SB | STATISTICS
THE SCIENCE OF HOUSING

SB–2026–01/I

Copying Transaction Prices: How Dutch Housing Appraisals Track Idiosyncratic Variation

Ramon F. A. de Punder
Mathijs R. G. Dijkstra

SB Statistics is the research unit of SlimBieden, a Dutch company that provides prospective home buyers with advisory and informative reports on bidding strategies in the housing market. SB Statistics conducts academic research on the residential property market.

Contact: research@slimbieden.nl
More information can be found at https://www.sb-statistics.com

SlimBieden
Toepad 95
3063 NJ Rotterdam
The Netherlands
Tel.: +31(0)10 307 2130

Copying Transaction Prices: How Dutch Housing Appraisals Track Idiosyncratic Variation*

Ramon F. A. de Punder†
University of Amsterdam, Tinbergen Institute and SB Statistics

Mathijs R. G. Dijkstra
University of Amsterdam, Tinbergen Institute and SB Statistics

November 4, 2026

Abstract

An appraisal is intended to provide an independent estimate of a dwelling's market value. We find instead that appraisals track the realized transaction price, including its idiosyncratic component, the part specific to this buyer and this sale, which carries no information about the expected price if the dwelling were resold almost immediately. This follows from the incentives in mortgage lending: all parties to the transaction benefit when the appraisal allows the purchase to proceed, so that a valuation formed with reference to the agreed price tends towards that price. In Dutch residential transactions, nearly all appraisals are multiples of €5,000, and up to rounding to that amount, 41.5% of appraisals equal the transaction price. A Monte Carlo test that accounts for the uncertainty in each dwelling's price shows decisively that appraisals lie far too close to realized prices to have been formed independently of them. Regressions show that appraisals pass on about four fifths of the deviation of the price from the mean of its estimated distribution, and we show that this coupling raises the collateral risk for both borrower and lender. These findings call into question physical appraisals as an independent safeguard, and point to statistical valuation models, formed independently of the transaction price, as a natural alternative.

Keywords: Idiosyncratic copying, Collateral risk, Predictive distributions

*This research was conducted in collaboration with Viisi. We are very grateful to Viisi for making the appraisal reports available and for their assistance in constructing and interpreting the dataset. In particular, we thank Hergen Dutrieux, Frank Tukker, and Ricardo den Brinker for their extensive support and contributions throughout the research project. We would additionally like to thank Marc Brouwers, Tim Bruinsma, Koen Cremers, Cees Diks, Floris Holstege, Steven Hommes, Stan Koobs, Gert-Jan Kristen, Sam van Meer, Bram Meilof, Rico Nederveen and Terri van der Zwan for sharing their valuable feedback and insights.

†Corresponding author. Email: `r.f.a.depunder@uva.nl`.

1 Introduction

Mortgage lending creates an incentive that runs against the independence an appraisal is meant to provide. The borrower benefits directly from a higher appraisal, since it can support a larger loan, while the seller, the real estate agent, and other parties to the transaction likewise benefit when the valuation allows the purchase to proceed. The bank itself faces a related tension: it has a commercial interest in originating the mortgage, while its protection against losses upon default depends on an independent assessment of the value of its collateral. This tension may be particularly easy to overlook during prolonged periods of rising house prices, when collateral risk appears limited. What the bank needs from the appraisal is an estimate of the price the dwelling is expected to realize if the borrower defaults and the dwelling is resold to cover the outstanding debt. This valuation should be insulated from transaction-level incentives and formed independently of the price agreed between buyer and seller.

How that value is established has changed repeatedly, in response to findings that earlier methods were unreliable; since 24 July 2026, Dutch lenders may no longer use the WOZ value, the municipal assessment of a dwelling's value, for a new mortgage at all (Ministerie van Financiën 2026). The reliability of the physical appraisal has also been a point of concern. In 2019, De Nederlandsche Bank examined 200,000 appraisals from 2012–2017 and found that appraisals matched the transaction price in a third of the cases and exceeded it in nearly 60%, with an average overvaluation of 2.9% of the sale price (De Nederlandsche Bank 2019; Minister van Binnenlandse Zaken en Koninkrijksrelaties 2019). The finding prompted parliamentary questions on whether subsequent safeguards had worked (Vaste commissie voor Binnenlandse Zaken en Koninkrijksrelaties 2022), and a partial answer arrived only in 2025, when the ministry reported figures compiled by NHG at its request: appraisals exceeded the transaction price in 46% of cases between 2021 and

mid-2025, matched it exactly in 19%, and overvalued the property by €2,566 on average (Minister voor Volkshuisvesting en Ruimtelijke Ordening 2025). The ministry welcomed this improvement but still considered the remaining pattern concerning. A 2026 study of 1.1 million appraisals commissioned by the Dutch appraisers' register described appraisals for 2019–2025 as balanced around the transaction price, with mean and median deviations near zero, and presented this as evidence on appraiser independence (Nederlands Register Vastgoed Taxateurs 2026). Yet in the same data, 21.6% of appraisals match the transaction price exactly and about 66% lie within 2% of it. Similar patterns have been documented in the United States, where about 30% of purchase appraisals equal the contract price and fewer than 10% fall below it, and where loans with appraisals equal to the price default more often (Calem et al. 2021); the paper explains the pattern by modelling the appraiser's incentives.

This paper adds to this literature in three ways. First, we provide formal inference for the dependence between appraisals and transaction prices: using dwelling-specific predictive distributions, we determine how closely appraisals and transaction prices should agree under independence and test whether the observed agreement is compatible with it, and we quantify the dependence by regression. Second, we show that this dependence induces collateral risk for both the lender and the borrower. Third, the predictive distributions provide an alternative valuation that is independent of the agreed price by construction.

Example 1. *A house is listed at €599,000, and a statistical model puts the expected transaction price at approximately €663,000. Laila is aware of this estimate but deliberately bids well above it to increase her chance of winning, offering €721,000, which lies at the 90% quantile of the predictive distribution. Her bid is accepted, and to fund the purchase she needs a mortgage. The size of that mortgage determines how much of her own savings she has to put in. What the bank is willing to lend rests on an appraisal of the property rather than on the price she agreed. A valuation at or above €721,000 lets the sale complete;*

one below it can force Laila to finance the shortfall herself. She therefore prefers a high figure, as does everyone whose part in the transaction depends on it completing. When the appraisal comes back it is reported at €725,000.

The appraised value in Example 1 lies very close to the realized transaction price. One might read this as accuracy, but it is not. Laila's bid is well above what the dwelling is expected to sell for, reflecting a willingness to pay that is idiosyncratic to this transaction. Such a deviation could also reflect model error rather than Laila's willingness to pay, but across our sample the predictive distributions show no evidence of miscalibration (Appendix A.4). Since nothing suggests that a future buyer would pay the same premium, this idiosyncratic deviation carries no information about the price upon immediate resale beyond what the dwelling characteristics already contain. This resale price is, however, exactly what determines the risk to which both parties to the mortgage are exposed. We conjecture that the transaction-level incentives described above nonetheless induce appraisers to couple their estimate of the market value to a single realization of the price, thereby importing its idiosyncratic deviation into the valuation.

To make this precise, let P_x denote the conditional distribution of the market value Y given $X = x$, where x is a realization of the dwelling characteristics X, such as living area, location, and time on market. Let $\mu_Y = \mathbb{E}[Y]$ and σ_Y^2 denote its mean and variance, where throughout $\mathbb{E}[\,\cdot\,]$ is taken under the true laws of the random variables it operates on, and a subscript marks expectations under a modeled distribution. The observed transaction price y is a realization from Y_{old}, whereas the lender is exposed to Y_{new}, the price realized upon an immediate resale, where we take Y_{old} and Y_{new} to be independent draws from P_x. The appraisal A is intended to estimate where this resale will land, so that Y and A are genuinely different random variables: Y is a draw from P_x, whereas A estimates a feature of P_x. The mean μ_Y is the natural choice of feature, since it weighs each possible outcome by its probability and minimizes the expected squared error $\mathbb{E}[(a - Y_{\text{new}})^2]$ over

all appraisal values a. Other loss functions induce other, possibly more robust, location parameters (Gneiting 2011); in each case, however, the target remains a functional of the conditional distribution rather than a particular draw from it.

Since P_x is unknown, μ_Y is an oracle appraisal value. To obtain a feasible counterpart, we specify a model F_x for P_x, with mean $m_Y := \mathbb{E}_{F_x}[Y]$, and estimate it from a training sample of n_0 other transactions, which yields the predictive distribution $\widehat{F}_x$ with mean $\widehat{m}_Y := \mathbb{E}_{\widehat{F}_x}[Y]$. We take $a^* := \widehat{m}_Y$ as the natural feasible appraisal value and assess its calibration in Appendix A.4.

Copying the transaction price, i.e., $A = Y_{\text{old}}$, yields an unbiased estimate of μ_Y, but a poor one. The realized price deviates from μ_Y by an amount that is partly idiosyncratic to the transaction, and copying imports this idiosyncratic part into the appraisal as noise, which lowers its signal-to-noise ratio. Since the estimator is unbiased, its mean squared error equals its variance, which is the full dispersion σ_Y^2 of the market. A statistical model instead pools information across comparable dwellings, so that the idiosyncratic deviations of individual transactions average out. Provided the model is correctly specified in its mean, its mean squared error therefore shrinks as n_0 grows, whereas that of the copied price remains σ_Y^2 however much data is available, which is unnecessarily large given the feasible alternatives.

Whether appraisals in practice import the idiosyncratic deviation of the transaction price, as conjectured above, is an empirical question. An appraisal that serves as an independent check on the agreed price should not draw on that price, so that, once the information contained in the dwelling characteristics X has been accounted for, it should contain no dependence on the realized transaction price. This motivates testing whether the appraisal value A is independent of the transaction price Y_{old}, conditionally on X; see Hypothesis 1. Rejection of this hypothesis provides evidence that the realized transaction price contains information about the appraisal beyond that already contained in X,

consistent with the appraiser following the observed transaction price.

