

The Execution-Certainty Wedge

How Measurement Choices Rank Housing-Access Instruments

Pablo Loschi · Independent researcher

Version 6.5 · Working paper · 14 October 2026 · Extended version: v5.1

Abstract

Sellers can prefer cash offers even when financed buyers value the home more. This paper compares housing-access instruments in an agent-based market set to New York City proxies: institutions win about 11% of sales, close to the share of company cash buyers among linked house sales, and financed offers carry a 1.5% discount. Removing that discount lowers the share of owner-occupiers who leave without buying only from 9.8% to 9.5%, so in this market the certainty wedge explains little of the access shortfall. The paper's results concern measurement. An efficiency benchmark computed over the buyers still waiting moves with the policy. The duty raises the ratio against this benchmark above an offer discount of about 2.2%, while the fixed comparison shows a loss at every tested discount up to 20%. Access gains for a cohort followed until purchase or exit range from about three quarters to two and a half times the full-run gains and reorder the instruments. A constant charge keeps more value than the duty when both are tuned to full-run access, and the duty keeps more when both are tuned to cohort access. A quota that resets each period cannot reach small access targets, and allowances that carry forward can. A stress test in which institutions win 58% of sales gives the same four results at larger magnitudes.

Keywords: execution certainty; housing access; policy evaluation; efficiency measurement; agent-based model; transaction taxes

JEL classification: D47; D61; D63; H23; R31

1. Introduction

A seller choosing between two offers for the same home may take the lower one because it is more likely to close. Cash offers usually are. When the buyers who can pay cash are mainly investors and the buyers who need a mortgage are mainly households who would live in the home, the seller's preference can decide who gets housing. In May 2026 New York legislators considered a 1% tax on all-cash residential purchases of \$1 million or more. It was reported as likely to be dropped a week later and was left out of the budget passed on 27 May (Bloomberg, 14, 21 and 27 May 2026). The proposal raises a design question that does not depend on its fate: if a charge on one class of buyer is meant to change who wins, how should its effects be measured?

The preference for certainty is not in itself a distortion. A failed sale costs the seller time, carrying costs and a relisting, and discounting a riskier offer can maximise expected surplus. Reher and Valkanov (2024) estimate a national mortgage-cash premium larger than their model of actuarial risk explains. Han and Hong (2024) find a smaller discount in Los Angeles and show that closing and relisting risks together can account for it. The companion New York City study (Loschi 2026a) finds a financing-related price gap of about 9 log points for houses and close to zero for condominiums. Within houses the gap is small when the cash buyer is a company, and it falls by almost half when pairs with a same-surname deed, a marker of possible transfers between relatives, are excluded. The observed gap varies with transaction composition and screening rules. Those comparisons do not separately identify the price of closing certainty.

That evidence limits what a model of the certainty wedge can claim about New York. This paper asks a narrower question. Suppose certainty does decide some allocations. Which instruments change them, at what cost, and how would an evaluator know? The answer depends on measurement choices that are easy to make without noticing.

The model is a stylised market with owner-occupiers, small investors and institutional buyers. Sellers rank offers by valuation times a certainty factor, and institutional offers are the most certain. Four instruments are compared with an unconstrained market: a cap on acquisitions per legal entity; a work-conserving fair share that limits institutional purchases only while owner-occupiers are waiting; a constant proportional charge on

institutional offers; and a duty whose rate rises with a local contention signal, the share of owner-occupiers who leave without buying. A fixed-offer result, Proposition 1, shows which acquisitions a proportional charge displaces. The aggregate comparison is computational. The main market is set to New York City proxies, with institutions winning about 11% of sales and a 1.5% offer discount. Earlier versions of this paper used a market in which institutions win 58% of sales at a 12% discount, and it is kept here as a stress test.

Four findings concern measurement, and each holds in both markets. First, an efficiency benchmark computed over the buyers present at each sale depends on the policy, because the policy decides who is still waiting. In the main market a fixed benchmark shows the contention duty losing value at every offer discount up to 20%, while the moving benchmark shows it gaining from a discount of about 2.2%. A two-period example shows why: a policy that serves a high-value buyer at once removes that buyer from the queue on which the benchmark is computed. Second, the usual access statistic, exits without purchase over the whole run, counts buyers still searching at the end as neither served nor unserved. Following a cohort until every member has bought or left changes measured gains by factors from about 0.75 to 2.5, and the two measures rank the instruments differently. Third, the comparison between the contention duty and a constant charge depends on which access population the instruments are tuned to. The constant charge keeps more value at common full-run targets, and the duty keeps more at common cohort targets. Fourth, a fair-share quota that resets each period cannot reach small access targets, because it admits institutional purchases only in whole numbers each period. Allowances that carry forward reach them.

Two results concern the instruments themselves. In the main market the certainty wedge matters little: removing the discount entirely lowers the full-run share of owner-occupiers who leave without buying from 9.8% to 9.5%. The cap is the expensive way to buy access there, costing four to eight times as much realised value per point of access as the charges at common targets. The stress test shows the same pattern at a larger scale. The duty lowers the share leaving without buying from 41.9% to 36.6% with almost no change in realised value, while the cap lowers it to 8.7% at an 11.3% loss, and owner-occupier weights of about 1.66 or more make the cap the best of the three original instruments.

Section 2 places the paper in the literature. Section 3 sets out the model, the two markets and the measures, and Section 4 the seller's problem and the selection result. Section 5 reports the main market, organised by the measurement choice that changes the comparison. Section 6 repeats the comparisons in the stress test and shows how the break-even discount moves with institutional participation. Section 7 varies the New York City proxies. Section 8 sets out what the model leaves out, Section 9 what each instrument would need in practice, and Section 10 concludes. Appendix A records how the results can be reproduced. The extended version of this paper (Loschi 2026b) contains further robustness exercises for the stress-test market, a longer discussion of implementation and objections, and the full replication record.

2. Related work

The paper sits between housing finance, market design and public finance. Reher and Valkanov (2024) and Han and Hong (2024) supply the empirical motivation and disagree on how much of the cash premium is excess over risk; that disagreement is why this paper sweeps the offer discount instead of transporting one estimate. Che and Gale (1998) show that auctions with budget-constrained bidders can allocate inefficiently. The mechanism here runs through the seller's preference for certainty instead, and it operates even when the buyer's willingness to pay is unconstrained. Budish, Cramton and Shim (2015) study how market design can reward an advantage other than valuation, in their case speed. Completion certainty differs because it can save real costs. Piazzesi, Schneider and Stroebel (2020) show that housing search is segmented across clienteles; the model's districts are a crude version of that segmentation. Kopczuk and Munroe (2015) show that New York's mansion tax, a notch in a transfer tax, moves prices and volume around its threshold, so a charge on transactions changes the transactions on which it is later measured.

The policy question is a second-best one in the sense of Lipsey and Lancaster (1956): a charge in an already distorted market can raise or lower welfare, and nothing general follows. Comparing a quantity rule (a cap or quota) with a price rule (a charge) follows Weitzman (1974), and the carry-forward quota borrows the logic of banking allowances analysed in Weitzman (2018). The treatment of distribution follows the practice of stating welfare weights explicitly (Atkinson 1970) and of reporting the trade-off between efficiency and

fairness instead of assuming one objective (Bertsimas, Farias and Trichakis 2012). The method is agent-based in the sense of Tesfatsion (2006) and Axtell and Farmer (2025), who survey its use in economics; Geanakoplos et al. (2012) use an agent-based model of the housing market to study systemic risk. Bidding is a fixed rule, not a learning or equilibrium strategy, which is the model’s largest limitation.

The measurement findings have close relatives outside housing. Arora, Dekel and Tewari (2012) show why judging a policy against the history that another policy produced can mislead, and define policy regret to correct for it. Perdomo et al. (2020) study decision rules that change the distribution on which they are evaluated. Bertsimas, Farias and Trichakis (2011) measure the cost of fairness against a common feasible optimum, the role the fixed benchmark plays here. In queueing, a waiting-time statistic computed only over served requests hides the ones that were dropped. Allocation through queues has its own literature in housing: Waldinger (2021) uses public-housing wait lists in Cambridge, Massachusetts to show that allocation rules trade targeting against tenant welfare, Leshno (2022) shows that welfare in overloaded waiting lists depends on how expected waits ration units, and Thakral (2019) studies efficiency and fairness when units arrive at random. Those papers allocate public units by rule; here sellers choose among private offers, and the queue matters through who is still waiting when a listing arrives. None of these results is applied here; the contribution is to show how much these choices matter for ranking housing-access instruments. A companion note (Loschi 2026e) develops the benchmark diagnostic on its own, with a second example in which the disagreement runs the other way and a reusable audit routine.

