observed known-unit sales before dropping ineligible endpoints and ambiguous financing. Timing uniqueness does not establish debtor identity or purchase purpose. No matched filing is a cash proxy and does not prove an unlevered purchase. Estimates use the September 2026 co-operative extract.

OA4 Credit conditions at both sale dates

If credit conditions shape financing risk, attaching a rate only to the resale misses conditions at the initial transaction. The full house design therefore adds separate terms $F_a(r_a - 4)$ and $F_z(r_z - 4)$, where r is the latest available weekly Freddie Mac 30-year PMMS rate, in percent, on or before each sale. Because quarterly effects do not absorb all within-quarter weekly-rate variation, I also include separate uninteracted rates at both endpoints. The object of interest is the financing-specific association with rates.

The slopes are small and negative at the first date and positive at the second date (Table OA3). Adding separate financing-specific linear calendar trends reverses both signs, and the two rate interactions cannot be jointly distinguished from zero in the calendar-dependence sensitivity (Table OA4). Rate movements, time trends and financing composition are hard to disentangle in these data, so the rate evidence does not identify the execution-risk mechanism. The PMMS methodology change on 17 November 2022 adds a further measurement caveat. The augmented model's reference contrast is evaluated at 4 percent and at the trend origin, and it should not be read as an unconditional average financing gap.

Table OA3: Credit conditions at both sale dates

Financing × rate	Without trends	SE	With trends	SE
First sale	-0.43	1.95	0.46	1.93
Second sale	2.98	0.62	-1.24	1.03

Log points per one percentage point of PMMS rates, centered on 4 percent; 6,006 pairs. Both models include separate uninteracted weekly rates at both endpoints, the full geographic controls and both financing-switch indicators. The second adds separate endpoint financing-specific linear calendar trends. Standard errors cluster by parcel (CR0). Appendix Table OA4 reports calendar-dependence sensitivity. The headline contrast is estimated in the model without rate interactions.

OA5 Historical market conditions and shared calendar shocks

The inventory extension uses a single indicator, the log base two of StreetEasy's prior-month borough sales inventory in its Single Family category, centred on each borough's mean log inventory over December 2015–November 2025, so a unit is a doubling. The September 2026 vintage covers both endpoints of all 6,006 pairs. Listing coverage need not match the administrative one-to-three-family sample, and inventory reflects demand, supply and

platform participation as well as anything related to closing risk. Because of revisions and unknown publication dates, a prior-month value may not have been public on the sale date.

Both endpoint inventory levels enter uninteracted, alongside separate financing-inventory interactions, the full house controls and endpoint financing-by-borough controls with Brooklyn omitted (the last keep fixed geographic differences in financing out of the inventory terms). I then add financing-specific linear calendar trends as a sensitivity check. I wrote this plan after checking coverage and before fitting, without registering it externally, and report both specifications.

Table OA4: Historical market conditions: endpoint and trend sensitivity

Model	First	95% CI	Second	95% CI	p_Q	p_Y
Credit	-0.43	[-5.91, 5.05]	2.98	[2.02, 3.94]	< 0.001	< 0.001
Credit + trends	0.46	[-4.63, 5.54]	-1.24	[-3.18, 0.70]	0.418	0.330
Inventory	-3.13	[-7.17, 0.91]	-13.38	[-22.54, -4.21]	0.005	0.003
Inventory + trends	-4.51	[-9.79, 0.77]	3.21	[-6.65, 13.08]	0.195	0.244

All 6,006 pairs; slopes in log points. Credit unit: one percentage point of the weekly rate. Inventory unit: a doubling of prior-month borough Single Family inventory relative to its borough reference. Inventory models also include separate endpoint financing-by-borough controls (Brooklyn reference). Both indicators enter uninteracted at both dates. Intervals use the union of parcel and either-endpoint calendar-quarter score dependence and a t_{39} reference. Columns p_Q and p_Y test both interaction slopes jointly, using calendar-quarter and calendar-year dependence respectively, with approximate $F_{2,39}$ and $F_{2,9}$ references. Negative covariance eigenvalues are set to zero: one is corrected in each quarter model with trends and in each year model. The unadjusted eigenvalues are recorded in the replication output. These are exploratory, pointwise diagnostics without exact small-cluster inference.