We test this hypothesis on 694 Dutch residential transactions from December 2025 to August 2026. For each dwelling, we observe the appraisal a, the transaction price y, and the characteristics x, and we construct the predictive distribution $\widehat{F}_x$. Descriptive patterns in the data are suggestive of such coupling. In 21.3% of the reports, the appraisal reproduces the transaction price exactly. This understates the agreement, because appraisals and prices are reported in different monetary denominations, that is, in different increments: 95.1% of appraisals are multiples of €5,000, against 47.7% of transaction prices, so that an appraiser who copies the price may report it rounded. Rounding the difference $a - y$ to the denomination of €5,000 in which appraisers report raises the agreement to 41.5%. More broadly, 60.8% of appraisals lie within €5,000 of the transaction price. When the appraisal does depart from the transaction price, the departure is moreover not symmetric. Among appraisals that depart by a multiple of €5,000, the step natural to the appraiser, 69.0% lie above the transaction price (120 of 174). How much agreement these shares represent depends on how uncertain the market value of each dwelling is; our test provides the reference for how closely independently formed appraisals and prices would agree.

We construct this reference with a Monte Carlo test in the sense of Barnard (1963) and Hope (1968). Holding every appraisal at its observed value, we redraw each transaction price from its dwelling-specific predictive distribution $\widehat{F}_x$. The resulting appraisal–price pairs are independent by construction, and repeating this exercise generates a reference distribution for how closely appraisals and prices agree when the appraisal draws no information from the realized price. Comparing the observed agreement with this reference distribution yields a test of Hypothesis 1. Since the reference rests on $\widehat{F}_x$, its validity depends on how well the predictive distributions approximate the true conditional distribution of the price, which we assess in Appendix A.4.

Rejection establishes that appraisals follow transaction prices, but not how strongly. To

quantify the strength of this coupling, we regress the deviation of the appraisal from the predictive mean on the corresponding deviation of the transaction price. Since appraisals and prices are reported in different denominations, we also repeat this regression with both amounts rounded to multiples of a common step κ. To assess whether the coupling persists among appraisals that are not exact or near copies, we further exclude the appraisals that equal the price up to rounding, $|a - y| \leq \kappa/2$. A copula representation, which separates the dependence between appraisal and price from their marginal distributions, places these slopes on the same footing as the benchmark cases of independence and exact copying. Furthermore, as a secondary result, we also test whether appraisals are on average centered at the expected market value.

Hypothesis 1 is rejected. Appraisals track transaction prices far more closely than independence permits. The regressions quantify this coupling: appraisals pass on about four fifths of the deviation of the transaction price from its expected value. The appraisal reported to the bank is hence not an independent estimate of the market value of a dwelling, so that the loan-to-value ratio rests on a figure less independent than its role assumes.

The coupling has important implications for the collateral risk of both the lender and the borrower, even though a perfectly copied transaction price is unbiased for μ_Y. The reason is an asymmetry in the consequences of a default: a shortfall upon resale leaves a residual debt, whereas a surplus accrues to the borrower and does not offset shortfalls elsewhere. The dispersion that a copied price inherits from the market therefore raises the expected residual debt (Section 6).

The predictive model used to construct our benchmark offers an alternative basis for valuing the collateral. Its mean $\widehat{m}_Y$ is formed without reference to the agreed price, and the calibration test in Appendix A.4 gives no evidence that it is displaced from the expected market value. Statistical valuation models therefore provide a natural independent basis for the bank's loan-to-value ratio, and their full predictive distributions additionally describe

the downside risk of the collateral that a single appraisal value discards (Section 2).

The rest of the paper is organized as follows. Section 2 sets out the two distributions and what independence requires of them, Section 3 develops the Monte Carlo test and the regressions, Section 4 describes the sample, Section 5 reports the results and Section 6 discusses what these results imply for collateral risk.

2 A model for appraisal value and transaction value

In this section we make explicit the probabilistic structure of the transaction price Y_i and the appraisal value A_i, together with their conditioning variables X_i, where $i = 1, \ldots, n$ indexes the dwellings in our sample. All random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We take the observed dwelling characteristics x_i as given throughout. Conditional on $X_i = x_i$, we assume that the transaction prices Y_i are independent across dwellings, and likewise for the appraisal values A_i. We require that the true conditional mean $\mu_i \equiv \mu_{Y_i}$ and model-implied conditional mean $m_i \equiv m_{Y_i}$ exist and are finite for every dwelling i. The same regularity is imposed on the second conditional moments, and we write $\sigma_i^2 \coloneqq \mathbb{V}(Y_i \mid X_i = x_i)$ and $\varsigma_i^2 \coloneqq \mathbb{V}(A_i \mid X_i = x_i)$.

The dwelling-specific conditional distributions of the transaction price are denoted by $P_i \equiv P_{x_i}$ and $F_i \equiv F_{x_i}$ for the true and model-implied distributions, respectively. Thus, $\mu_i = \mathbb{E}[Y_i]$ and $m_i = \mathbb{E}_{F_i}[Y_i]$. In practice, F_i is estimated from a training sample of size n_0, and we write $\widehat{F}_i$ and $\widehat{m}_i \coloneqq \mathbb{E}_{\widehat{F}_i}[Y_i]$ for the estimated counterparts. We define the true error $\varepsilon_i \coloneqq Y_i - \mu_i$, which has conditional mean zero and variance σ_i^2 by construction, and the model-implied residual $e_i \coloneqq Y_i - m_i$.

Remark 1. *In reality, the conditional distributions of the transaction price and the appraisal, as well as the model-implied distribution, condition on different information sets. Buyers and appraisers both observe characteristics that are not contained in x_i, such as*

the state of maintenance or the quality of the interior. Such shared information induces dependence between A_i and Y_i conditional on $X_i = x_i$, even when appraisals are formed independently of the transaction price, in the sense that A_i and Y_i are independent conditional on all information shared by buyers and appraisers. We abstract from this difference and assume that all three distributions condition on the same characteristics x_i, so that any dependence between A_i and Y_i conditional on $X_i = x_i$ is attributable to the appraisal process.

The observed appraisal value is a realization of the random variable A_i, which, conditional on the same characteristics $X_i = x_i$ as Y_i, has true distribution $Q_i \equiv Q_{x_i}$ and model distribution $G_i \equiv G_{x_i}$ with mean $m_{A_i} \coloneqq \mathbb{E}_{G_i}[A_i]$. Both variables are conditioned on the same x_i because P_i arises from the bids of buyers, and the appraiser's task is to estimate the mean price these bids would generate; the relevant characteristics are thus those on which buyers base their bids.

Remark 2. *The appraisal target μ_i is a scalar, or, represented as a random variable, a degenerate random variable with distribution δ_{μ_i}. The transaction price Y_i differs from this target by the deviation ε_i, which is idiosyncratic to the transaction given x_i. The appraisal A_i, by contrast, is a single-observation estimator of the target and need not be degenerate; under an unbiased appraisal process, its distribution Q_i is centered at μ_i, and an appraisal that tracks the realized price imports the idiosyncratic deviation ε_i into its error ξ_i.*

The properties of the appraisal process that we study can now be expressed in terms of the appraisal error $\xi_i \coloneqq A_i - \mu_i$ and the transaction-price error ε_i. The first property, independence from the realized transaction price, requires $\xi_i \perp \varepsilon_i \mid X_i = x_i$, equivalently $A_i \perp Y_i \mid X_i = x_i$. As a secondary property, an unbiased appraisal satisfies $\mathbb{E}[\xi_i] = 0$, which fails, for instance, when the appraisal departs from the price upward more often, or by larger amounts, than downward. The two properties are distinct, since exact copying,

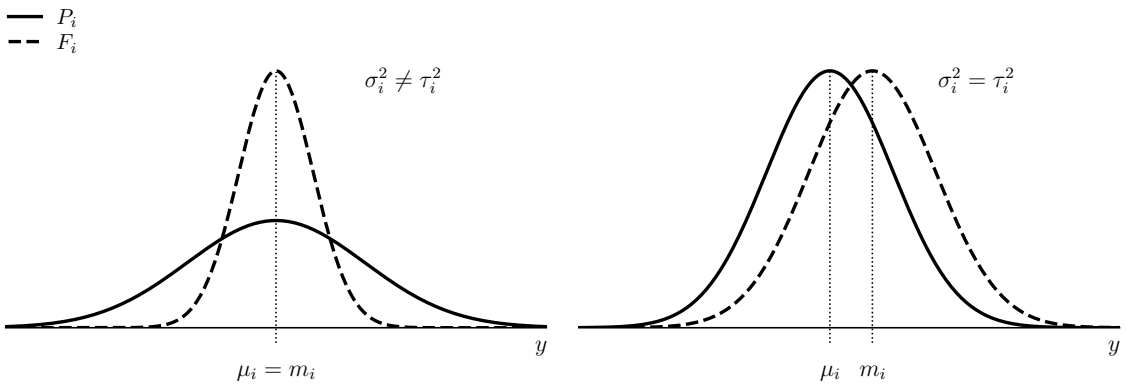
for instance, is unbiased yet fully dependent on the realized price.

Beyond providing the framework for these tests, the distinction between P_i and its location μ_i also clarifies what a statistical valuation can offer beyond a single number. An appraisal reduces the valuation to a point estimate of μ_i, whereas $\widehat{F}_i$ describes the entire range of possible resale prices and their probabilities, in line with the broader shift from point to distributional forecasting (Dawid 1984; Granger and Pesaran 2000; Garratt et al. 2003), including its use in financial risk management. For a bank, the features beyond location are directly relevant: the dispersion determines how likely a resale is to fall short of the outstanding loan, and the lower tail determines how large that shortfall can become, which Section 6 formalizes through the expected residual debt. Figure 1 illustrates why location and dispersion both matter for this purpose. In Panel (a), the predictive distribution is correctly centered but understates the dispersion of possible transaction prices, thereby understating market uncertainty and the probability of more extreme resale outcomes. In Panel (b), the dispersion is correctly captured but the predicted center is displaced from the true expected market value, leading to a systematically mislocated valuation. A predictive distribution is therefore only a useful alternative if it is calibrated in both respects, which we assess in Appendix A.3.