3. The model

3.1 Market

The simulation is a discrete-time market with five districts over 120 periods, with illustrative supply, arrivals and valuations. Each district lists a Poisson-distributed number of units averaging 20 per period. Three classes of buyer bid:

- **Owner-occupiers** arrive at 17 per district per period, draw valuations from a lognormal distribution centred near 400,000 ($\sigma = 0.34$), and leave unserved after 36 periods.
- **Small investors** arrive at 4 per district per period, centred near 372,000 ($\sigma = 0.30$), with 24 periods of patience.
- **Institutional buyers** are three standing entities with valuations centred near 366,000 ($\sigma = 0.18$). In the main market each entity places one bid per district per period; in the stress test the three together place 42.

Sellers rank offers by effective value, the bid multiplied by a certainty factor: 1.00 for institutional cash offers, c_{OO} for financed owner-occupiers and $c_{SI} = \min(1, c_{OO} + 0.07)$ for small investors. The owner-occupier offer discount is $w = 1 - c_{OO}$: 1.5% in the main market and 12% in the stress test, and Sections 5.2 and 6.2 sweep it. A financed price premium p over the cash price corresponds to $w = p/(1 + p)$, and a log-price gap β to $w = 1 - \exp(-\beta)$. Units go to the highest effective offer in the eligible pool. No bargaining equilibrium is solved.

3.2 Two markets

The main market is set to New York City proxies from the companion study and dataset (Table 1). In the linked dataset, company cash buyers account for 19,019 of 166,227 house sales of at least \$100,000 with an ORDINARY linked deed, or 11.4%; all company buyers account for 22.8% (Loschi 2026d). One bid per entity per district-period gives an unconstrained institutional share of 11.0% in the model. The company-cash gap of 1.53 log points sets the offer discount to 1.5%; its upper confidence endpoint gives 5.3%. The price-gap interval includes zero, and Section 7 varies both inputs. The comparison is a calibration check, not a test of any tax.

These proxies are imperfect. Company buyers include LLCs used by individuals, the absence of a matched mortgage does not prove cash financing, and transaction-price gaps do not separate a certainty premium from composition or completion costs. Neither proxy identifies certain institutional buyers. The main market is therefore a realistic order of magnitude, not an estimate of New York City’s market, and supply, arrivals, patience and valuations remain illustrative in both markets.

The stress test keeps the calibration of earlier versions: 42 institutional bids per district-period, an unconstrained institutional share of 58.1% and a 12% offer discount. Institutional purchases run at a few per cent of

sales nationally, so this market is far from observed ones. It shows how the measurement results behave when institutions dominate and every instrument has large effects. Of the house groups in Table 1, only cash buyers other than companies show a discount near 12%, at 12.9%; excluding same-surname deeds reduces it to 6.9%. Mortgage denials, which precede accepted offers, cannot identify subsequent failure risk.

Market	Log-price gap, 100 β	95% interval	Approx. cash discount
One- to three-family houses, headline	9.25	[7.17, 11.35]	8.8%
Houses, excluding same-surname deeds	4.87	[2.69, 6.76]	4.8%
Houses, cash buyer is a company	1.53	[-1.79, 5.44]	1.5%
Houses, cash buyer is another party	13.77	[10.96, 15.98]	12.9%
Houses, other party, excluding same-surname deeds	7.11	[4.48, 9.12]	6.9%
Condominiums	0.77	[-0.01, 1.56]	0.8%
Co-operatives	-1.42	[-3.70, 0.86]	-1.4%

Table 1. Financing-related price gaps, New York City 2016–2025, in log points; positive values mean higher adjusted prices for financed purchases. Cash discounts are $100[1 - \exp(-\beta)]$. Source: Loschi (2026a); the figures are unchanged from version 3.4 to the current version 3.7. House intervals are parcel-bootstrap intervals; the co-operative interval clusters by building. Property types differ in sample construction, so the rows are not a matched comparison across tenures.

3.3 Instruments

Unconstrained. Sellers accept the highest effective offer.

Static per-entity cap. Each legal entity may acquire at most 50 units per district over the run.

Work-conserving fair share. Institutional buyers may take at most 30% of a district’s listings in a period while owner-occupiers are waiting. When none is waiting, the limit does not apply.

Contention-triggered duty. Institutional effective offers are multiplied by $(1 - \tau_{\max} \cdot s)$, with $\tau_{\max} = 0.22$. The signal s is measured on exits from the local owner-occupier cohort. Over a trailing window of six periods, let U be the number who left without buying and V the number who bought; then $s = U/(U + V)$, and $s = 0$ when the window is empty. Buyers still searching count in neither term. The signal counts exits because a bid-based version, the share of owner-occupier bids that lose, saturates near one in a market with hundreds of waiting buyers and twenty units per period, and would define a different instrument. In the main market the realised mean of s is 0.094 without intervention, 0.092 under the duty and 0.076 under the cap. In the stress test it runs from 0.078 under the cap to 0.406 without intervention, and is 0.341 under the duty. The multiplier is a reduced-form offer haircut. It is not a statutory tax rate, and no revenue is computed.

Constant charge. Institutional effective offers are multiplied by $(1 - h)$, whatever the contention. Its rate h is set either to the duty’s mean scheduled haircut or by tuning to an access target. It needs no registry of entries and exits, although classifying eligible buyers and enforcing the charge still require administration.

Combined. The fair share and the duty together.

3.4 Measures

The primary outcome is the **realised-value index**: the sum of successful buyers’ valuations, normalised to the unconstrained regime under matched seeds. It measures realised willingness to pay, not social welfare. It omits renters, the use of any revenue, administrative costs and differences in the marginal utility of income.

Two efficiency ratios put that sum over a benchmark. **Oracle efficiency** divides by the sum of the L highest valuations among all bidders generated in the run, where L is the number of units listed. With full search and exogenous listings, matched seeds give every policy the same denominator. The oracle sum ignores arrival, expiry and geography, so it is an upper bound F on the best feasible realised value V^* , with $V \leq V^* \leq F$. A ratio V/F near one therefore shows that little value is left unrealised under this objective, but it does not identify the exact feasible gap. **Contemporaneous efficiency** divides by the sum, over listings, of the highest valuation

among the bidders present at that listing. That denominator depends on who is still waiting, and so on the policy.

Access is measured by the **starvation rate**: owner-occupiers who leave without buying, divided by all owner-occupier arrivals during the run. Buyers still searching at the horizon stay in the denominator. Section 5.3 contrasts this with a completed-cohort measure. Results are means over twelve independent replications unless stated otherwise.

4. The seller's problem and which bids a charge displaces

4.1 How much discount risk justifies

A seller holds a unit worth P . A cash offer C closes at horizon τ_C . A financed offer F closes with probability $1 - q$ at $\tau_F > \tau_C$; if it fails, the seller incurs a relisting loss δ and eventually sells at P . With holding cost k per period and risk neutrality, the seller is indifferent when

$$(1 - q)(F - k\tau_F) + q(P - \delta - k\tau_F) = C - k\tau_C. \quad (1)$$

Setting $P = C$, $d = \delta/C$ and $\kappa = k(\tau_F - \tau_C)/C$, the justified discount on the financed price, $\pi^* = (F - C)/F$, is

$$\pi^* = (qd + \kappa) / (1 - q + qd + \kappa). \quad (2)$$

It depends on the failure probability, the cost of a failed sale and the difference in closing times, and it is positive even when no sale fails (at $q = 0$ it equals $\kappa/(1 + \kappa)$). Only a discount beyond π^* is a candidate distortion, and neither a price gap nor a mortgage denial rate identifies π^* without local evidence on accepted offers that fail.

4.2 Which bids a proportional charge displaces

Consider one listing with fixed offers and eligibility. A targeted buyer j has valuation v_j and certainty factor c_j ; an untargeted buyer i has v_i and c_i . Moving the unit from j to i changes realised valuation by $-G$, where $G = v_j - v_i$. This local measure ignores real completion costs.

Proposition 1 (fixed-offer selection). *Let $a = v_j c_j$ be the highest effective offer among targeted buyers and $b = v_i c_i$ the highest among untargeted buyers. Suppose $a \geq b > 0$, ties go to the targeted group, and offers and eligibility are fixed. Multiplying every targeted offer by $(1 - \theta)$, with $0 < \theta < 1$, changes the winning group exactly when*

$$r = 1 - b/a < \theta. \quad (3)$$

If $c_j > c_i$ and $\theta \leq 1 - c_i/c_j$, every such displacement has $v_i > v_j$.