Parcel-only uncertainty is in the replication output, and the table also allows dependence whenever two pairs share a parcel or either endpoint calendar period, whichever sale the period belongs to, counting overlapping memberships once. The calendar-quarter calculation has 40 nodes and the calendar-year version ten. Negative eigenvalues in the covariance are set to zero and their magnitudes reported. The t and F references are approximate, intervals are pointwise, and dependence across non-overlapping periods is left out.

With trends, the inventory interactions are jointly indistinguishable from zero under either calendar aggregation, and the credit exercise gives the same answer once common weekly rates are included. The associations are sensitive to the time specification, which is weaker than showing that market conditions are irrelevant. The headline contrast comes from the main model and should not be confused with the reference contrast of either interaction model.

OA6 Separate-band property-value and assessment splits

Table OA5 reports the separate-band companion to Table 7 of the paper, refitting all ZIP-year, borough-quarter, and party controls separately inside each price or pre-sale assessment band across the same 999 common parcel-bootstrap draws.

Table OA5: Financing contrast by property-value band (separate-band controls, 999 draws)

Band definition	Switching pairs all / screened	All pairs [95% CI]	Surname screen [95% CI]	Paired change [95% CI]
Geometric mean of sale prices ($\geq$ \$1M)				
Below \$1 million	1,175 / 1,059	13.44 [10.32, 15.57]	7.38 [4.48, 9.83]	−6.06 [−7.45, −4.10]
At least \$1 million	606 / 575	2.72 [−1.75, 6.09]	1.27 [−3.27, 4.20]	−1.46 [−4.00, 0.64]
First-sale price ($\geq$ \$1M)				
Below \$1 million	1,323 / 1,207	9.63 [7.23, 11.88]	5.32 [3.00, 7.57]	−4.31 [−5.57, −2.89]
At least \$1 million	458 / 427	3.43 [−1.45, 8.42]	0.38 [−4.39, 5.37]	−3.05 [−5.66, −0.31]
Prior DOF market value ($\geq$ \$1M, $N = 5{,}992$)				
Below \$1 million	1,431 / 1,321	8.80 [6.64, 11.01]	4.94 [2.83, 7.07]	−3.86 [−5.03, −2.65]
At least \$1 million	343 / 307	8.48 [−0.80, 19.52]	5.00 [−4.43, 17.68]	−3.48 [−10.33, 4.48]
Prior DOF market value sensitivity ($\geq$ \$850k, $N = 5{,}992$)				
Below \$850,000	1,254 / 1,161	7.99 [5.74, 10.21]	4.53 [2.32, 6.83]	−3.46 [−4.60, −2.27]
At least \$850,000	520 / 467	11.19 [5.13, 16.76]	5.82 [0.31, 11.64]	−5.37 [−9.76, −0.99]

Log points across 999 common parcel-bootstrap draws (8,127 clusters, seed 20260909), refitting all ZIP-year, borough-quarter, and party controls separately within each band. Sparse ZIP-year cells among the 871 pairs at or above \$1 million in prior `FullVal` widen those two separate-band intervals to 20.32 and 22.11 log points; they are moved here under the 20-log-point main-table precision rule.

OA7 House building class and loan-to-value mechanism checks

Exact matching on parcel identifier (BBL), sale date, and rounded sale price links all 8,208 house pairs (and all 6,006 headline pairs) to the NYC Department of Finance building-class categories and ACRIS purchase-window mortgage amounts in the linked dataset (Loschi, 2026). Recorded financing flags and mortgage amounts agree on 100 percent of financed endpoints with the original house panel (`panel2.csv.gz`), and the compact extract `scripts/house_mechanism_attributes.csv.gz` allows offline verification alongside `scripts/MECHANISM_ATTRIBUTES_MANIFEST.json`.