3 Dependence

This section develops the two analyses of Hypothesis 1 announced in Section 1, a test of independence and a regression that quantifies the coupling, and closes with a secondary test of the difference between the unconditional means of A_i and Y_i. Both analyses use the predictive distributions $\widehat{F}_i$ of Section 2, with mean $\widehat{m}_i$ and standard deviation $\widehat{\tau}_i := \sqrt{\mathbb{V}_{\widehat{F}_i}(Y_i)}$, as substitutes for the unknown conditional price distributions P_i. Appendix A details their construction, and Appendix A.3 evaluates their calibration against realized transaction

Figure 1: Illustrative misspecification of the predictive market-value distribution



(a) Different variances (b) Different means

NOTE: The true conditional distribution P_i is shown alongside a model-implied distribution F_i. In Panel (a), the means coincide but the model-implied spread τ_i is smaller. In Panel (b), the spreads coincide but the means differ.

prices. Section 3.1 uses the $\widehat{F}_i$ to construct the distribution of the appraisal–price discrepancy under conditional independence and thereby tests Hypothesis 1, and Section 3.2 derives the reference values against which the test statistic can be read. Section 3.3 quantifies the coupling by regressing the deviation of the appraisal from $\widehat{m}_i$ on that of the transaction price, and Appendix E.5 assesses how errors in $\widehat{m}_i$ affect the estimated slopes. Section 3.4 relates these slopes to the dependence structure of the pair (A_i, Y_i) in a copula framework, and Section 3.5 tests Hypothesis 2.

3.1 The coupling statistic

Appraisal independence is formalized by the following hypothesis.

Hypothesis 1. *Conditional on the characteristics X of the dwelling, the appraisal A is independent of the transaction price Y. More specifically, we test*

$$H_0 : \; A \perp Y \mid X \qquad \textit{vs.} \qquad H_1 : \; A \not\perp Y \mid X.$$

Let J_i denote the conditional law of (A_i, Y_i) given $X_i = x_i$, with marginals Q_i and P_i

respectively. Under the null of Hypothesis 1, J_i is the product of its marginals, $J_i = Q_i \otimes P_i$, where $\otimes$ denotes the product measure, so that the appraisal carries no information about the realized price beyond what the characteristics already contain. The observable content of Hypothesis 1 is that conditioning on the appraisal leaves the conditional law of the price unchanged,

$$Y_i \mid (A_i = a_i,\ X_i = x_i) \overset{H_0}{\sim} P_i, \tag{1}$$

Since P_i is unknown, we implement (1) with the predictive distribution $\widehat{F}_i$, and refer to the result as the working null. We test it with a Monte Carlo test in the sense of Barnard (1963) and Hope (1968), in which the observed value of a statistic is ranked within a reference set of values simulated under the null. Hypothesis 1 leaves the law Q_i of the appraisal unrestricted, but conditioning on the observed appraisals and characteristics removes Q_i, because under Hypothesis 1 the price then has law P_i whatever Q_i is. Since $\widehat{F}_i$ is estimated on a separate training sample, it is, conditional on that sample, a known distribution, so that with $\widehat{F}_i$ in place of P_i the working null fully specifies the distribution to simulate from. This is the situation in which such a test is exact, as observed by Barnard (1963) and formalized by Hope (1968). The reference set is generated by holding each a_i fixed and drawing a new price from $\widehat{F}_i$ for every dwelling, so that prices are formed independently of the appraisals without any assumption on how the latter are formed. The draws use inverse-transform sampling; Appendix A.2 gives the details.

Anchoring appears as an unusually small difference between A_i and Y_i. This discrepancy is standardized because $\widehat{\tau}_i$ varies substantially across dwellings, so that the same absolute difference is stronger evidence of anchoring when the conditional price distribution is wide, since agreement is then less likely to arise by chance. This gives

$$S_A = \frac{1}{n} \sum_{i=1}^{n} \frac{(a_i - Y_i)^2}{\widehat{\tau}_i^2},$$

with small values as evidence against Hypothesis 1 in the direction predicted by transaction-price anchoring.

Under the working null, the terms of S_A are independent but not identically distributed, since each depends on the observed appraisal a_i and on the shape of $\widehat{F}_i$. The distribution of S_A is therefore not available in closed form and is obtained by simulation, which also preserves the exactness of the test. For replication $b \in \{1, \ldots, B\}$, we draw $Y_i^{(b)} \sim \widehat{F}_i$ independently across dwellings, keep every appraisal at its observed value, and compute

$$S_A^{(b)} = \frac{1}{n} \sum_{i=1}^{n} \frac{\left(a_i - Y_i^{(b)}\right)^2}{\widehat{\tau}_i^2}.$$

Under the working null Y_i and $Y_i^{(b)}$ have the same conditional distribution for every dwelling, with a_i, x_i and $\widehat{\tau}_i$ fixed, so $S_A^{\text{obs}}, S_A^{(1)}, \ldots, S_A^{(B)}$ are exchangeable and the rank of S_A^{obs} among them is uniform on $\{1, \ldots, B+1\}$ (Besag and Clifford 1989). Since the alternative predicts small values, the lower-tail Monte Carlo p-value is

$$p_{\text{MC}} = \frac{1 + \sum_{b=1}^{B} \mathbb{1}\{S_A^{(b)} \leq S_A^{\text{obs}}\}}{B+1},$$

with $\mathbb{1}\{\cdot\}$ the indicator function. Let $\mathcal{G}$ denote the σ-algebra generated by $(A_i, X_i)_{i=1}^n$ and the training sample, and let $\mathbb{P}_0$ denote the law under which, conditional on $\mathcal{G}$, the prices are independent with $Y_i \sim \widehat{F}_i$. Then $\mathbb{P}_0(p_{\text{MC}} \leq \alpha \mid \mathcal{G}) \leq \alpha$ for every $\alpha \in (0, 1)$, without any distributional assumption on S_A (Hope 1968).

Validity under the null of Hypothesis 1 itself depends on how well the $\widehat{F}_i$ approximate the P_i. Predictive distributions that are systematically too wide, for instance, would make the simulated values $S_A^{(b)}$ large relative to S_A^{obs}, and so produce evidence of anchoring for reasons unrelated to the appraiser. Appendix A.3 examines this through the price-calibration statistic S_P.

The test rejects for small S_A and is not omnibus against Hypothesis 1: as in Hope (1968),

the statistic is chosen to measure a specific departure from the null, and the direction of rejection follows from the alternative. Consider an appraiser who systematically copies the transaction price and adds a constant markup $\gamma \in \mathbb{R}$, so that $A_i = Y_i + \gamma$. Then $\xi_i = \varepsilon_i + \gamma$, implying perfect dependence between the two errors. Nevertheless, $S_A = n^{-1} \sum_{i=1}^{n} \gamma^2 / \widehat{\tau}_i^2$, which need not be small when γ is large relative to the typical predictive spread. The test is therefore specifically powered against transaction-price anchoring: a small S_A is evidence against independence in that direction, whereas a large value does not constitute evidence that independence holds.

3.2 Implied lower bound

The simulation calibrates S_A against the observed appraisals, but the statistic also has a fixed reference point. Under Hypothesis 1 the cross term in the conditional second moment of the appraisal-price difference, $\Delta_i \coloneqq A_i - Y_i = \xi_i - \varepsilon_i$, vanishes, leaving

$$\mathbb{E}\big[\Delta_i^2\big] = \mathbb{E}\big[\xi_i^2\big] + \sigma_i^2 \;\geq\; \sigma_i^2, \tag{2}$$

so that $\mathbb{E}[S_A] \geq 1$ when $\widehat{\tau}_i = \sigma_i$. An appraiser who locates μ_i exactly still faces the market variation ε_i, and any appraisal error adds to the discrepancy.

The randomization distribution sits above this reference, because it conditions on the observed appraisals. Define the product measure $\widehat{F}^{(n)} \coloneqq \bigotimes_{i=1}^{n} \widehat{F}_i$. Each simulated price is drawn from $\widehat{F}_i$, which has mean $\widehat{m}_i$ and variance $\widehat{\tau}_i^2$, so

$$\mathbb{E}_{\widehat{F}^{(n)}}\big[S_A^{(b)}\big] = 1 + \frac{1}{n} \sum_{i=1}^{n} \frac{(a_i - \widehat{m}_i)^2}{\widehat{\tau}_i^2}, \tag{3}$$

which equals one only if every appraisal coincides with the center of its predictive distribution. Neither reference assumes an unbiased appraiser. In Equation (3) displacement raises the benchmark rather than S_A^{obs}, and Appendix D shows the same for Equation (2).

3.3 Regression

The test developed in Section 3.1 formally tests the independence of the appraisals and the transaction prices. In this section we regress the deviation of the appraisal from the predictive (conditional) mean on the corresponding deviation of the transaction price, to quantify how strongly appraisals depend on transaction prices. We estimate

$$a_i - \widehat{m}_i = \alpha + \beta\,(y_i - \widehat{m}_i) + \zeta_i, \tag{4}$$

where both sides are centered on $\widehat{m}_i$, removing the value component that Y_i and A_i share through the characteristics x_i. The slope β measures the share of the price deviation that passes into the appraisal. If the appraisal is formed independently of the realized transaction price it is close to 0, whereas it equals 1 under exact copying. Since $\widehat{m}_i$ need not equal μ_i, its error enters both sides of (4) and pulls the slope up. Appendix E.5 shows that the part of this error due to estimation shrinks with the size n_0 of the training sample and disappears as n_0 tends to infinity. The intercept α measures the average deviation of the appraisal from $\widehat{m}_i$ that the price does not explain.

As discussed in Section 1 (see also Section 4), appraisals and transaction prices are reported in different denominations: 95.1% of appraisals are multiples of €5,000, against only 47.7% of transaction prices. An appraiser who copies the price may therefore report it rounded and agree with it only up to that rounding. To account for this, we introduce a rounding step $\kappa \geq 0$ and the rounding operator $[z]_\kappa \coloneqq \kappa\,\mathrm{round}(z/\kappa)$ for $\kappa > 0$, which rounds z to the nearest multiple of κ, with halfway cases rounded to the nearest even multiple, and set $[z]_0 \coloneqq z$. A first approach brings both amounts to the same scale before taking their difference, $\tilde{\Delta}_i(\kappa) \equiv \tilde{\Delta}(a_i, y_i; \kappa) \coloneqq [a_i]_\kappa - [y_i]_\kappa$. However, since our interest is in the difference between the appraisal and the transaction price, we instead round this difference directly, $\Delta_i(\kappa) \equiv \Delta(a_i, y_i; \kappa) \coloneqq [a_i - y_i]_\kappa$, so that $\Delta(a_i, y_i; \kappa) = 0$ if and only if

$|a_i - y_i| \leq \kappa/2$. Neither form of rounding is stronger than the other, but the two never differ by more than one step, $|\Delta_i(\kappa) - \tilde{\Delta}_i(\kappa)| \leq \kappa$. Appendix B proves this result and reports the shares based on $\tilde{\Delta}_i(\kappa)$ as a robustness check. We use the rounding operator in two ways to examine what drives the estimate of β: the first asks whether $\hat{\beta}$ depends on the difference in denominations, the second whether the dependence extends beyond appraisals that copy the price.