Proof. Scaling all targeted offers by the same factor preserves their ranking, so the best targeted offer after the charge is $(1 - \theta)a$. The untargeted group wins if and only if $b > (1 - \theta)a$, which is (3). A displacement implies $v_i/v_j > (1 - \theta)c_j/c_i$, which is at least one under the stated condition. ■

The proposition holds offers fixed. A strategic extension must specify bidders' information, payoffs, the charge's treatment and tie-breaking. Public knowledge of valuations and score weights bounds the offers a buyer can support; it does not by itself supply an equilibrium. With private valuations and asymmetric bidders, first-price shading can change which buyer wins (Maskin and Riley 2000; Hafalir and Krishna 2008). Condition (3) describes displacement of given offers, while the response of those offers to a charge requires a separate bidding model. A duty whose rate follows a published signal adds a further response, choosing where and when to bid; Section 8 reports a limited version of it.

The charge therefore selects listings by their normalised effective-offer margin. This does not rank its average loss against a cap's. Take two equally likely listings with targeted and untargeted valuations (10,000, 9,700) and (100, 80), and certainty factors 1 and 0.88. A 20% charge changes the winner only at the first listing, at a valuation loss of 300. Excluding the targeted buyer with probability one-half at each listing gives the same

expected number of displacements and a conditional average loss of 160. Aggregate comparisons have to be computed.

In the stress-test simulation they come out as follows. Removing the current instrument from the candidate pool at each listing and comparing winners gives a mean G of +143,789 for the cap, +22,188 for the fair share and -12,905 for the duty, with 9.0%, 32.8% and 80.4% of displacements replacing the winner with a higher-value buyer. The duty mostly moves units to buyers who value them more; the cap mostly does the opposite. These local comparisons cannot be summed into a policy effect, because a buyer served at one listing might otherwise have bought later.

5. Results in the main market

5.1 The baseline comparison

Table 2 compares the instruments at their default settings in the main market. Without intervention, full-run starvation is 9.8%, against 41.9% in the stress test.

Regime	OO share	Inst. share	Unserved, full run	Unserved, cohort	Value index	Oracle eff.	Contemp. eff.
Unconstrained	72.3%	11.0%	9.83%	14.87%	100.00	99.54%	99.93%
Static per-entity cap	76.1%	6.3%	8.06%	10.47%	98.64	98.18%	94.53%
Work-conserving fair share	72.3%	11.0%	9.83%	14.87%	100.00	99.54%	99.93%
Contention-triggered duty	72.7%	10.5%	9.63%	14.42%	99.95	99.49%	99.91%
Constant charge	72.6%	10.6%	9.60%	14.50%	99.95	99.49%	99.95%

Table 2. Allocation regimes in the main market at default settings, means of 12 replications. Shares are of units bought. Full-run unserved is the starvation rate; the cohort follows owner-occupiers entering in periods 36 to 83 until they buy or leave by period 119. The constant charge equals the duty ceiling times the duty's mean signal ($h = 0.020$). The fair share's default limit of 30% of listings does not bind in any replication, because the three entities place at most three bids per district-period; its row equals the unconstrained market, and the combined regime equals the duty.

The default instruments do little here. The duty and the constant charge each lower full-run starvation by about 0.2 points and cost about 0.05% of realised value. The cap lowers it by 1.77 points at a 1.36% loss: the cap costs 0.77% of realised value per percentage point of full-run access, against 0.34% at baseline in the stress test. Among entrants followed to purchase or exit, unserved rates are 14.9% without intervention, and the duty and constant charge improve access by 0.45 and 0.37 points. The constant charge does slightly better over the full run (0.24 against 0.20 points) and the duty slightly better in the cohort. Both differences are under 0.1 point, although each exceeds three paired Monte Carlo standard errors.

The certainty wedge explains little of the shortfall. With the offer discount set to zero, full-run starvation remains 9.5%, against 9.8% at the company-cash gap, and the duty still improves access by 0.18 points. Institutional purchases still matter: excluding institutions altogether lowers full-run starvation by 6.6 points on the four tuning seeds of Section 5.3. What drives unserved exits in this market is how many units institutions buy, and the charges work mainly by limiting those purchases. Small investors retain the convention $c_{SI} = \min(1, c_{OO} + 0.07)$, making them as certain as institutions at the company-gap inputs. Setting their certainty equal to owner-occupiers' reduces the duty's full-run gain from 0.20 to 0.18 points.

5.2 Which efficiency benchmark

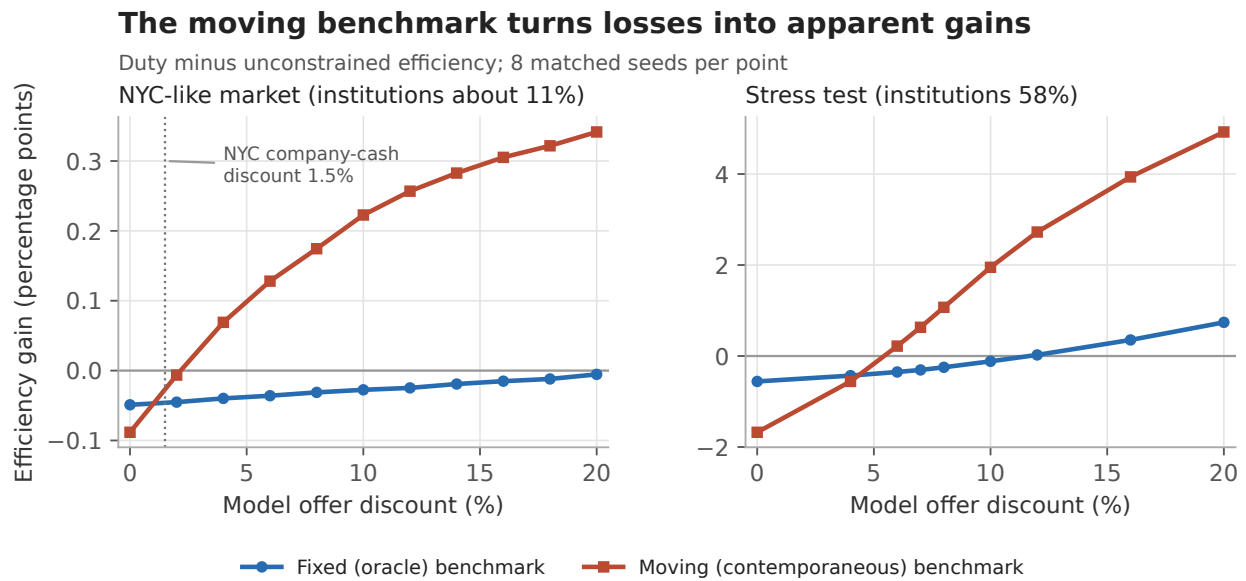
The two efficiency ratios in Table 2 differ systematically. In a market where buyers wait, the policy decides who is waiting. A regime that serves high-value buyers quickly leaves a thinner queue, and a benchmark computed over that queue falls with it. In the main market the contemporaneous benchmark averages about 444,000 per unit without intervention and under the duty, and 463,000 under the cap. Ratios against it are not comparable across regimes.

A two-period example shows the mechanism. One unit is sold in each period. In period 1 an institution with valuation v_I and an owner-occupier with valuation $v_O > v_I$ are present; the institution offers with certainty 1 and leaves if not served, and the owner-occupier's effective offer is $c \cdot v_O < v_I$, but the owner-occupier stays for period 2. In period 2 a late owner-occupier with valuation $v_L < v_I$ arrives. Without intervention the institution wins period 1 and the first owner-occupier wins period 2, for a realised value of $v_I + v_O$. A duty large enough to make $(1 - \theta)v_I < c \cdot v_O$ gives the owner-occupier the unit in period 1; the institution leaves, and the late buyer wins period 2, for $v_O + v_L$. The duty lowers realised value by $v_I - v_L$, and the fixed oracle benchmark, $v_O + v_I$ in both regimes, shows the loss. The moving benchmark counts the highest valuation present at each sale: $2v_O$ without intervention, because the waiting owner-occupier is present at both sales, and $v_O + v_L$ under the duty. Contemporaneous efficiency therefore rises from $(v_I + v_O)/(2v_O)$ to 1. With $v_O = 10$, $v_I = 6$, $v_L = 1$ and $c = 0.5$, it rises from 0.80 to 1.00 while realised value falls from 16 to 11. The policy removed a high-value buyer from the queue, and the benchmark fell with it.