One-family versus two-to-three-family houses. By first-sale building-class category, the headline sample contains 2,951 one-family pairs (category 01) and 3,055 two-to-three-family pairs (categories 02 and 03), with 69 pairs crossing between categories across sales. Panel A of Table OA6 shows that the screened financing contrast is at least as large among one-family houses as among two-to-three-family houses across pooled headline controls (5.70 [3.19, 8.07] versus 4.11 [1.12, 6.89]), constant-class pairs (5.39 versus 4.13), separate-group headline controls (4.95 versus 4.82), and long-hold ($\geq$ 5 years) trimmed specifications. Reusing the exact 999 union-of-parcels bootstrap draws of Section OA1.2, one-family houses alone yield 4.92 [2.84, 7.64] in the long-hold trimmed borough-quarter comparison, exceeding condominiums by 5.95 log points [3.76, 8.66], against 4.67 [2.17, 7.23] for two-to-three-family houses (paired difference -0.26 [-3.76, 3.03]).

Recorded loan-to-value (LTV) gradient. Following the standard validity screen ($0.05 < \text{LTV} \leq 1.05$) used for recorded mortgage amounts, 9,172 of 9,349 financed endpoints in the headline sample (98.1 percent) carry valid LTVs (median 77.4 percent; 3 missing or non-positive, 26 at or below 5 percent, and 148 above 105 percent). Among 1,634 screened switching pairs, 1,586 have valid LTVs; the 48 out-of-range switching pairs are absorbed in a separate indicator pair so the sample and nuisance design remain fixed (dropping them alters the $\leq 80\%$ and $> 80\%$ point estimates by less than 0.01 log points). Panel B of Table OA6 reports binary, four-bin, and continuous-slope fits under both screened OLS and 2.5th/97.5th annualized-growth trimming (cutoffs recomputed inside every bootstrap draw without financing labels). Because $\text{LTV} = M/P$ places the recorded sale price in the denominator (correlation -0.44 with log first-sale price and -0.44 with log resale price), transitory price variation generates a mechanical negative LTV slope, confirmed directly within 3,505 always-financed pairs (-0.74 [-1.19, -0.43] per 10 percentage points of ΔLTV in OLS; -0.51 [-0.81, -0.23] trimmed). Against that downward mechanical tilt, both OLS and trimmed estimates exhibit a U-shaped profile with a sharp concentration above 95 percent LTV (8.36 [3.57, 13.58] OLS; 10.07 [5.94, 14.60] trimmed), where 124 of 196 screened switching pairs sit between 97.5 and 98.5 percent LTV (consistent with FHA 96.5 percent financing plus financed upfront mortgage insurance). Excluding switching pairs above 95 percent LTV still leaves a positive screened house contrast (4.50 [2.25, 6.49] OLS; 1.88 [0.38, 3.40] trimmed).

Direction asymmetry ($C\to F$ versus $F\to C$) and party anatomy across LTV bands. Table OA7 separates the cash-to-financed ($C\to F$) and financed-to-cash ($F\to C$) switching arms within each LTV band on the same 999 common parcel-bootstrap draws used in Ta-

Table OA6: House building-class splits and financed-endpoint loan-to-value (LTV) checks

Panel A. One-family versus two-to-three-family houses (surname-screened)

Specification	1-family	2–3-family	2–3 minus 1-fam	1-fam minus condo
Headline pooled ($N = 5,648$)	5.70 [3.19, 8.07]	4.11 [1.12, 6.89]	−1.59 [−5.20, 2.01]	–
Constant class ($N = 5,584$)	5.39 [2.87, 7.80]	4.13 [1.16, 6.88]	−1.26 [−4.85, 2.39]	–
Headline separate controls	4.95 [2.41, 7.45]	4.82 [1.27, 8.17]	−0.14 [−4.62, 3.70]	–
Headline ≥ 5 yr trimmed	5.02 [1.51, 7.68]	2.96 [−0.36, 6.34]	−2.06 [−6.26, 2.90]	–
Common BQ, untrimmed	9.56 [7.53, 11.83]	5.81 [3.69, 8.15]	−3.76 [−7.00, −0.72]	9.93 [7.78, 12.33]
Common BQ, ≥ 5 yr trimmed	4.92 [2.84, 7.64]	4.67 [2.17, 7.23]	−0.26 [−3.76, 3.03]	5.95 [3.76, 8.66]