First, we round both amounts to multiples of κ before centering them and estimate

$$[a_i]_\kappa - \widehat{m}_i = \alpha(\kappa) + \beta(\kappa)\big([y_i]_\kappa - \widehat{m}_i\big) + \zeta_i, \tag{5}$$

where $\kappa = 0$ reproduces (4). This puts the appraisal and the price on the same scale, so that an appraisal that copies the price up to rounding, $[a_i]_\kappa = [y_i]_\kappa$, is treated as an exact copy. Both amounts are rounded in the same way and move by at most $\kappa/2$. If $\hat{\beta}(\kappa)$ remains close to $\hat{\beta}(0)$, the estimated dependence does not arise from the different denominations in which appraisals and prices are reported. Section 5.2 reports the estimates for $\kappa \in \{€2{,}500, €5{,}000, €10{,}000\}$.

Second, we estimate (4) on the dwellings with $\Delta_i(\kappa) \neq 0$ only, that is, with $|a_i - y_i| > \kappa/2$. This excludes the appraisals that equal the price up to rounding and shows whether the dependence persists among appraisals that lie away from the price. For the exclusion, we report the results for a fine grid of κ in Figure 8.

Since the centered regression (4) measures deviations in euros, expensive dwellings, whose deviations tend to be larger in absolute terms, receive more weight in its estimates. To verify that our results are not driven by these dwellings, we rescale both deviations by the predictive spread $\widehat{\tau}_i$ as a robustness check, which makes them comparable across dwellings and yields

$$\frac{a_i - \widehat{m}_i}{\widehat{\tau}_i} = \alpha + \beta \, \frac{y_i - \widehat{m}_i}{\widehat{\tau}_i} + \zeta_i, \tag{6}$$

so that each deviation is measured relative to the market uncertainty for that dwelling. We apply both checks to (6) as well, dividing the rounded deviations by $\widehat{\tau}_i$. All standard errors are heteroskedasticity-robust.

3.4 Dependence structure

The estimated slope in Section 3.3 summarizes the dependence between appraisal and transaction price in a single number, but a single slope cannot distinguish between different forms of dependence, such as exact copying of a fraction of prices and partial tracking of all of them. A copula representation of the independence restriction formulated in Hypothesis 1 separates the dependence structure from the two marginal distributions and thereby makes these forms explicit.

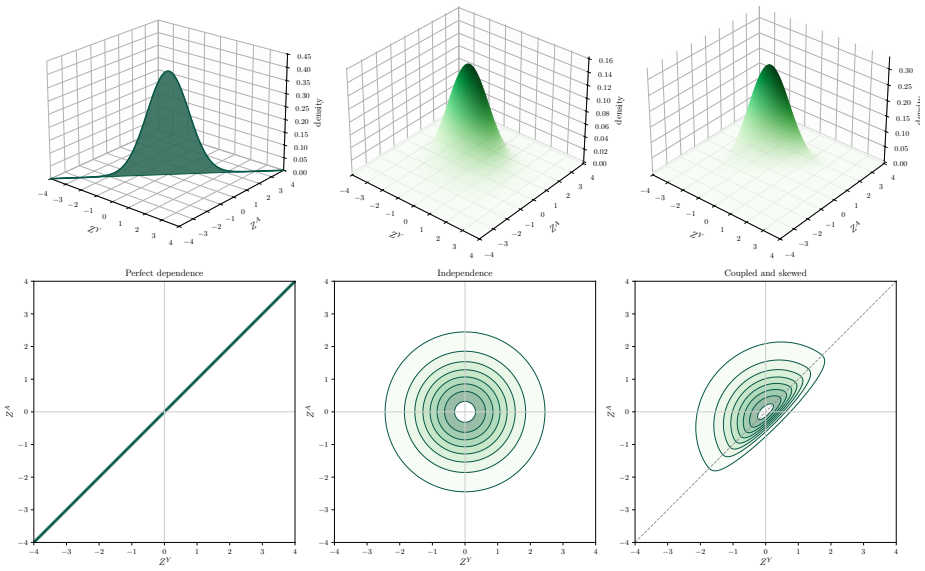
Recall that J_i denote the conditional law of (A_i, Y_i) given $X_i = x_i$, with marginals Q_i and P_i. By the theorem of Sklar (1959), for almost every x_i there exists a conditional copula C_i such that

$$J_i(a, y) = C_i\big(Q_i(a), P_i(y)\big), \qquad \text{for all } (a, y) \in \mathbb{R}^2,$$

which is unique when the conditional marginals are continuous. See Patton (2009, 2012) for reviews of copula models in financial and economic applications. Conditional independence, $J_i = Q_i \otimes P_i$, corresponds to $C_i(u, v) = uv$ and is thus a restriction on the dependence structure and not on either marginal. In particular, this means that an appraiser may be correctly centered on average while still tracking the particular transaction-price realization too closely.

Standardizing A_i and Y_i by the conditional mean and standard deviation of Y_i, one obtains $Z_i^A := \frac{A_i - \mu_i}{\sigma_i} = \frac{\xi_i}{\sigma_i}$ and $Z_i^Y := \frac{Y_i - \mu_i}{\sigma_i} = \frac{\varepsilon_i}{\sigma_i}$. Conditional on $X_i = x_i$ these are affine transformations of A_i and Y_i, so they preserve the conditional copula.

Figure 2: Copula dependence structures



NOTE: Conditional dependence structures for the standardized error pair (Z^A, Z^Y). The left panel depicts perfect positive dependence, illustrated by the special case $Z^A = Z^Y$ a.s., the middle panel illustrates conditional independence and the right panel contains positive coupling combined with upward displacement, the pattern predicted when appraisals anchor on the realized transaction price and tend to depart from it upward.

Figure 2 illustrates the dependence between the two variables through Z_i^A and Z_i^Y. These variables are the infeasible counterparts of the dependent and independent variable in (6), respectively. The diagonal line in the left panel is the special case $Z^A = Z^Y$, in which the appraisal reproduces the dwelling's realized deviation from μ_i; in the infeasible version of (6) it corresponds to $\alpha = 0$, $\beta = 1$ and $\zeta_i = 0$, whereas $\beta = 1$ alone is also compatible with independent noise around the line. The case is an instance of perfect positive dependence, the copula $C_i(u, v) = \min(u, v)$, which more generally describes any appraisal that is a strictly increasing function of the price, such as the constant markup $A_i = Y_i + \gamma$ of Section 3.1. The middle panel shows conditional independence, under which the conditional joint distribution factorizes. The resulting copula corresponds to the null distribution of Hypothesis 1. The right panel shows an intermediate structure, where concentration around the 45° line represents transaction-price anchoring: the appraisal follows the realized price more closely than conditional independence permits. A systematic

shift above that line, so that $Z_i^A - Z_i^Y = (A_i - Y_i)/\sigma_i$ is positive on average, represents directional displacement, which the intercept α picks up.

3.5 Direction

Besides independence, we examine the direction of appraisal errors, formulated through the following hypothesis.

Hypothesis 2. *On average, the appraisal A equals the expected market value μ of the dwelling. Otherwise, the appraisals bear a positive bias. In mathematical terms, we test*

$$H_0 : \ \mathbb{E}[A - \mu] = 0 \qquad \textit{vs.} \qquad H_1 : \ \mathbb{E}[A - \mu] > 0.$$

This hypothesis concerns $\mathbb{E}[\xi_i]$, with $\xi_i \coloneqq A_i - \mu_i$, which is unobserved because μ_i is unobserved. The appraisal–price difference $\Delta_i = A_i - Y_i$ replaces the unobserved center by the realized price and is observed. Since $\Delta_i = \xi_i - \varepsilon_i$ and $\mathbb{E}[\varepsilon_i] = 0$ by construction, linearity gives $\mathbb{E}[\Delta_i] = \mathbb{E}[\xi_i]$, so the two have the same mean. Taking expectations over X_i therefore makes Hypothesis 2 equivalent to

$$H_0 : \mathbb{E}[\Delta_i] = 0 \qquad \text{vs.} \qquad H_1 : \mathbb{E}[\Delta_i] > 0,$$

which no longer involves the predictive model. Inference uses $\bar{\Delta} = n^{-1} \sum_{i=1}^{n} \Delta_i$ and the t statistic $\bar{\Delta}/(s_\Delta/\sqrt{n})$, with s_Δ the sample standard deviation of the differences. Confidence intervals follow from resampling dwellings with replacement.

4 Data

Through Viisi, an independent Dutch mortgage advisory firm, we obtained 800 appraisal reports covering the period from December 2025 to August 2026. The sample is therefore

a snapshot of the housing market in which each dwelling enters once, at the time of its transaction, which complements time-series studies such as Bolt et al. (2019), who identify long-lasting temporary deviations of Dutch and other house prices from a rent-based fundamental value. Such reports are typically validated by the Nederlands Woning Waarde Instituut (NWWI), which does not itself judge the appraised market value, but investigates the difference when a second report on the same dwelling with a substantially different value is submitted within a short period (Nederlands Woning Waarde Instituut 2024). These safeguards apply to the report once it is submitted, and do not exclude consultation with the parties to the transaction beforehand.

We matched each appraisal report to the corresponding transaction and dwelling information required for our analysis. We then excluded 106 dwellings, either because a characteristic required by the predictive model was missing, so that the predictive distribution would rely on a default value, or because the dwelling falls outside the model's training and application domain. The final sample contains $n = 694$ dwellings, for which we construct dwelling-specific predictive price distributions using the latest version of the prediction model by SlimBieden. This model constructs a predictive distribution from characteristics of the dwelling and macroeconomic variables. Importantly, these inputs contain only information available before the transaction of the dwelling took place, so that neither the transaction price nor the appraisal enters the prediction. The predictive distribution is therefore formed without reference to either quantity.

A potential concern is selection into completed transactions. De Nederlandsche Bank (2019, p. 20) notes that, although quantitative evidence is unavailable, lenders and real estate agents report that transactions are rarely terminated because of a low appraisal, and therefore judges the resulting selection effect unlikely to materially affect its findings. Mortgage advisers at Viisi independently expressed the same view for our setting. We therefore likewise expect such selection effects to play only a limited role.

4.1 Summary statistics

Table 1 reports summary statistics for the final sample. These statistics are pooled over all dwellings and thereby averaged over the characteristics X. The median transaction price and the median appraisal coincide at €550,000, while the mean of the appraisal is slightly larger. Furthermore the table reports the standard deviation and some quantiles of the appraisals and transaction price.