Table 3 and Figure 1 show the consequence in the main market. On the fixed benchmark the duty loses value at every offer discount from 0% to 20%. The loss shrinks from 0.049 to 0.005 points as the discount rises, and in 3,943 of 4,000 bootstrap resamples of the eight seeds the gain never turns positive on that grid. On the moving benchmark the duty appears to add value from a discount of about 2.2% (95% interval [2.1, 2.3]). Above that point the two measures disagree on the sign of the policy's effect. At the upper company-cash discount of 5.3% and at the headline 8.8%, the moving benchmark would report a gain that the fixed benchmark does not; at the company-cash discount of 1.5% both report a small loss.

Offer discount w	Inst. share	Gain, oracle (pp)	Gain, contemporaneous (pp)	Value index, duty
0%	10.7%	-0.049	-0.088	99.95
2%	11.0%	-0.045	-0.006	99.95
4%	11.3%	-0.040	+0.069	99.96
6%	11.6%	-0.036	+0.128	99.96
8%	11.9%	-0.031	+0.174	99.97
10%	12.3%	-0.028	+0.223	99.97
12%	12.6%	-0.025	+0.257	99.98
16%	13.2%	-0.015	+0.305	99.98
20%	13.8%	-0.005	+0.342	99.99

Table 3. The duty's efficiency gain over the unconstrained market in the main market as the owner-occupier offer discount varies, 8 replications per row, duty ceiling 0.22. At each row $c_{OO} = 1 - w$ and $c_{SI} = \min(1, c_{OO} + 0.07)$; institutional certainty stays at one. The institutional share is the unconstrained share at that discount. Standard errors of the gains are below 0.008 points; all eleven grid nodes are in `realistic_market_results.json`.



Lines join simulation means. Left: the fixed benchmark shows a loss at every discount up to 20%. Right: the published stress-te

Figure 1. The duty's efficiency gain on the fixed (oracle) and moving (contemporaneous) benchmarks as the offer discount varies, in the main market (Table 3) and in the stress test (Section 6.2). The horizontal axis is the model's offer discount, not an empirical price premium; lines join simulation means.

The fix is to hold the benchmark independent of the policy. The value index needs only identical arrivals across regimes, and oracle efficiency needs the full arrival stream. A simple diagnostic makes the problem visible: report the benchmark's value per unit alongside each ratio. When this denominator differs across treatment arms, a change in the ratio cannot be attributed solely to total realised value.

5.3 Which access population

The default settings differ in strength as well as design, so the comparison calibrates each instrument to common access targets. Achieved gains on the evaluation seeds remain visible because they need not match exactly. Each was tuned on four simulation seeds to lower full-run starvation by 0.25, 0.5 or 1 percentage point, and the chosen settings were evaluated on sixteen separate seeds (Table 4, Figure 2). The fair share appears twice, with the period-reset allowance of Section 3.3 and with allowances that carry forward (Section 5.4). The duty cannot reach the one-point target: at its highest tested ceiling, 0.95, it lowers full-run starvation by 0.88 points on the tuning seeds, because its signal stays low when few owner-occupiers leave unserved.

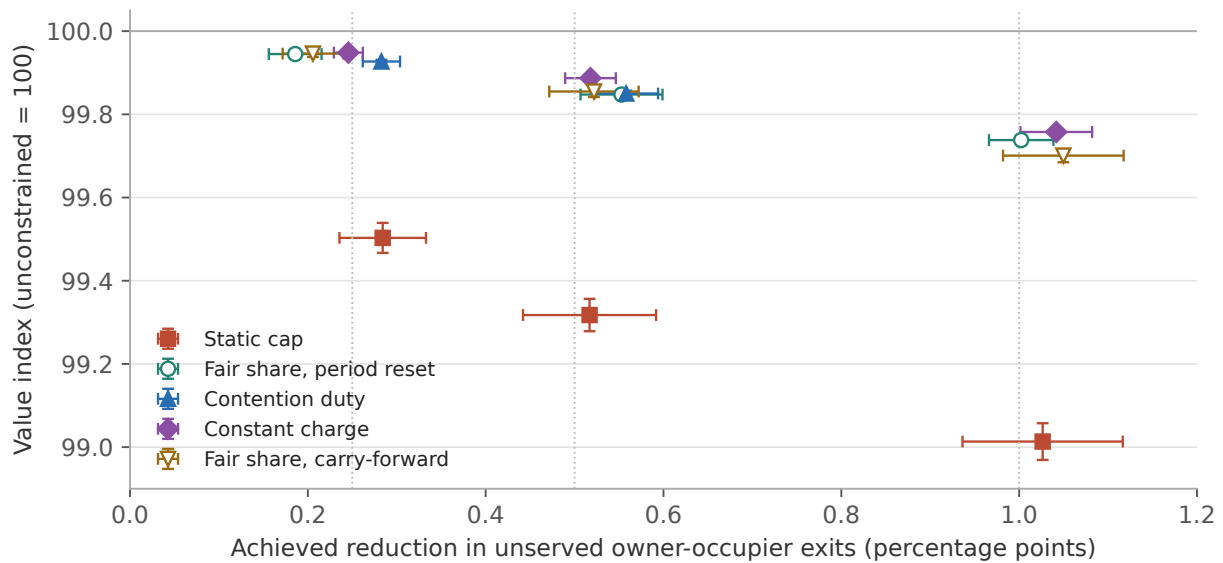
The cap is the expensive instrument at every target, giving up 0.50, 0.68 and 0.99 index points. Per point of access it costs four to eight times as much realised value as the charges, which differ from the fair share by hundredths of an index point. At the two smaller targets the constant charge keeps more value than the duty, by 0.021 (paired SE 0.0017) and 0.037 points (paired SE 0.0044), while achieving 0.04 points less full-run access (paired SE 0.013 and 0.019).

Target (pp)	Instrument	Setting	Full-run gain (pp; SE)	Cohort gain (pp; SE)	Value index (SE)
0.25	Static cap	73	0.28 (0.023)	0.71 (0.057)	99.50 (0.017)
0.25	Fair share, period re-set*	0.1172	0.19 (0.014)	0.25 (0.021)	99.94 (0.004)
0.25	Contention duty	0.3006	0.28 (0.010)	0.64 (0.029)	99.93 (0.002)
0.25	Constant charge	0.0213	0.25 (0.008)	0.34 (0.016)	99.95 (0.002)
0.25	Fair share, carry-forward	0.1113	0.21 (0.016)	0.15 (0.025)	99.95 (0.004)
0.5	Static cap	68	0.52 (0.035)	1.29 (0.087)	99.32 (0.018)
0.5	Fair share, period re-set*	0.0977	0.55 (0.022)	0.74 (0.037)	99.85 (0.004)
0.5	Contention duty	0.5641	0.56 (0.017)	1.24 (0.039)	99.85 (0.003)
0.5	Constant charge	0.0424	0.52 (0.013)	0.75 (0.026)	99.89 (0.003)
0.5	Fair share, carry-forward	0.1036	0.52 (0.024)	0.61 (0.045)	99.85 (0.006)
1	Static cap	60	1.03 (0.042)	2.56 (0.105)	99.01 (0.021)
1	Fair share, period re-set	0.0859	1.00 (0.017)	1.37 (0.029)	99.74 (0.005)
1	Contention duty	—	not reachable (at most 0.88)	—	—
1	Constant charge	0.0802	1.04 (0.019)	1.56 (0.036)	99.76 (0.004)
1	Fair share, carry-forward	0.0946	1.05 (0.032)	1.38 (0.064)	99.70 (0.007)

Table 4. Instruments tuned to common full-run access targets in the main market: settings tuned on four seeds, results on sixteen held-out seeds. Gains are reductions in the full-run and completed-cohort unserved rates. The cap setting is acquisitions per entity, the fair-share settings are institutional period shares, and the charges are rates. An asterisk marks a period-reset fair share whose tuning could not come within the larger of 0.02 points and 10% of the target (Section 5.4).

In the NYC-like market the cap is the costly way to buy access

Settings tuned on 4 seeds to full-run targets of 0.25, 0.5 and 1 point; outcomes on 16 separate seeds



Dotted lines: targets. Bars: approximate 95% Monte Carlo intervals.

Hollow markers: fair-share rules; the period-reset rule misses the two smaller targets.

Figure 2. Settings tuned to common full-run access targets in the main market and evaluated on sixteen separate seeds. Bars are approximate 95% Monte Carlo intervals conditional on the settings. Points are not joined into a frontier.

The full-run starvation rate has a blind spot. Buyers still searching when the run ends count as neither served nor unserved, and instruments differ in how many they leave in that state. Following every owner-occupier who enters in periods 36 to 83 until they buy or leave by period 119 resolves the whole cohort. On the evaluation seeds, 15.1% of that cohort leaves unserved without intervention, against a full-run rate of 10.1%.