Panel B. Financed-endpoint LTV gradient in the headline screened sample ($N = 5,648$)

LTV band / statistic	OLS N	Screened OLS	Trim N	Screened 2.5/97.5 trim
LTV $\leq 80\%$	1,151	6.44 [3.87, 8.51]	1,040	2.19 [0.56, 3.75]
LTV $> 80\%$	435	3.42 [0.27, 7.26]	397	4.85 [2.31, 8.53]
Diff. ($> 80\%$ minus $\leq 80\%$)		−3.02 [−6.50, 1.73]		2.66 [−0.08, 6.66]
LTV $(5\%, 70\%]$	626	9.27 [6.14, 11.92]	570	3.31 [1.16, 5.08]
LTV $(70\%, 80\%]$	525	2.78 [−0.46, 6.09]	470	0.87 [−1.85, 3.44]
LTV $(80\%, 90\%]$	148	−1.37 [−7.00, 4.68]	137	0.06 [−4.17, 5.22]
LTV $(90\%, 105\%]$	287	6.03 [2.20, 10.52]	260	7.52 [4.51, 11.59]
LTV $(80\%, 95\%]$ (with $> 95\%$)	239	−1.16 [−5.87, 3.88]	223	−0.04 [−3.38, 4.50]
LTV $(95\%, 105\%]$	196	8.36 [3.57, 13.58]	174	10.07 [5.94, 14.60]
Diff. ($> 95\%$ minus 70–80%)		5.56 [0.11, 11.39]		9.19 [4.55, 14.35]
Switch slope (per +10 pp LTV)	5,498	−1.20 [−2.15, −0.10]		0.35 [−0.38, 1.20]
Always-fin. FF (per +10 pp ΔLTV)	3,505	−0.74 [−1.19, −0.43]		−0.51 [−0.81, −0.23]

Log points; 95% percentile intervals from 999 common parcel-bootstrap draws refitting all controls and growth-trim cutoffs in each draw. Panel A classifies houses by first-sale NYC DOF building-class category (01 versus 02–03; 100% exact match on BBL, sale date, and price against the linked sales extract). Screened pair counts are 2,776 (1-family) and 2,872 (2–3-family) in the headline sample; 3,762 and 3,997 in untrimmed common borough-quarter (BQ); and 1,729 and 1,856 in $\geq$ 5-year trimmed BQ. The common BQ rows reuse the union-of-parcels draws of Online Appendix OA1.2, where the screened condominium contrast is -0.37 untrimmed and −1.03 in the $\geq$ 5-year trimmed sample (2–3-family minus condo is 6.17 [4.03, 8.55] untrimmed and 5.70 [3.16, 8.53] long-hold trimmed). Panel B uses the largest purchase-window mortgage divided by sale price, keeping switching pairs with $0.05 < \text{LTV} \leq 1.05$ (1,586 of 1,634 screened switching pairs) and absorbing the 48 out-of-range switching pairs in a separate indicator pair. In Panel B, N reports switching pairs in band rows and total screened pairs with valid LTV at every financed endpoint in the two slope rows ($5,498$ overall; $3,505$ always-financed pairs).

ble OA6. Above 95 percent LTV, the two arms move strongly in opposite directions and are close to symmetric in magnitude: 9.52 [3.00, 16.31] ($C \to F$) versus −7.19 [−14.87, −0.22] ($F \to C$) in screened OLS (sum 2.33 [−7.82, 11.78]), and 9.33 [2.97, 15.13] versus −10.82 [−17.78, −5.11] under 2.5th/97.5th growth trimming (sum −1.49 [−11.05, 6.78]). In the conventional 70–80 and 80–95 percent bands, by contrast, both arms are positive under trimming (5.12 and 3.35; 3.22 and 3.30), yielding half-differences near zero. Panel C shows that corporate-entity sellers are more prevalent in > 95 percent pairs on both sides of the FHA segment (63.7 versus 53.3 percent in $C \to F$, with median holding period 4.40 versus 4.84 years; 40.4 versus 18.6 percent at the initial financed purchase in $F \to C$). Splitting > 95 percent switching pairs by whether the grantor at the financed sale is a corporate entity confirms that the symmetric trimmed contrast appears in both entity-seller (12.79 [6.29, 18.43]; arms 11.39 and −14.18) and non-entity-seller (6.94 [1.88, 14.04]; arms 5.29 and −8.59) transactions (paired difference 5.85 [−3.71, 12.80]).