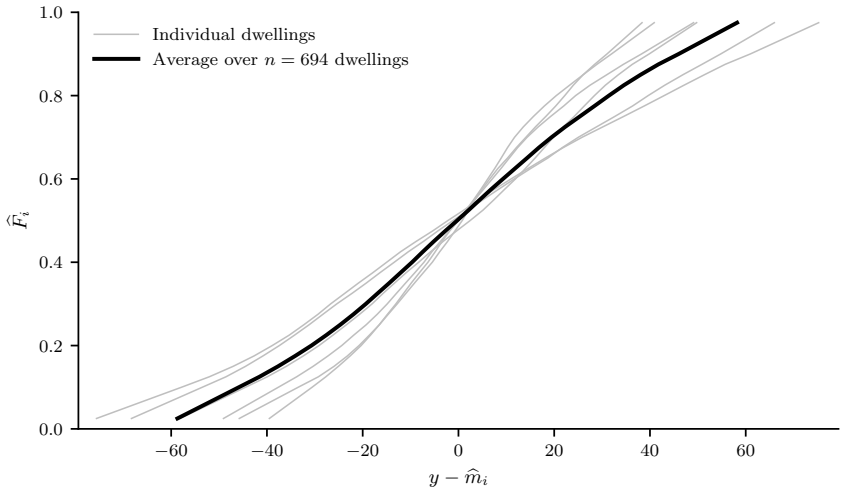
Table 1: Appraisals and transaction prices

	Mean	SD	Min	5%	25%	50%	75%	95%	Max
Appraisal	609.7	220.6	305.0	354.3	460.0	550.0	700.0	1,076.8	1,550.0
Price	607.8	219.7	300.0	353.0	460.0	550.0	697.5	1,071.1	1,500.0

NOTE: Summary statistics of the appraisals a_i and transaction prices y_i of the $n = 694$ dwellings, in thousands of euros. SD is the sample standard deviation and the quantiles are empirical quantiles.

Conditional on the characteristics $X_i = x_i$ of each dwelling, the predictive distribution $\widehat{F}_i$ provides an estimate of the conditional price distribution P_i. Figure 3 shows six examples of these distribution functions after centering by $\widehat{m}_i$, together with the average of each centered quantile over all 694 dwellings.

Figure 3: Example of distribution functions



NOTE: The centered distribution functions of six dwellings in the data set as gray lines accompanied with the average centered distribution function in a solid line; the horizontal axis shows $y_i - \widehat{m}_i$ in thousands of euros.

4.2 Appraisal–price differences

Appraisals exceed the transaction price more often than they fall below it, 43.8% against 34.9% of dwellings, and coincide with it exactly for 21.3%. Exact agreement understates how often the appraisal follows the price, however, because appraisals and prices are reported in different monetary denominations: 95.1% of appraisals are multiples of €5,000, against 47.7% of transaction prices (Table 2). An appraiser who adopts the price but reports it as a round amount therefore produces an appraisal that differs from the price by up to half a step, even though it is set relative to the price.

To compare appraisals and prices in a common denomination, Table 2 uses the rounded difference $\Delta_i(\kappa) = [a_i - y_i]_\kappa$ of Section 3.3. It equals zero if and only if $|a_i - y_i| \leq \kappa/2$, and equals $\pm\kappa$ when the appraisal lies roughly one step of κ above or below the price.

Table 2: Agreement between appraisals and transaction prices under rounding

	$\kappa =$ €2,500	$\kappa =$ €5,000	$\kappa =$ €10,000
Appraisal a_i is a multiple of κ	97.1%	95.1%	53.3%
Price y_i is a multiple of κ	52.3%	47.7%	25.5%
$\Delta_i(\kappa) = 0$	31.1%	41.5%	60.8%
$0 \leq \Delta_i(\kappa) \leq \kappa$	40.6%	56.3%	71.0%
$-\kappa \leq \Delta_i(\kappa) \leq 0$	38.2%	50.6%	69.0%

NOTE: Shares of all $n = 694$ dwellings. Amounts are rounded to the euro, and an amount is a multiple of κ when it equals $j\kappa$ for an integer j. $\Delta_i(\kappa) = [a_i - y_i]_\kappa$ is defined in Section 3.3; $\Delta_i(\kappa) = 0$ if and only if $|a_i - y_i| \leq \kappa/2$. The last two rows count the appraisals that equal the price up to rounding or lie one step of κ above, respectively below, it, where one step above corresponds to $\kappa/2 < a_i - y_i < 3\kappa/2$. Without rounding, $a_i = y_i$ for 21.3% of dwellings. Appendix B reports the corresponding shares for $\tilde{\Delta}_i(\kappa) = [a_i]_\kappa - [y_i]_\kappa$.

With $\kappa =$ €5,000, the denomination of nearly all appraisals, the appraisal equals the price up to rounding for 41.5% of dwellings, almost twice the share of exact copies. This agreement is not symmetric: 14.8% of appraisals lie one step above the price, against 9.1% one step below it. The same asymmetry appears for the other values of κ; Figure 6 in Appendix B reports these statistics, alongside those using $\tilde{\Delta}_i(\kappa)$, for more thresholds κ and two steps of κ above or below the price. We take as reference denomination the largest κ

of which nearly all appraisals are multiples, which is $\kappa = €5{,}000$; for $\kappa = €10{,}000$, this share drops to 53.3%.

Table 3: Appraisal–price differences at $\kappa = €5{,}000$

Panel A: Copies of the transaction price	Count	Percentage
Exact copy, given y_i is a multiple ($n = 331$)	145	43.8%
Exact copy, given y_i is not a multiple ($n = 363$)	3	0.8%
Rounded copy, given y_i is not a multiple ($n = 363$)	131	36.1%
Panel B: Departures by a multiple of €5,000	Count	Percentage
Above the price ($n = 174$)	120	69.0%
Below the price ($n = 174$)	54	31.0%
Panel C: Contribution to the mean of $a_i - y_i$	Count	Contribution
Departure by a multiple of €5,000	174	€1,996
Other departure	372	−€90
All dwellings	694	€1,906

NOTE: The sample contains $n = 694$ dwellings, and amounts are rounded to the euro. An exact copy has $a_i = y_i$. A rounded copy has $a_i = [y_i]_{5{,}000}$ for a price y_i that is not a multiple of €5,000. A departure is a multiple of €5,000 when $a_i - y_i = k\times$ €5,000 for an integer $k \neq 0$; Figure 7 shows these departures by step size. Percentages are shares of the subgroup of size n stated in the row. The contribution is $n^{-1}\sum_i (a_i - y_i)$ over the respective group, with $n = 694$, so that the contributions add up to the mean difference; exact copies contribute nothing.

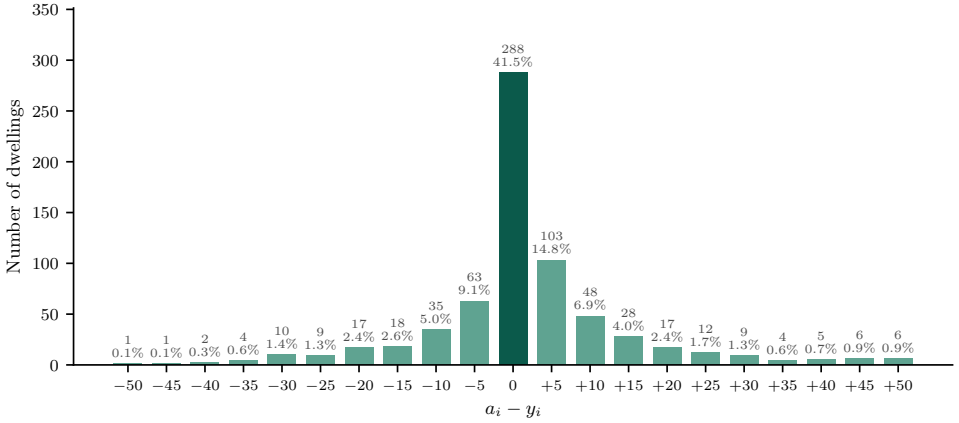
Table 3 looks more closely at the reference denomination $\kappa = €5{,}000$. Exact copies are almost entirely confined to prices that are themselves multiples of €5,000: they occur for 43.8% of these prices, against 0.8% of the other prices (Panel A). For 36.1% of these other prices, the appraisal instead equals the price rounded to the nearest multiple of €5,000.

Since nearly all appraisals are multiples of €5,000, departures from the price are often a multiple of €5,000 as well, and these lie above the price in 69.0% of cases (Panel B), at each step size (Figure 7 in Appendix B). The other departures are evenly split: 49.5% have an appraisal above the transaction price. Figure 4 groups all differences by $\Delta_i(5{,}000)$ and shows that the upward asymmetry persists once appraisals and prices are compared on the same denomination, so that it is not confined to departures by exact round amounts.

The departures by multiples of €5,000 account for more than the entire mean appraisal–price difference of €1,906, while the other departures contribute slightly negatively (Panel

C). On average, appraisals thus lie close to the transaction price, as an unbiased estimate of the market value would, but these patterns suggest that this agreement arises at least partly from copying the transaction price, which would violate the null of Hypothesis 1. These groupings are descriptive and were defined after inspection of the data; formal tests of independence and bias are presented in Section 5.

Figure 4: Appraisal–price differences, grouped by $\Delta_i(5{,}000)$



NOTE: Number of dwellings by the rounded difference $\Delta_i(5{,}000) = [a_i - y_i]_{5{,}000}$, so that the middle bin covers $|a_i - y_i| \leq €\,2{,}500$; it consists almost entirely of exact and rounded copies. Not shown are 2 dwellings with $a_i - y_i < -52{,}500$ and 6 with $a_i - y_i > 52{,}500$. The horizontal axis is in thousands of euros; percentages are shares of all 694 dwellings.

4.3 Comparison with earlier findings

Table 4 places the shares in Panel A of Table 3 alongside the figures of De Nederlandsche Bank (2019), Nederlands Register Vastgoed Taxateurs (2026) and Minister voor Volkshuisvesting en Ruimtelijke Ordening (2025) discussed in Section 1. Our shares of 43.8% above, 21.3% equal to and 34.9% below the transaction price are close to both the NRVT figures for 2019–2025 and the NHG figures for 2021 to mid-2025. Relative to the period studied by DNB, this suggests an improvement, but the same pattern persists: appraisals remain more likely to lie above than below the transaction price, and approximately one in five coincide with it exactly. In the United States, Calem et al. (2021) find that of 1.3 million appraisals for home purchase mortgages in 2013–2015, 64% lie above the contract

price, 30% equal it and only 6% lie below it, so that the asymmetry is considerably more pronounced there. Part of this difference may be institutional, since US lenders base the loan-to-value ratio on the lesser of the appraisal and the contract price (Calem et al. 2021, p. 138), so that an appraisal above the price brings no benefit.