At the full-run settings of Table 4, cohort gains range from about three quarters of the full-run gain, for the carry-forward fair share at the smallest target, to two and a half times it, for the cap at every target. The two measures also rank the instruments differently. At the half-point target the cap and the constant charge give the smallest full-run gains, 0.52 points each, while the cap gives the largest cohort gain, 1.29 points, against 0.75 for the constant charge. Matched to the duty on full-run access, the constant charge delivers 0.29 and 0.48 points less cohort access (paired SE 0.028 and 0.039).

Retuning to the cohort measure reverses the comparison between the two charges (Table 5). Tuned on the same four seeds to cohort gains of 0.5 and 1 point and evaluated on sixteen further seeds, the duty keeps 0.018 and 0.027 more index points than the constant charge (paired SE 0.0020 and 0.0052). Their cohort gains are within 0.03 points of each other (paired SE 0.022 and 0.044), while the constant charge delivers 0.13 and 0.23 points more full-run access (paired SE 0.011 and 0.021). Each charge keeps more value on the population it was tuned for, as in the stress test (Section 6.3), and neither result generalises beyond its stated population.

Target (pp)	Instrument	Setting	Cohort gain (pp; SE)	Full-run gain (pp; SE)	Value index (SE)
0.5	Static cap	75	0.44 (0.055)	0.18 (0.022)	99.60 (0.019)
0.5	Fair share, period re-set	0.1016	0.56 (0.035)	0.40 (0.018)	99.89 (0.005)
0.5	Contention duty	0.2542	0.50 (0.020)	0.23 (0.006)	99.94 (0.002)
0.5	Constant charge	0.0297	0.52 (0.029)	0.36 (0.012)	99.92 (0.002)
0.5	Fair share, carry-forward	0.1040	0.51 (0.053)	0.44 (0.034)	99.87 (0.010)
1	Static cap	70	0.88 (0.077)	0.35 (0.032)	99.42 (0.020)
1	Fair share, period re-set	0.0938	0.95 (0.049)	0.67 (0.025)	99.81 (0.006)
1	Contention duty	0.5307	1.07 (0.040)	0.50 (0.017)	99.86 (0.004)
1	Constant charge	0.0561	1.07 (0.050)	0.72 (0.017)	99.84 (0.004)
1	Fair share, carry-forward	0.0974	1.05 (0.060)	0.80 (0.038)	99.76 (0.011)

Table 5. Instruments tuned to common completed-cohort access targets in the main market: settings tuned on four seeds, results on sixteen seeds disjoint from those of Table 4.

5.4 Which quota accounting

With three institutional bids per district-period and about 20 listings, a period share s lets institutions make $[s \cdot L]$ purchases in a period of L listings while an owner-occupier waits. The allowance moves in whole purchases, so the full-run gain jumps between steps: on the tuning seeds, from 0.21 to 0.31 points near the 0.25 target and from 0.40 to 0.55 near the 0.5 target. No period share reaches either target (asterisks in Table 4). Allowances that carry forward, one credit spent per institutional purchase while an owner-occupier waits, accumulate fractional credit and reach all three targets on the tuning seeds, with held-out gains of 0.21, 0.52 and 1.05 points. The gap is a property of the accounting, not of fair-share quotas. Unlimited banking is not proposed as a design: it changes the timing of institutional purchases, and limits on accumulation, expiry and borrowing are untested.

6. Stress test: a market dominated by institutions

This section repeats the comparisons in the market of earlier versions, where institutions win 58.1% of sales at a 12% offer discount. Every instrument has large effects here, and so do the measurement choices.

6.1 The baseline comparison

Regime	OO share	Inst. share	Starvation	Value index	Oracle eff.	Contemp. eff.
Unconstrained	34.0%	58.1%	41.9%	100.00	99.5%	95.4%
Static per-entity cap	75.8%	6.3%	8.7%	88.72	88.3%	78.5%
Work-conserving fair share	54.5%	32.2%	24.7%	98.22	97.7%	95.2%
Contention-triggered duty	41.1%	49.4%	36.6%	100.03	99.5%	98.1%
Combined (share + duty)	54.7%	32.1%	24.6%	98.20	97.7%	96.3%

Table 6. Allocation regimes in the stress test, means of 12 replications. OO and Inst. share are the shares of units bought by owner-occupiers and institutions. The value index is the sum of winners' valuations relative to the unconstrained regime under matched seeds.

The instruments deliver very different amounts of access (Table 6). The duty lowers starvation from 41.9% to 36.6% with a value index of 100.03; its gain of about 0.026 index points has a paired Monte Carlo standard error of 0.0054, so it is precisely estimated and economically negligible. Starvation falls to 24.7% under the

fair share, at a 1.8% loss of value, and to 8.7% under the cap, at an 11.3% loss. Combining the fair share and the duty adds almost nothing, because the fair share lowers the contention signal and with it the duty's rate. The unconstrained market reaches 99.5% of the oracle benchmark, so the room for aggregate efficiency gains is small; the differences that are large are distributional.

Whether a larger access gain is worth its cost depends on the welfare weights. Scoring each allocation with $W(\omega) = \omega V_{OO} + \sqrt{\omega} V_{SI} + V_I$, where V is the realised valuation of each class, the fair share beats the duty from $\omega \approx 1.16$ and the cap becomes the best of the three original instruments from $\omega \approx 1.66$. The cap breaks even against no intervention at $\omega \approx 1.35$. These weights are normative scenarios, not estimates; `replication/welfarist.py` reports the full schedule.

Two constraints from the extended version change the cap's case in particular. When institutional capital is split across legal entities at no cost, the cap's effect decays and disappears at 25 entities; beneficial-owner aggregation would be needed to enforce it. When listings respond to seller proceeds with unit elasticity and the seller bears the whole charge, the cap's value loss rises to 17.3% and its starvation rate to 16.0%, while the duty loses 2.0 index points. The duty stays the least costly in value at every elasticity tested, but it is not costless.

Waiting times carry the same trap. When costless entity splitting erodes a cap, the 99th-percentile wait of owner-occupiers who complete a purchase falls, even as starvation rises: the buyers who would have waited longest leave without buying and drop out of the sample. A waiting-time statistic for completed purchases means little without the share of entrants who never complete.

6.2 Which efficiency benchmark

The benchmark result is larger in the stress test (Figure 1, right panel). Here the contemporaneous benchmark averages about 521,000 per unit without intervention, 562,000 under the cap and 506,000 under the duty, a spread of about 11%. On the fixed benchmark the duty starts adding value at an offer discount of about 11.7% (95% interval [11.5, 11.8]). On the moving benchmark it appears to add value from about 5.4% [5.4, 5.5]. At the tested discounts of 6%, 8% and 10% the two measures disagree on the sign of the policy's effect: the moving benchmark reports gains of 0.22 to 1.95 points where the fixed one reports losses of 0.12 to 0.35. Choosing the moving benchmark more than doubles the range of discounts over which the duty seems to pay for itself.

Institutional share at $w = 12\%$	Interpolated break-even w	95% interval
12.6%	none up to 20%	—
22.4%	15.9%	[15.6, 16.1]
29.9%	14.2%	[13.8, 14.7]
40.2%	12.5%	[12.2, 12.7]
46.8%	12.1%	[11.8, 12.3]
52.8%	11.9%	[11.8, 12.1]
58.1%	11.7%	[11.5, 11.8]
63.5%	11.9%	[11.7, 12.0]
68.3%	11.8%	[11.7, 11.9]
73.7%	12.0%	[11.8, 12.1]
79.9%	12.3%	[12.2, 12.4]

Table 7. Offer discount at which the duty stops losing value on the fixed benchmark, across eleven institutional-capacity settings, with the duty ceiling fixed at 0.22. The first row is the main market's capacity, where the fixed-benchmark gain stays negative across the whole grid up to 20% (Table 3); the other ten rows are the stress test's capacity sweep. Institutional shares are unconstrained shares at a 12% discount. Crossings are linearly interpolated between grid nodes two points apart, with eight replications per node. Intervals are 95% percentile intervals from 4,000 bootstrap resamples of the eight matched seeds (`threshold_intervals.py`); they reflect simulation noise, not error from the linear interpolation or the model.

Across the stress-test settings in Table 7 the break-even offer discount ranges from 11.7% [11.5, 11.8] to 15.9% [15.6, 16.1], and at the main market's participation it is not reached below 20%. Below the stress test's

institutional share, the fewer the institutional purchases, the larger the discount at which the duty starts to pay for itself on the fixed benchmark.