Condominium versus house LTV incidence and conventional-only comparisons. Exact matching on unit parcel (`linked_bbl`), sale date, and rounded price links all 13,672 of 13,672 corrected condominium endpoints (100.0 percent; 13,670 of 13,672 endpoints, 99.99 percent, agree on the base financing flag, and duplicate DOF rows show zero mortgage-amount conflicts) to `nyc_linked_dataset_v1_2_0` (bundled offline in `scripts/condo_ltv_attributes.csv.gz` with `scripts/CONDO_LTV_MANIFEST.json`). Panel A of Table OA8 documents a stark incidence gap: valid LTVs ($0.05 < \text{LTV} \leq 1.05$) exceed 95 percent in 16.0 percent of financed endpoints in the 8,208-pair house sample (10.3 percent in the 97.5–98.5 percent FHA peak) and 16.5 percent in the headline house sample, versus 1.0 percent (0.15 percent in the FHA peak) in condominiums. Among screened switching pairs, 12.4 percent of houses (177 pairs in the FHA peak) exceed 95 percent LTV against 0.9 percent of condominiums (1 pair in the FHA peak). Panel B refits the paired borough-quarter comparison of Section OA1.2 across the same 999 union-of-parcels draws under four LTV filters. Excluding > 95 percent switching pairs in the long-hold (≥ 5 years) trimmed comparison lowers the house contrast from 4.78 to 3.94 [2.29, 5.97] (paired change −0.85 [−1.50, 0.01]) and trims the house-minus-condominium gap from 5.81 to 4.93 [3.28, 7.12] (paired gap change −0.88 [−1.51, −0.01]). Restricting switching pairs to conventional 70–95 percent LTV yields 3.50 [1.28, 5.69] for houses and −2.24 [−3.13, −1.33] for condominiums, a difference of 5.74 [3.18, 8.20] (5.37 [2.99, 8.18] when absorbing non-conventional switches). In untrimmed common borough-quarter fits, the conventional 70–95 percent difference is 5.89 [3.47, 8.38] (4.64 versus −1.26). High-LTV FHA loans therefore account for a measurable share of the house contrast, but a large conventional-only house-versus-condominium divergence remains.

Table OA7: Direction decomposition ($C{\to}F$ versus $F{\to}C$) and party anatomy across LTV bands