These comparisons are descriptive, in that they document how often appraisals agree with transaction prices, but they do not by themselves provide a statistical test of whether appraisals are formed independently of those prices. Whether a given degree of agreement is unusual depends on the dispersion of possible transaction prices for the dwelling: agreement that is unremarkable when the conditional price distribution is narrow is unlikely when it is wide. None of the earlier studies accounts for this dispersion or formally tests independence; the Monte Carlo test of Section 3 does both, and Section 5 reports its outcome.

Table 4: Comparison with DNB, NRVT and NHG

	Period	Above price	Exact match	Below price
DNB	2012–2017	ca. 60%	ca. 33%	ca. 8%
NRVT	2019–2025	42.8%	21.6%	35.6%
NHG	2021–mid-2025	46%	19%	30%
This paper	Dec 2025–Aug 2026	43.8%	21.3%	34.9%

NOTE: Shares of appraisals above, equal to and below the transaction price. DNB figures are from De Nederlandsche Bank (2019, p. 16), who report them as approximate: nearly 60% above, one third equal and less than 8% below the transaction price. NRVT figures are averages over 1,064,141 appraisals from Nederlands Register Vastgoed Taxateurs (2026, Table 3.1); for 2025 alone they are 44.3%, 17.0% and 38.7%. NHG figures are from Minister voor Volkshuisvesting en Ruimtelijke Ordening (2025); their reported shares add up to 95% with no mention of what the remaining 5% consists of.

5 Results

This section reports the results of the analyses of Section 3. Section 5.1 reports the outcome of the Monte Carlo test of Hypothesis 1 developed in Section 3.1, and Section 5.2 presents the regression estimates of Section 3.3 across rounding levels κ. Section 5.3 then reports the model diagnostics that support the use of the predictive distributions $\widehat{F}_i$ in place of P_i,

together with the results for the test of Hypothesis 2.

5.1 Test of appraisal independence

The descriptive patterns in Section 4.2 suggest that appraisals depend on transaction prices, and the Monte Carlo test establishes this formally. The observed statistic is $S_A^{\text{obs}} = 0.2535$, with a bootstrap 95% confidence interval of (0.2049, 0.3076), and the Monte Carlo p-value is 0.00001. Hypothesis 1 is therefore rejected: appraisals and transaction prices are not independently formed. Given the descriptive patterns, the strength of this rejection is not surprising. The test, however, also applies where such patterns are less pronounced, and can thus establish whether a future reform of the appraisal system restores independence.

The size of the departure is clear against both references of Section 3.2. First, independence implies $\mathbb{E}[S_A] \geq 1$ by (2), whatever the accuracy of the appraiser, and the entire bootstrap interval lies below this value. Second, the randomization distribution has mean 1.8656 and a central 95% range of (1.7074, 2.0273), so that the observed statistic is approximately 14% of its benchmark. This benchmark exceeds one by the average squared displacement of the appraisals from the predictive centers in (3), and thus reflects the observed appraisals rather than a hypothetical appraiser who reports μ_i exactly. Against either reference, realized transaction prices lie far closer to their appraisals than independently formed prices would.

The rejection cannot be attributed to displacement of the appraisals from the predictive centers. The randomization conditions on the observed a_i, so that any displacement is held fixed across replications while only the transaction price is redrawn from $\widehat{F}_i$, and Appendix D shows that displacement can only raise, never lower, the population bound.

5.2 Anchoring regressions

The Monte Carlo test establishes that appraisals depend on transaction prices, but not how strongly. Table 5 reports the estimates of regressions (4) and (5), which quantify this dependence, alongside the estimates for the exclusion regression described in Section 3.3. On all 694 dwellings, the slope is $\hat{\beta} = 0.802$ (0.023): appraisals pass on roughly four fifths of the deviation of the realized price from the predictive center, far more than independent formation allows, yet less than exact copying would imply (Panel A). Rounding both amounts to a common denomination before estimation leaves this estimate essentially unchanged: for $\kappa = €5{,}000$, the share of dwellings whose rounded appraisal equals the rounded price rises from 21.3% to 39.9%, while the slope remains 0.803 (0.023) (Panel B). The dependence thus does not arise from the different denominations in which appraisals and prices are reported.

Table 5: Anchoring regression

	n	$\hat{\alpha}$	$\hat{\beta}$	$p_{\beta=1}$	R^2
Panel A: Centered regression, (4)					
All dwellings	694	814 (538)	0.802 (0.023)	< 0.001	0.778
Panel B: κ-rounding, (5)					
$\kappa = €2{,}500$	694	870 (538)	0.803 (0.023)	< 0.001	0.779
$\kappa = €5{,}000$	694	970 (540)	0.803 (0.023)	< 0.001	0.778
$\kappa = €10{,}000$	694	829 (555)	0.795 (0.024)	< 0.001	0.764
Panel C: κ-exclusion, $\Delta_i(\kappa) \neq 0$					
Exact copies, $a_i = y_i$	546	1,388 (659)	0.763 (0.027)	< 0.001	0.739
$\kappa = €2{,}500$	478	1,670 (745)	0.748 (0.029)	< 0.001	0.722
$\kappa = €5{,}000$	406	2,034 (863)	0.726 (0.031)	< 0.001	0.699
$\kappa = €10{,}000$	272	2,062 (1,214)	0.621 (0.041)	< 0.001	0.587

NOTE: Heteroskedasticity-robust (HC1) standard errors in parentheses; $\hat{\alpha}$ is in euros. Panel A estimates (4) on all dwellings. Panel B estimates (5) on all dwellings, with both amounts rounded to multiples of κ before centering. Panel C estimates (4) on the dwellings with $\Delta_i(\kappa) \neq 0$, that is, $|a_i - y_i| > \kappa/2$; its first row excludes the exact copies only. $p_{\beta=1}$ is the two-sided p-value of the t test of $\beta = 1$; $\beta = 0$ is rejected with $p < 0.001$ in every row. Appendix C reports the rescaled specification (6) and the exclusion for a fine grid of κ (Figure 8).

Excluding appraisals that equal the price lowers the slope, but does not remove it

(Panel C). Without the 148 exact copies, $\hat{\beta} = 0.763$ (0.027), and without all appraisals that equal the price up to rounding, $|a_i - y_i| \leq €2{,}500$, it is 0.726 (0.031). The slope declines further as the excluded range widens, since the excluded appraisals are those closest to the diagonal (Figure 8 in Appendix C). However, $\hat{\beta}$ remains significantly positive throughout, also for the 62 dwellings whose appraisal lies more than €25,000 from the price, with $\hat{\beta} = 0.423$ (0.108). The dependence on the transaction price is therefore not confined to appraisers who copy it: appraisals that differ from the price still move with it. A deviation from the price thus need not reflect an independent valuation, and screening for appraisals equal or close to the price would miss part of the dependence. The rescaled specification (6) leads to the same conclusions, with a full-sample slope of 0.790 (0.020), although its slope declines more steeply along the exclusion path (Appendix C).

The intercept is positive throughout but small, on the full sample $\hat{\alpha} = €814$ (538), and rounding leaves it of similar size. Once appraisals close to the price are excluded, it rises to about €2,000 for κ from €5,000. Appraisals that depart from the price thus tend to lie above the predictive center beyond what the price explains, consistent with the upward displacement examined in Section 5.3.

5.3 Model validity and direction

Conditional on the predictive distributions $\widehat{F}_i$, the Monte Carlo p-value is exact. However, inference for Hypothesis 1 itself requires that these distributions adequately approximate the conditional price distributions P_i. We therefore assess their calibration in both location and dispersion. Neither diagnostic provides evidence of material miscalibration, so that the small value of S_A cannot be attributed to predictive distributions that are too wide. Appendix A.3 reports the full calibration analysis, and Appendix A.5 compares the point-prediction accuracy of the model with the statutory accuracy thresholds for model-based residential valuations (Minister van Financiën 2026b).

The appraisal–price differences are also systematically positive. The mean difference is €1,906, with a bootstrap 95% confidence interval of (€779, €3,040), and the one-sided t test of Hypothesis 2 yields $t = 3.31$ and $p = 0.0005$, rejecting a zero mean difference in favor of upward displacement.

6 Collateral risk implications

Because the appraisal determines the maximum loan, the coupling documented in Section 5 directly affects the risk borne by borrower and lender. An appraisal that copies the transaction price no longer ties the loan to the value the dwelling can be expected to fetch on resale. When the price paid exceeds that value, the loan may exceed the expected resale value from the day of purchase, leaving the borrower with negative equity that the appraisal does not reveal. This risk materializes upon resale after a default. If the dwelling sells for Y_{new}, the lender recovers at most the outstanding loan L, and any surplus accrues to the borrower. A shortfall $(L - Y_{\text{new}})^+$, with $z^+ \coloneqq \max\{0, z\}$, remains owed by the borrower, since Dutch mortgages carry full recourse. The borrower thus loses the dwelling and still carries a residual debt. This debt becomes a loss for the lender to the extent that the borrower cannot repay it, which is likely because a default typically stems from an inability to pay (Gerardi et al. 2018). Given a default, the expected residual debt $\mathbb{E}[(L - Y_{\text{new}})^+]$ therefore measures the risk that the loan places on borrower and lender together. We show that an appraisal coupled to the transaction price raises this risk, first for a single loan in Example 2 and then under general conditions in Propositions 1 and 2, which compare valuations that copy the price with one that targets the expected resale value.

Example 2. *Returning to Example 1, suppose Laila borrows the full purchase price, so that the loan is $L^{\text{obs}} =$ €721,000. Relative to the appraisal of $a =$ €725,000, the loan-to-value ratio is 99.4%, but relative to the expected market value of €663,000 it is 108.7%. If she*

defaults and the dwelling is resold at its expected value, a residual debt of €58,000 remains, before any costs of a forced sale.