6.3 Which access population

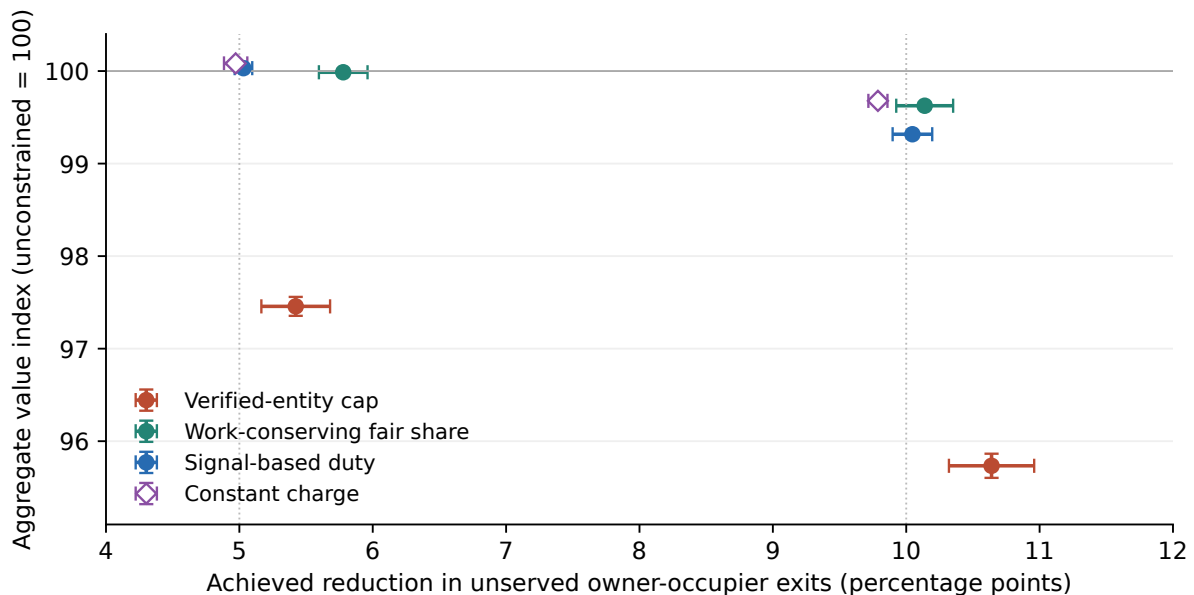
In the stress test the instruments were tuned in the same way to five- and ten-point full-run targets (Table 8, Figure 3). The cap is again the most expensive way to buy access, giving up about 2.5 index points for five points of access and 4.3 for ten. The fair share and the duty cost almost nothing at five points. At ten points the fair share delivers slightly more access than the duty and keeps more value, so the default ordering between them does not survive tuning.

Target (pp)	Instrument	Setting	Access gain (pp; SE)	Value index (SE)	Constant – duty, value (SE)
5	Static cap	336	5.42 (0.121)	97.46 (0.048)	—
5	Fair share	0.5000	5.78 (0.086)	99.99 (0.007)	—
5	Signal duty	0.2115	5.03 (0.031)	100.03 (0.005)	—
5	Constant charge	0.0779	4.97 (0.041)	100.08 (0.005)	+0.051 (0.003)
10	Static cap	273	10.64 (0.150)	95.73 (0.061)	—
10	Fair share	0.4121	10.14 (0.100)	99.63 (0.014)	—
10	Signal duty	0.4657	10.05 (0.070)	99.32 (0.009)	—
10	Constant charge	0.1466	9.79 (0.034)	99.68 (0.008)	+0.361 (0.008)

Table 8. The stress test's instruments at common full-run access targets: settings tuned on four seeds, results on sixteen held-out seeds. The cap setting is acquisitions per entity, the fair share is the institutional period share, and the charges are rates. The last column is the paired value difference between the constant charge and the duty, using the delta method on matched seeds.

Instrument rankings change with the access target

Settings tuned on 4 seeds; outcomes on 16 separate seeds



Dotted lines: tuning targets. Bars: approximate 95% Monte Carlo intervals.

Fixed verified entities and supply. Intervals condition on chosen settings; no policy frontier is estimated.

Figure 3. Stress test: settings tuned to common full-run access targets and evaluated on sixteen separate seeds, including the constant charge. Bars are approximate 95% Monte Carlo intervals conditional on the settings. Points are not joined into a frontier.

The constant charge, which ignores contention, keeps more value than the duty at both targets, by 0.05 and 0.36 index points (Table 8, Figure 3). Its access gains fall slightly short of the duty's, so the comparison is close but not exact.

Regime	Setting	Full-run gain (pp)	Cohort gain (pp; SE)	Cohort unserved (%)
Unconstrained	—	—	—	60.09
Static cap	336	5.42	13.53 (0.279)	46.56
Fair share	0.5000	5.78	7.91 (0.134)	52.18
Signal duty	0.2115	5.03	9.34 (0.072)	50.75
Constant charge	0.0779	4.97	7.15 (0.065)	52.94
Static cap	273	10.64	26.44 (0.318)	33.65
Fair share	0.4121	10.14	13.97 (0.164)	46.12
Signal duty	0.4657	10.05	18.04 (0.140)	42.05
Constant charge	0.1466	9.79	14.03 (0.091)	46.06

Table 9. Stress test: completed-cohort access at the settings of Table 8 and the common-target exercise, without retuning. The cohort averages 4,061 entrants per replication.

The full-run measure understates every instrument’s effect on this cohort, by different amounts (Table 9). The unconstrained row shows the size of the gap: 60.1% of the cohort leaves without buying, against a full-run rate of 41.9%, partly because the full-run denominator includes entrants who are still searching when the run ends. For the cap, a 5.42-point full-run gain is a 13.53-point cohort gain, 2.5 times larger; the fair share’s 5.78 becomes 7.91, 1.4 times larger. The rankings also differ: at the five-point settings the fair share leads on full-run access and the duty leads on cohort access. The two measures describe different populations, and a comparison needs to state which one it targets.

Retuning to the cohort measure changes the settings and the comparison. Tuned on the original four seeds to five- and ten-point cohort gains and evaluated on sixteen new seeds, a five-point cohort gain needs a cap of 385 acquisitions per entity, not 336. At the ten-point cohort target the order of the two charges flips. The duty keeps 0.032 more index points than the constant charge (paired SE 0.0033), with cohort gains within 0.03 points of each other (paired SE 0.044), while the constant charge still delivers 1.6 points more full-run access (paired SE 0.044). Each charge wins on the population it was tuned for. Near five points the value difference, 0.0012 (paired SE 0.0017), is not resolved. Neither result generalises beyond its stated population.

6.4 Quota accounting and robustness

The quota limit of Section 5.4 also appears in a third configuration, with a single institutional entity placing one bid per district-period and an unconstrained institutional share of 4.47%. Any positive period share then admits the single bid, and a zero share excludes it whenever an owner-occupier is waiting, so the quota can only overshoot a 0.5- or 1-point target or miss it entirely. Carry-forward credits reach gains of 0.48 and 0.96 points at value indices of 99.737 and 99.477. In the stress test itself, carry-forward shares of 0.5103 and 0.4351 reach gains of 5.22 and 10.18 points at value indices of 99.997 and 99.600 (`carry_share_results.json`).

The extended version reports the remaining sensitivity exercises; the main points are these. Letting 10% of accepted financed sales fail leaves the value indices and ranking similar and the fixed-benchmark break-even almost unchanged (11.8% against 11.7%), but it raises failed transactions under the duty by 21.6% and more than doubles them under the cap. Reduced search, with inspection probabilities of 0.5 and 0.25, preserves the ranking. Sweeping institutional buyers’ mean valuation from 340,000 to 410,000 moves the duty’s index by at most 0.06 points and raises the cap’s cost from 8.6% to 16.0%. The baseline institutional share of 58% is far above observed markets; with the share lowered to 22.5% the duty and fair share barely act, reducing starvation by 0.7 and 0.1 points, while the cap still costs the most value at every share tested. The rate sweep finds the highest mean value index at $\tau_{\max} = 0.11$ (100.12), not at the baseline 0.22 (100.02); at an offer discount of 5.6%, the ceilings 0.11, 0.22 and 0.44 all lose value.

7. Sensitivity to the New York City proxies

Table 10 varies the proxy inputs, holding instrument settings at their defaults. The constant charge equals the duty ceiling times its mean signal across district-periods. This matches an unweighted scheduled haircut, without equating charges on completed purchases or access. Sections 5.3 and 6.3 give the comparisons at common access targets.