Panel A. Switching arms ($C{\to}F$, $F{\to}C$), contrast, and sum, screened OLS ($N = 5,648$)					
LTV band	N_{CF}/N_{FC}	$C{\to}F$	$F{\to}C$	$\frac{1}{2}(CF{-}FC)$	$CF{+}FC$
LTV $(5\%, 70\%]$	367/259	5.69 [1.58, 8.81]	−12.91 [−17.94, −8.13]	9.30 [6.13, 11.94]	−7.22 [−13.88, −1.69]
LTV $(70\%, 80\%]$	326/199	8.28 [3.85, 12.43]	2.69 [−2.24, 7.80]	2.80 [−0.48, 6.10]	10.97 [4.20, 17.80]
LTV $(80\%, 95\%]$	158/81	2.62 [−2.67, 8.60]	4.93 [−3.26, 12.98]	−1.16 [−5.87, 3.88]	7.56 [−2.30, 17.99]
LTV $(95\%, 105\%]$	102/94	9.52 [3.00, 16.31]	−7.19 [−14.87, −0.22]	8.36 [3.57, 13.58]	2.33 [−7.82, 11.78]
$> 95\%$: entity seller	65/38	5.57 [−2.62, 13.89]	−15.00 [−24.83, −5.12]	10.28 [3.48, 17.04]	−9.43 [−22.10, 3.32]
$> 95\%$: non-entity	37/56	15.06 [4.35, 26.11]	−1.25 [−11.88, 7.67]	8.16 [1.37, 15.40]	13.81 [−1.93, 28.14]
Panel B. Switching arms, contrast, and sum, screened 2.5/97.5 growth trim ($N = 5,366$)					
LTV band	N_{CF}/N_{FC}	$C{\to}F$	$F{\to}C$	$\frac{1}{2}(CF{-}FC)$	$CF{+}FC$
LTV $(5\%, 70\%]$	343/227	2.62 [−0.21, 5.15]	−4.03 [−6.65, −0.96]	3.33 [1.19, 5.11]	−1.41 [−5.53, 2.74]
LTV $(70\%, 80\%]$	290/180	5.12 [1.74, 8.06]	3.35 [−0.98, 7.52]	0.88 [−1.83, 3.45]	8.47 [3.02, 13.56]
LTV $(80\%, 95\%]$	148/75	3.22 [−0.86, 8.23]	3.30 [−3.93, 9.48]	−0.04 [−3.38, 4.50]	6.52 [−1.59, 14.44]
LTV $(95\%, 105\%]$	87/87	9.33 [2.97, 15.13]	−10.82 [−17.78, −5.11]	10.07 [5.94, 14.60]	−1.49 [−11.05, 6.78]
$> 95\%$: entity seller	56/35	11.39 [2.34, 18.53]	−14.18 [−23.05, −4.98]	12.79 [6.29, 18.43]	−2.79 [−15.21, 8.69]
$> 95\%$: non-entity	31/52	5.29 [−2.83, 15.21]	−8.59 [−17.83, −1.09]	6.94 [1.88, 14.04]	−3.30 [−16.37, 9.26]
Panel C. Party composition and holding period of screened switching pairs					
Arm and LTV band	N	Fin. ent. seller	Cash ent. buyer	Med. hold	Hold < 5 yr
$C{\to}F$, $(5\%, 70\%]$	367	38.4%	37.9%	4.96 yr	50.7%
$C{\to}F$, 70–95% conv.	484	53.3%	53.9%	4.84 yr	53.7%
$C{\to}F$, $> 95\%$ FHA	102	63.7%	59.8%	4.40 yr	63.7%
$F{\to}C$, $(5\%, 70\%]$	259	21.2%	38.2%	5.18 yr	46.3%
$F{\to}C$, 70–95% conv.	280	18.6%	46.1%	4.74 yr	53.6%
$F{\to}C$, $> 95\%$ FHA	94	40.4%	61.7%	5.58 yr	38.3%

Log points; 95% percentile intervals from the same 999 common parcel-bootstrap draws as Table OA6 (headline ZIP-year, borough-quarter, and buyer/seller party controls; all four bands estimated jointly under the 70/80/95 split; growth-trim cutoffs recomputed inside every draw without financing labels). Panels A and B report the separate cash-to-financed ($C{\to}F$) and financed-to-cash ($F{\to}C$) arm coefficients, their half-difference contrast $\frac{1}{2}(\hat{\beta}_{CF}{-}\hat{\beta}_{FC})$, and their sum $\hat{\beta}_{CF}{+}\hat{\beta}_{FC}$. Out-of-range switching pairs (LTV ≤ 0.05 or > 1.05) are absorbed separately. The entity-seller rows split $(95\%, 105\%]$ switching pairs by whether the grantor at the financed endpoint (z for $C{\to}F$; a for $F{\to}C$) is a corporate entity; the trimmed entity-minus-non-entity contrast difference is 5.85 [−3.71, 12.80]. Panel C summarizes unadjusted shares of corporate-entity sellers at the financed sale, corporate-entity buyers at the cash sale, and holding-period quantiles in the screened headline sample.