For illustration, let the conditional price distribution P_i of her dwelling be normal with mean $\mu_i = €663{,}000$ and standard deviation $\sigma_i = €45{,}000$. A resale then falls short of the loan with a probability of approximately 90%, and the expected residual debt is $\mathbb{E}[(L - Y_{\text{new}})^+] \approx €60{,}100$. Had the appraisal equaled the expected market value of €663,000, Laila would have had to finance the remaining €58,000 from her own savings; the probability of a shortfall would then have been 50%, with an expected residual debt of approximately €18,000.

To the extent that the appraisal follows the agreed price, a bid above the expected value tends to validate itself: the price supports the appraisal, the appraisal supports the loan, and the loan finances the price. Where the buyer borrows up to the appraisal, the part of the price paid above the expected value is then financed by the loan rather than from the buyer's own savings, and the resulting risk is borne by borrower and lender together upon resale.

To analyze this mechanism more generally, we consider a loan determined by a valuation $\hat{\mu}$ of the dwelling, an estimate of its expected resale value μ_Y. Dutch rules permit this valuation to be established through the purchase price itself, an appraisal or, under conditions, a statistical model (Minister van Financiën 2026a, art. 115, lid 6), and cap the mortgage at it (Minister van Financiën 2026b, art. 5, lid 1). A lender therefore extends a loan $L = \ell\hat{\mu}$ with $\ell \leq 1$, intending ℓ to describe the loan relative to μ_Y.

We compare three choices of $\hat{\mu}$, which mirror these valuation bases, by the expected residual debt they imply. The first is exact copying, $\hat{\mu}_1 = Y_{\text{old}}$, which corresponds to valuing the dwelling at its purchase price; as noted in Section 1, it is unbiased for μ_Y. The second is noisy copying, $\hat{\mu}_2 = Y_{\text{old}} + \eta$, with $\mathbb{E}[\eta \mid Y_{\text{old}}] = 0$ and (Y_{old}, η) independent of the resale price Y_{new}; it describes an appraisal that starts from the agreed price and departs

from it for reasons unrelated to that price, such as adjustments in steps of €5,000 that are equally likely in either direction. The third reports the mean of the predictive distribution, $\hat{\mu}_3 = \widehat{m}_Y$, which is formed from the characteristics alone and, for a model calibrated at the first moment, satisfies $m_Y \approx \mu_Y$.

Proposition 1. *Let Y_{old} and Y_{new} be independent with distribution function P and finite mean μ_Y, and let $\ell > 0$ be the loan-to-value ratio. Let η be integrable with $\mathbb{E}[\eta \mid Y_{\text{old}}] = 0$ and (Y_{old}, η) independent of Y_{new}. Then the strategies $\hat{\mu}_1 = Y_{\text{old}}$ and $\hat{\mu}_2 = Y_{\text{old}} + \eta$ satisfy*

$$\mathbb{E}\big[(\ell\mu_Y - Y_{\text{new}})^+\big] \le \mathbb{E}\big[(\ell\hat{\mu}_1 - Y_{\text{new}})^+\big] \le \mathbb{E}\big[(\ell\hat{\mu}_2 - Y_{\text{new}})^+\big].$$

Proposition 1 shows that copying the transaction price raises the expected residual debt above the level attained by a valuation at μ_Y, which the model mean $\hat{\mu}_3$ targets, even though the copied price is unbiased for μ_Y, and that noisy copying raises it further. Proposition 2 shows that any appraisal displaced upward on average, whether copied or not, likewise exceeds this level.

Proposition 2. *Let Y_{new} have distribution function P with finite mean μ_Y, let $\ell > 0$ be the loan-to-value ratio, and let A be an integrable appraisal independent of Y_{new}. If $\mathbb{E}[A] \ge \mu_Y$, then*

$$\mathbb{E}\big[(\ell\mu_Y - Y_{\text{new}})^+\big] \le \mathbb{E}\big[(\ell A - Y_{\text{new}})^+\big].$$

The effect stems from the precision of the valuation as an estimator of μ_Y. A copied price is unbiased, but as a single draw from P it disperses around μ_Y as widely as the market itself, and the convexity of the residual debt turns this dispersion into a larger expected loss. The predictive mean m_Y draws on the many transactions on which the model is estimated, so that it disperses far less around μ_Y and, for instance when both valuations are normally distributed, leaves a smaller expected loss; Appendix E.4 illustrates

the difference in dispersion for a linear model.

Through the reported loan-to-value ratio, the coupling thus gives each mortgage a risk label that describes its collateral more favorably, on balance, than the collateral supports. Even symmetric coupling, without any bias, has asymmetric consequences upon default, since a loan that exceeds the resale price leaves a residual debt, while a loan that falls below it leaves a surplus that accrues to the borrower and does not offset shortfalls elsewhere. In a rising housing market these consequences are largely hidden, because appreciation covers the gap before a default occurs. Recent reports, however, point to a cooling of the Dutch housing market, led by Amsterdam (ABN AMRO 2026), which makes the findings more pressing. Labels that overstate the quality of the collateral are reminiscent of the run-up to the financial crisis of 2007–2008, when mortgages were treated as safe because the figures on which their risk was assessed described the underlying collateral more favorably than it warranted. Although the risks in the current market may be of a different scale, the underlying mechanism is similar: confidence in the collateral rests on a figure less independent than its role assumes.

6.1 Time effects

The size of the expected residual debt depends on the distribution of the market value of the dwelling at the time of default. If a dwelling is bought and a loan is taken out at time t_0, its market value has distribution $P_i^{(t_0)}$. Denote the time of default by $t_1 \geq t_0$ and the corresponding distribution of the market value of the dwelling by $P_i^{(t_1)}$.

If default occurs at $t_1 > t_0$, the two distributions need not be equal, that is, $P_i^{(t_0)} \neq P_i^{(t_1)}$. In a growing housing market, a plausible scenario is an increase in the center of the distribution, so that the value of the dwelling increases on average. Together with the repayment of the loan, this reduces the expected residual debt of the borrower. If the housing market cools, this scenario could just as well be reversed: the dwelling is expected

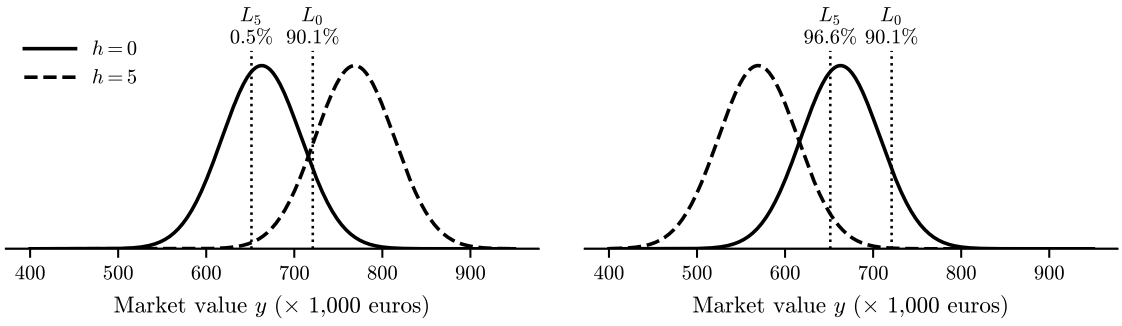
to fetch less on resale, and the expected residual debt increases. ABN AMRO (2026) expects price growth to slow from 3.8% in 2026 to 2.5% in 2027. Such a cooling need not stop at slower growth, since Dutch house prices can deviate from their fundamental value for long periods (Bolt et al. 2019), so that a period of rising prices can be followed by a sustained correction. Between August 2008 and mid-2013, prices of existing owner-occupied dwellings fell by 21.4% (Langenberg and Jonkers 2022).

Example 3. *Returning to Example 2, suppose Laila defaults after h years. Let the market value of her dwelling then be normal with standard deviation $\sigma = €45{,}000$ and mean $\mu_h = (1+g)^h\mu$, where $\mu = €663{,}000$ and g is a constant annual change in market value, and let her loan of $€721{,}000$ be repaid as a 30-year annuity at 4%, leaving an outstanding balance L_h (Appendix E.3). We consider a rising market with $g = 3\%$, in line with ABN AMRO (2026), and a declining market with $g = -3\%$, which over five years amounts to a decline of approximately 14%, milder than the 21.4% between 2008 and 2013 (Langenberg and Jonkers 2022).*

Figure 5 shows the market value at origination and after five years. At origination, a resale falls short of her loan with probability 0.90. If prices rise, the market absorbs her overbid, and after five years this probability has fallen to 0.005. If prices fall, it has risen to 0.97: despite five years of repayment, she is at more risk than on the day of purchase. Had the appraisal been $€663{,}000$, her loan would have started at the center of the distribution rather than in its tail, with a shortfall probability of 0.50 at origination and 0.74 after five years of decline.

The example shows that the consequences of a coupled appraisal depend on the market that follows. A rising market absorbs an overbid: within a few years the loan falls below the expected market value, and the excess risk disappears without borrower or lender having noticed it. A declining market does the opposite and adds to the shortfall that the

Figure 5: Market value of the dwelling and outstanding loan over time



(a) rising market, $g=3\%$

(b) declining market, $g=-3\%$

NOTE: Distribution $P_i^{(t_0+h)}$ of the market value of the dwelling in Example 2 at origination ($h=0$) and after five years ($h=5$), when the market value changes by g per year, with mean €663,000 at origination and standard deviation €45,000. Dotted lines show Laila's outstanding loan L_h on a loan of €721,000, repaid as a 30-year annuity at 4%. The percentages give $P_i^{(t_0+h)}(Y < L_h)$, the probability that a resale h years after origination leaves residual debt. The scenarios are illustrative.

overbid created at origination. Since a loan that followed an overbid starts in the tail of the distribution rather than at its center, it has no margin to absorb a decline, and even a moderate fall in prices makes a resale short of the loan nearly certain. The risk of a coupled appraisal thus materializes precisely when house prices fall, which is when borrowers who can no longer pay cannot sell their way out of the loan and defaults rise (Mayer et al. 2009).

7 Conclusion

An appraisal is meant to provide an independent estimate of a dwelling's market value, rather than reproduce a transaction price that is only one realization of it. Yet all parties to a mortgage transaction benefit from an appraisal that allows the purchase to proceed, so that the appraisal tends to be anchored to the agreed price, including its idiosyncratic part, which carries no information about the expected price if the dwelling were resold almost immediately.