Raising the offer discount to the upper endpoint of 5.3% or to the headline gap’s 8.8% changes little: full-run starvation without intervention rises to 10.6% and 11.1%, and the duty still lowers it by about 0.2 points. Counting all company buyers as institutions raises their simulated share to 25.0% and full-run starvation to 18.8%; every instrument then acts more strongly, and the cap lowers starvation by 11.29 points at a 5.16% loss. In the upper-gap and headline-gap scenarios, the contemporaneous benchmark gives duty gains of 0.11 and 0.19 points, while the fixed oracle gives losses of 0.04 and 0.03 points.

Scenario	Discount / model share (%)	Unserved: full / cohort (%)	Cap	Fair share	Duty	Constant
Company cash	1.5 / 11.0	9.8 / 14.9	−1.77 / −1.36	0.00 / 0.00	−0.20 / −0.05	−0.24 / −0.05
Upper gap	5.3 / 11.5	10.6 / 16.0	−2.15 / −1.47	0.00 / 0.00	−0.21 / −0.04	−0.26 / −0.04
All companies	1.5 / 25.0	18.8 / 27.2	−11.29 / −5.16	−0.46 / −0.06	−1.09 / −0.18	−1.21 / −0.17
Headline gap	8.8 / 12.1	11.1 / 16.8	−2.53 / −1.56	0.00 / 0.00	−0.20 / −0.03	−0.25 / −0.03
Stress test	12.0 / 58.1	41.9 / 59.8	−33.16 / −11.28	−17.14 / −1.78	−5.24 / +0.03	−4.85 / +0.10

Table 10. NYC proxy scenarios and the stress test, twelve matched seeds. Each instrument cell is change in full-run starvation (percentage points) / realised value (percent), relative to its scenario’s unconstrained market; negative starvation means improved access. The unserved columns give unconstrained full-run and completed-cohort rates. The cohort enters in periods 36 through 83 and is fully resolved by period 119. Instruments retain the baseline settings; constant charges match the duty’s mean scheduled haircut. Shares are simulated purchase shares, not exact empirical targets. `nyc_anchored.json` supplies full-run results; `nyc_cohort_check.json` supplies cohort results, seed-level runs and paired Monte Carlo standard errors.

Starvation is a model quantity, and the company shares concern a selected set of linked transactions. The two headline inputs cannot establish New York City’s policy effects.

New York also charges purchase mortgages. Applying headline borrower-side rates gives averages of 1.41% of price for houses and 1.28% for condominiums before credits, exemptions and consolidation adjustments (Loschi 2026c). The model does not compare replacing that charge with taxing cash purchases.

8. What the model leaves out

Bidding is not strategic. Valuations are private, fixed and bid truthfully, so the simulation tests what the instruments do to allocations, not how buyers would respond to them. In the stress test, a limited exercise lets institutions reallocate a fixed acquisition budget across districts in proportion to $(1 - \tau_{\max} s_d)^\lambda$. For λ from zero to eight, the duty’s value index ranges from 99.60 to 100.22 and its starvation rate from 36.5% to 36.7%, so even the sign of its small aggregate effect is not stable once buyers move. Bid shading, entry and manipulation of the contention signal are not modelled. This is the largest gap.

Outside the two proxy inputs, the market’s parameters are illustrative. In the stress test, patience, the number of districts, valuation dispersion, the contention window and search friction leave the full ordering unchanged; arrival and listing rates do not. With owner-occupier arrivals doubled to 34 per district-period, which lowers the unconstrained institutional share to 38%, the fair share scores 100.20 and overtakes the duty at 100.09, while the duty still lowers starvation more (4.0 against 2.9 points). The cap is the most costly instrument at every point tested in either market, at 97.3 in that case. Its place relative to the contention-linked instruments holds across the reported exercises; the ordering within that pair does not. These sensitivity exercises have not been repeated in the main market.

The value index is not welfare. It counts realised willingness to pay and leaves out renters, the use of any revenue, administrative costs and differences in the marginal utility of income. The supply exercise lets listings respond to seller proceeds; it has no construction sector, rents or entry. The horizon is finite, so the full-run starvation rate is not a lifetime purchase probability, and entity splitting is costless, so the cap’s decay is an upper bound on avoidance without enforcement.

9. Implementation

Each instrument needs information the model takes for granted.

The duty needs an exit registry. Its signal counts owner-occupiers who leave without buying. Recording that requires a register that observes entry into the market and distinguishes purchase, withdrawal and a move to another market; routine transaction records do not show it. The signal can be pushed in either direction. Institutions could depress it by withdrawing bids during the measurement window. Anyone who gains from a higher duty could inflate it with entries or withdrawals staged to register as unserved exits. A workable version would count only verified participation, such as a mortgage pre-approval or a deposit, keep the records with a party that has no stake in the rate, and limit how far the rate can move in one period. Björkegren, Blumenstock and Knight (2026) show how to design decision rules that stay accurate when the people they score change their behaviour to move the score; a contention signal faces the same problem. The constant charge needs none of this, which is the practical side of Section 5.3: the duty's value advantage is small, depends on the access population, and has to be weighed against the cost of measuring contention.

The cap needs ownership aggregation. A per-entity limit binds a legal person, and in the stress test a single owner who can form entities at no cost defeats it with 25 of them. An enforceable cap would need a definition of common control, disclosure of beneficial owners, verification and sanctions. Federal beneficial-ownership reporting has exempted domestic companies since 2025, so a state rule would have to supply that information itself. A charge on each purchase avoids the gain from multiplying entities, although it still needs an enforceable definition of who pays.

The fair share's release rarely binds. The quota is work-conserving: it lifts when no owner-occupier is waiting. In the stress test someone is always waiting, so the simulation tests the quota and not its release; in the main market the default quota does not bind at all. Its behaviour in slack markets is untested.

Classification decides who is reached. Payment method, institutional status and owner-occupation are different categories. Some owner-occupiers pay cash and some investors borrow. The model charges institutions. The 2026 New York proposal would have charged cash purchases of \$1 million or more, and a payment-method test can be avoided with a short-term loan repaid after closing. A charge on all non-owner-occupier purchases would also reach small investors and would need its own simulation.

Coverage and geography matter. A tax reaches only what its statute defines. New York's recording tax applies to mortgages on real property, so co-operative share loans fall outside it, while the mansion tax expressly includes co-operative units. Buyers can also move purchases to places the policy does not cover; the district exercise in Section 8 captures only a limited version of that response.

10. Conclusion

In this model, instrument rankings depend on the efficiency benchmark and the access population, both in a market set to New York City proxies and in a stress test dominated by institutions. A benchmark computed over waiting buyers moves with the policy: in the main market it shows the contention duty paying for itself at discounts above about 2.2%, where a fixed benchmark shows a loss up to 20%. Following entrants to purchase or exit changes measured access gains and can reverse their ordering. The constant charge and the duty swap places when retuned to a different access population. Carrying quota allowances forward resolves a limit created by resetting them each period.

At the company-cash proxy values, the duty lowers full-run unserved exits by 0.2 points, while the cap lowers them by 1.8 points at a 1.4% loss in realised value. Removing the cash discount there lowers unserved exits only from 9.8% to 9.5%, so at those inputs certainty explains little of the access shortfall. In the stress test the instruments' effects are larger, and buyer-class weights above about 1.66 favour the cap among the three original instruments; entity splitting and elastic supply weaken it.

The New York City inputs leave the model's other market conditions illustrative. Transaction composition affects the observed price gap, and company status does not identify certainty or beneficial ownership. Policy evaluation needs a fixed efficiency benchmark and an explicit access population. Waiting-time reports should include exits without purchase, and quota comparisons should specify how allowances accumulate.

Declarations

Declaration of generative AI use. During preparation the author used Anthropic’s Claude for drafting, coding and literature assistance, including the preparation of this shortened version, and OpenAI’s Codex for review, reproduction checks and figure revision. For version 6.1, Claude ran the New York City sensitivity exercise, and Codex checked it and ran the completed-cohort diagnostic. Computational reviews reported with earlier versions were performed by separate AI agents, not by human referees. For version 6.2, Claude updated the citations, added the related literature and the statement on the oracle bound, and edited wording; no result changed. For version 6.3, Claude renamed the holding cost in Section 4.1, added the remark on bid shading after Proposition 1, moved the Section 5.3 tuning results into Table 5 and edited that section’s wording; no result changed. For version 6.4, Claude made the market set to New York City proxies the main case, ran the measurement exercises there (`realistic_market.py`, 1,224 runs) under a plan recorded before the run, built Tables 2 to 5 and Figures 1 and 2 from its outputs, moved the earlier calibration to a stress-test section whose results are unchanged, renamed the holding cost k to avoid a clash with the charge rate h , and added references on housing search, transfer taxes, waiting lists, agent-based models, asymmetric auctions and manipulation-robust rules, each checked against its journal page. The author is responsible for the content.