Table OA8: House versus condominium LTV incidence and conventional-only paired comparisons

Panel A. Recorded LTV distribution across valid endpoints ($0.05 < LTV \leq 1.05$*)*

Sample	Valid N	$\leq 70\%$	70–80%	80–95%	$> 95\%$ (97.5–98.5%)
Houses (headline), fin. endpts.	9,172	37.5%	30.7%	15.2%	16.5% (10.82%)
Houses (BQ 8,208), fin. endpts.	12,004	39.0%	30.4%	14.6%	16.0% (10.27%)
Condos (BQ 6,836), fin. endpts.	7,890	46.4%	43.4%	9.2%	1.0% (0.15%)
Houses (BQ), scr. switches	2,436	40.4%	32.8%	14.5%	12.4% (7.27%)
Condos (BQ), scr. switches	2,907	46.4%	44.1%	8.6%	0.9% (0.03%)

Panel B. Paired house-versus-condominium contrasts under LTV filters (common BQ, 999 union draws)

Specification	Sw. H / C	Houses	Condos	House − condo	Δ gap
Long-hold (≥ 5 yr), within-type 2.5/97.5 growth-trimmed (surname-screened)					
Baseline (all switches)	1,063/1,469	4.78 [3.17, 6.63]	−1.03 [−1.64, −0.36]	5.81 [4.00, 7.77]	–
Excl. > 95% LTV switches	937/1,457	3.94 [2.29, 5.97]	−0.99 [−1.66, −0.33]	4.93 [3.28, 7.12]	−0.88 [−1.51, −0.01]
Valid ≤ 95% LTV only	909/1,451	4.04 [2.40, 6.01]	−0.99 [−1.63, −0.33]	5.03 [3.31, 7.15]	−0.78 [−1.55, 0.10]
Conv. 70–95% (restr.)	484/717	3.50 [1.28, 5.69]	−2.24 [−3.13, −1.33]	5.74 [3.18, 8.20]	−0.07 [−2.27, 1.91]
Conv. 70–95% (abs.)	490/715	3.15 [0.99, 5.77]	−2.22 [−3.11, −1.33]	5.37 [2.99, 8.18]	−0.44 [−2.12, 1.73]
Untrimmed common borough-quarter (surname-screened)					
Baseline (all switches)	2,518/2,917	7.50 [6.00, 9.10]	−0.37 [−1.07, 0.27]	7.86 [6.25, 9.60]	–
Excl. > 95% LTV switches	2,216/2,890	7.06 [5.40, 8.72]	−0.27 [−0.99, 0.36]	7.33 [5.57, 9.18]	−0.54 [−1.17, 0.07]
Conv. 70–95% (restr.)	1,151/1,530	4.64 [2.43, 6.89]	−1.26 [−2.36, −0.32]	5.89 [3.47, 8.38]	−1.97 [−3.92, −0.16]
Conv. 70–95% (abs.)	1,151/1,530	4.61 [2.37, 6.77]	−1.27 [−2.35, −0.31]	5.88 [3.56, 8.34]	−1.98 [−3.84, −0.15]

Panel A links all 13,672 of 13,672 corrected condominium endpoints (100.0%) exactly on unit BBL, sale date, and rounded price to `nyc_linked_dataset_v1_2_0` (13,670 of 13,672 endpoints, 99.99%, agree on the base financing flag; 2 early-2016 endpoints lack a matched dataset mortgage). Parentheses in the final column report the share inside the 97.5–98.5% FHA peak (1,233 of 12,004 house endpoints versus 12 of 7,890 condominium endpoints; 177 house switching pairs versus 1 condominium switching pair). Panel B reuses the exact 999 union-of-parcels bootstrap draws of Section OA1.2 (14,874 parcels), refitting borough-quarter controls and within-type 2.5/97.5 annualized-growth cutoffs in every draw. Restricted conventional rows keep always-financed and always-cash pairs for market controls while restricting switching pairs to $0.70 < \text{LTV} \leq 0.95$; absorbed rows keep the full BQ sample fixed and absorb non-70–95% switching pairs in separate arm indicators.

References

Han, L. and Hong, S.-H. (2024). Cash is king? understanding financing risk in housing markets. *Review of Finance*, 28(6):2083–2118.

Loschi, P. (2026). NYC residential sales linked to ACRIS deeds and mortgages, 2016–2025. Zenodo, dataset version 1.2.0.