Competing interests. The author founded HolaCasaClara, a residential real estate service in Latin America with no operations or interests in the jurisdictions discussed here. No party funded, reviewed or approved the manuscript.

Funding. None.

Data availability. The package retains the extended version’s model and results (Loschi 2026b) and adds the scripts and outputs for the main market (Section 5) and the proxy scenarios (Section 7). The version family is <https://doi.org/10.5281/zenodo.22164073>. New York City inputs come from the companion study (Loschi 2026a) and linked dataset (Loschi 2026d). Their source hashes and the share calculation accompany this version.

Prior dissemination. Earlier versions are posted on Zenodo under the concept DOI above and on SSRN.

References

- Arora, R., Dekel, O., & Tewari, A. (2012). Online bandit learning against an adaptive adversary: From regret to policy regret. *Proceedings of the 29th International Conference on Machine Learning*.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3), 244–263.
- Axtell, R. L., & Farmer, J. D. (2025). Agent-based modeling in economics and finance: Past, present, and future. *Journal of Economic Literature*, 63(1), 197–287.
- Bertsimas, D., Farias, V. F., & Trichakis, N. (2011). The price of fairness. *Operations Research*, 59(1), 17–31.
- Bertsimas, D., Farias, V. F., & Trichakis, N. (2012). On the efficiency–fairness trade-off. *Management Science*, 58(12), 2234–2250.
- Björkegren, D., Blumenstock, J. E., & Knight, S. (2026). Manipulation-robust prediction. *American Economic Review*, 116(9), 3263–3293.
- Bloomberg News (2026). NY plans new tax on NYC homes over \$1 million bought with cash (14 May); NYC cash home purchases tax is likely to be dropped from New York state budget (21 May); New York lawmakers approve second-home tax in state budget (27 May). Bloomberg.
- Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, 130(4), 1547–1621.
- Che, Y.-K., & Gale, I. (1998). Standard auctions with financially constrained bidders. *Review of Economic Studies*, 65(1), 1–21.
- Geanakoplos, J., Axtell, R., Farmer, J. D., Howitt, P., Conlee, B., Goldstein, J., Hendrey, M., Palmer, N. M., & Yang, C.-Y. (2012). Getting at systemic risk via an agent-based model of the housing market. *American Economic Review*, 102(3), 53–58.
- Hafalir, I., & Krishna, V. (2008). Asymmetric auctions with resale. *American Economic Review*, 98(1), 87–112.

- Han, L., & Hong, S.-H. (2024). Cash is king? Understanding financing risk in housing markets. *Review of Finance*, 28(6), 2083–2118. <https://doi.org/10.1093/rof/rfae025>
- Kopczuk, W., & Munroe, D. (2015). Mansion tax: The effect of transfer taxes on the residential real estate market. *American Economic Journal: Economic Policy*, 7(2), 214–257.
- Leshno, J. D. (2022). Dynamic matching in overloaded waiting lists. *American Economic Review*, 112(12), 3876–3910.
- Lipsey, R. G., & Lancaster, K. (1956). The general theory of second best. *Review of Economic Studies*, 24(1), 11–32.
- Loschi, P. (2026a). *Who Gets the House? Related-Party Transfers and the Cash–Mortgage Price Gap in New York City* (Version 3.7) [working paper and replication package]. Zenodo. <https://doi.org/10.5281/zenodo.22421850>
- Loschi, P. (2026b). *The Execution-Certainty Wedge: Allocation, Distribution, and the Design of Housing-Market Instruments* (Version 5.1, extended version) [working paper and replication package]. Zenodo. <https://doi.org/10.5281/zenodo.22977667>
- Loschi, P. (2026c). *The Shape of the Tax: Round Prices and Bunching at New York City’s 2019 Transfer-Tax Notches* (Version 3.0) [working paper]. Zenodo. <https://doi.org/10.5281/zenodo.22925301>
- Loschi, P. (2026d). *NYC residential sales linked to ACRIS deeds and mortgages, 2016–2025* (Version 1.0) [dataset]. Zenodo. <https://doi.org/10.5281/zenodo.22977532>
- Loschi, P. (2026e). *Auditing policy dependent benchmarks in allocation simulations* (Version 1.0) [working paper and code]. Zenodo. <https://doi.org/10.5281/zenodo.23257315>
- Maskin, E., & Riley, J. (2000). Asymmetric auctions. *Review of Economic Studies*, 67(3), 413–438.
- Perdomo, J. C., Zrnic, T., Mendler-Dünner, C., & Hardt, M. (2020). Performative prediction. *Proceedings of the 37th International Conference on Machine Learning*, PMLR 119, 7599–7609.
- Piazzesi, M., Schneider, M., & Stroebel, J. (2020). Segmented housing search. *American Economic Review*, 110(3), 720–759.
- Reher, M., & Valkanov, R. (2024). The mortgage-cash premium puzzle. *The Journal of Finance*, 79(5), 3149–3201.
- Tesfatsion, L. (2006). Agent-based computational economics: A constructive approach to economic theory. In L. Tesfatsion & K. L. Judd (Eds.), *Handbook of Computational Economics*, Vol. 2, 831–880. Elsevier.
- Thakral, N. (2019). Matching with stochastic arrival. *AEA Papers and Proceedings*, 109, 209–212.
- Waldinger, D. (2021). Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation. *American Economic Review*, 111(8), 2660–2696.
- Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies*, 41(4), 477–491.
- Weitzman, M. L. (2018). *Prices or Quantities Dominate Banking and Borrowing*. NBER Working Paper 24218.

Appendix A. Reproducibility

The package retains the extended version’s simulation code and adds the main-market and proxy scripts. `realistic_market.py` runs the main-market benchmark sweep, the common full-run and cohort targets and the carry-forward quota (Tables 3 to 5 and the first row of Table 7) under the plan in evidence/`realistic_market_plan.json`; `realistic_tables.py` builds those tables, `build_figures_v64.py` draws Figures 1 and 2, and `check_realistic_text.py` checks every main-market number in the text against the outputs. The main-market exercises use the same tuning and evaluation seeds as the stress test, and the benchmark sweep uses the first eight baseline seeds. `clean_reproduce.py` regenerates the baseline, the offer-discount and rate sweeps, the supply and entity experiments, the capacity and valuation sweeps, the break-even schedule of Table 7 and the class-weight schedule. The main scripts are `replication/sim_v6.py` (baseline and failed-transaction variants), `replication/sim_v6_supply.py` (listing response), `replication/sim_v6_search_fixed.py` (reduced search), `replication/displacement.py` (Section 4.2), `replication/welfarist.py` (Section 6.1),

`replication/breakeven_schedule.py` (Table 7), `equal_access.py` (common targets), `constant_charge.py` (Table 8), `check_completed_cohorts.py` (Table 9), `cohort_targets.py` (cohort retuning) and `carry_share.py` (Section 6.4). `threshold_intervals.py` reruns seed by seed the grid nodes around each break-even crossing, checks that their means reproduce the published schedule exactly, and bootstraps the matched seeds for the intervals in Section 6.2 and Table 7.

The baseline uses seed 20260824 and twelve replications, advancing the seed by 100 between runs. The common-target exercises tune on seeds $22260824 + 100i$ for $i = 0, \dots, 3$ and evaluate on $23260824 + 100i$ for $i = 0, \dots, 15$; the cohort retuning is evaluated on sixteen further seeds disjoint from both. Within a seed every policy sees identical arrivals, valuations and listings, and the oracle denominators match; the reduced-search runs draw search from a separate generator so that this also holds there.

On 21 September 2026 the package was extracted into a fresh directory with reads from the original workspace blocked, and every result was regenerated, including the 1,020 runs behind the constant-charge, cohort and carry-forward experiments. All matched the shipped outputs to floating-point tolerance. The New York City estimates in Table 1 come from the companion study’s version 3.4 (Loschi 2026a) and are unchanged in its version 3.7. For Tables 2 and 10 run `replication/nyc_anchored.py`, then `replication/nyc_cohort_check.py`. The original model is unchanged; the cohort diagnostic reproduces every original summary exactly. `replication/verify_nyc_inputs.py` recomputes the share proxies from the separately deposited dataset; instructions and input hashes are in the README and `nyc_inputs.json`.