Using 694 Dutch residential transactions and dwelling-specific predictive price distributions, we find overwhelming evidence against independence: the Monte Carlo test rejects

Hypothesis 1 with a p-value far below 0.01, and appraisals pass on roughly four fifths of the deviation of the realized price from the predictive center. Where earlier findings by De Nederlandsche Bank (2019) and Minister voor Volkshuisvesting en Ruimtelijke Ordening (2025) documented descriptive patterns consistent with this concern, our test establishes formally that the appraisal process fails the independence requirement it is intended to satisfy. Exact copies moreover understate the agreement, since appraisals and transaction prices are found to be reported in different monetary denominations: 95.1% of appraisals are multiples of €5,000, against 47.7% of transaction prices. Once the difference is rounded to this denomination, the share of appraisals that equal the price rises from 21.3% to 41.5%.

A coupled appraisal raises the collateral risk for both borrower and lender. We prove that a valuation that copies the transaction price raises the expected residual debt upon default relative to a valuation at the expected market value, even though the copied price is an unbiased estimator of that value, so that borrowers and lenders are exposed to more risk than the reported loan-to-value ratios suggest. In a rising market this excess risk can remain largely hidden, but a cooling market, as recently reported for the Netherlands, would increase it further.

The current appraisal system therefore appears not to have withstood the incentives built into the transaction process. Strengthening appraiser independence is an obvious response, but may be difficult when appraisers work closely with parties whose interests depend on the transaction proceeding. A more direct alternative is to base collateral valuation on transparently backtested statistical models, or at least to make such a model-based valuation the core of the appraisal report. The well-calibrated model used here demonstrates the feasibility of this approach, while its predictive distributions quantify the uncertainty that a single appraisal value cannot capture. As financial forecasting increasingly moves beyond point estimates toward full predictive distributions, residential real estate valuation should not be left behind.

Disclosure statement

This paper uses the predictive model of SlimBieden to obtain dwelling specific market value distributions, which are essential to the construction of our tests. These distributions enter the analysis as a given predictive dataset; the paper neither evaluates nor promotes the model, and its results are not used for any commercial purpose. Ramon de Punder and Mathijs Dijkstra are affiliated with SlimBieden. SlimBieden had no influence on the research question, the design of the tests, or the interpretation and reporting of the results. The authors declare no conflict of interest.

The views expressed in this paper are those of the authors and do not necessarily reflect the views of SlimBieden. Nothing in this paper constitutes financial, legal or bidding advice.

References

ABN AMRO (2026). "Woningmarkt presteert beter in 2026, sterkere afkoeling verwacht in 2027". Press release on the Woningmarktmonitor, 1 October 2026.

Barnard, G. A. (1963), "Discussion of "The Spectral Analysis of Point Processes" by M. S. Bartlett", *Journal of the Royal Statistical Society: Series B*, *25*(2), 294.

Besag, J. and P. Clifford (1989), "Generalized Monte Carlo Significance Tests", *Biometrika*, *76*(4), 633–642.

Bolt, W., M. Demertzis, C. Diks, C. Hommes, and M. van der Leij (2019), "Identifying booms and busts in house prices under heterogeneous expectations", *Journal of Economic Dynamics and Control*, *103*, 234–259.

Calem, P., J. Kenney, L. Lambie-Hanson, and L. Nakamura (2021), "Appraising Home Purchase Appraisals", *Real Estate Economics*, *49*(S1), 134–168.

Dawid, A. P. (1984), "Present Position and Potential Developments: Some Personal Views.

Statistical Theory. The Prequential Approach", *Journal of the Royal Statistical Society: Series A (General)*, *147*(2), 278–292.

De Nederlandsche Bank (2019). "De kwaliteit en onafhankelijkheid van woningtaxaties". DNB Occasional Study / research note. Published 12 March 2019, accompanying Minister van Binnenlandse Zaken en Koninkrijksrelaties (2019).

Garratt, A., K. Lee, M. H. Pesaran, and Y. Shin (2003), "Forecast Uncertainties in Macroeconomic Modeling: An Application to the UK Economy", *Journal of the American Statistical Association*, *98*(464), 829–838.

Gerardi, K., K. F. Herkenhoff, L. E. Ohanian, and P. S. Willen (2018), "Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default", *The Review of Financial Studies*, *31*(3), 1098–1131.

Gneiting, T. (2011), "Making and Evaluating Point Forecasts", *Journal of the American Statistical Association*, *106*(494), 746–762.

Granger, C. W. J. and M. H. Pesaran (2000), "Economic and Statistical Measures of Forecast Accuracy", *Journal of Forecasting*, *19*(7), 537–560.

Hope, A. C. A. (1968), "A Simplified Monte Carlo Significance Test Procedure", *Journal of the Royal Statistical Society: Series B*, *30*(3), 582–598.

Langenberg, H. and W. Jonkers (2022). "De ontwikkeling van de huizenprijzen". In: Achtergrond bij de huizenprijsstijgingen vanaf 2013, Centraal Bureau voor de Statistiek.

Mayer, C., K. Pence, and S. M. Sherlund (2009), "The Rise in Mortgage Defaults", *Journal of Economic Perspectives*, *23*(1), 27–50.

Minister van Binnenlandse Zaken en Koninkrijksrelaties (2019). "Kamerbrief inzake reactie op het onderzoek van De Nederlandsche Bank naar de kwaliteit van woningtaxaties". Kamerstuk 32 847, nr. 501.

Minister van Financiën (2026a). "Besluit Gedragstoezicht financiële ondernemingen Wft".

Besluit van 12 oktober 2006, geldend vanaf 22 september 2026. BWBR0020421.

Minister van Financiën (2026b). "Tijdelijke regeling hypothecair krediet". Regeling van 12 december 2012, kenmerk FM/2012/1887 M, geldend vanaf 1 januari 2026. BWBR0032503.

Minister voor Volkshuisvesting en Ruimtelijke Ordening (2025). "Kamerbrief inzake de leennormen 2026". Kamerstuk 32 847, nr. 1387.

Ministerie van Financiën (2026). "Wijzigingsbesluit financiële markten 2026". Staatsblad van het Koninkrijk der Nederlanden, 2026, no. 223.

Nederlands Register Vastgoed Taxateurs (2026). "Taxatiewaarden en gerealiseerde koopsommen van woningen: Een onderzoek van NRVT naar de ontwikkeling van de verhouding tussen marktwaarde en transactieprijs (2019–2025)". Technical report, Nederlands Register Vastgoed Taxateurs.

Nederlands Woning Waarde Instituut (2024). "Richtlijn grote waardeverschillen". Guideline 5.2.1.7, version 2.2, Nederlands Woning Waarde Instituut. 16 September 2024.

Patton, A. J. (2009). "Copula-Based Models for Financial Time Series". In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 767–785. Springer.

Patton, A. J. (2012), "A Review of Copula Models for Economic Time Series", *Journal of Multivariate Analysis*, *110*, 4–18.

Sklar, A. (1959), "Fonctions de répartition à n dimensions et leurs marges", *Publications de l'Institut de Statistique de l'Université de Paris*, *8*, 229–231.

Vaste commissie voor Binnenlandse Zaken en Koninkrijksrelaties (2022). "Verslag van een schriftelijk overleg over de stand van zaken op het gebied van woningtaxaties". Kamerstuk 32 847, nr. 928. Vragen voorgelegd 16 september 2021; beantwoord bij brief van 7 juli 2022.

A Implementation

This section summarizes how the moments of $\widehat{F}_i$ are obtained from the predictive model, and how the two tests are carried out in practice.

A.1 Estimating the center and the spread

Writing $\widehat{q}_i$ for the quantile function of $\widehat{F}_i$, the center and spread are

$$\widehat{m}_i = \int_0^1 \widehat{q}_i(u)\,\mathrm{d}u, \qquad \widehat{\tau}_i^2 = \int_0^1 \big(\widehat{q}_i(u) - \widehat{m}_i\big)^2\,\mathrm{d}u.$$

These are computed from the predictive model's output for each dwelling, using the same working distribution throughout: it supplies $\widehat{m}_i$ and $\widehat{\tau}_i$, and it is also the distribution from which the simulated prices in Section 3 are drawn. Because the test conditions on this working distribution rather than on the true, unobserved law of Y_i, the null distribution and the standardization are internally consistent by construction. Section A.4 assesses how well the working distribution describes realized prices, which is what justifies its use in place of P_i.

A.2 Simulation and resampling

Simulated prices are generated by inverse-transform sampling: for each replication a uniform draw on $[0, 1]$ is passed through the interpolated $\widehat{q}_i$. The reported results use $B =$ 100,000 replications, computed in blocks to limit memory use, from a fixed random seed so that they are exactly reproducible.

The direction test of Section 3.5 uses $\bar{\Delta} = n^{-1}\sum_i \Delta_i$, with $\Delta_i = \Delta_i(0)$ denoting unrounded differences, with the usual t statistic. One feature of the data deserves comment: in a substantial share of dwellings the appraisal equals the realized price exactly, contributing a difference of precisely zero. These observations contribute nothing to the sum of the

differences, but they place enough mass at the origin for the median difference to be zero while the mean is not, so the median is uninformative about the displacement here.

Sampling uncertainty in the reported statistics is summarized by a non-parametric bootstrap. Complete dwelling-level observations are resampled with replacement, so that A_i, Y_i, m_i and τ_i stay together, and the statistics are recomputed for each bootstrap sample; the reported 95% intervals are the 2.5th and 97.5th percentiles of 20,000 replications. The bootstrap is used for confidence intervals only. The p-value for the coupling test comes from the randomization procedure, and that for the direction test from the t statistic.

A.3 Calibration and scale

The randomization test replaces the unknown conditional price distributions P_i by the fitted distributions $\widehat{F}_i$. Inference for Hypothesis 1 therefore depends on whether the latter reproduce the relevant features of the transaction-price distribution. For this purpose we assess their first two moments separately. The location of $\widehat{m}_i$ determines whether the predictive center is a credible approximation to the expected market value, while the scale $\widehat{\tau}_i$ determines the standardization of the coupling statistic S_A.

We assess location on a relative rather than an absolute scale. A given proportional prediction error corresponds to a larger euro error for a more expensive dwelling, so an unscaled statistic would place greater weight on the upper end of the price distribution. Define the first-moment calibration statistic, based on the relative error, as

$$D_P = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i - \widehat{m}_i}{\widehat{m}_i}. \tag{7}$$

A correctly located predictive center satisfies $\mathbb{E}\left[\frac{Y_i - \widehat{m}_i}{\widehat{m}_i}\right] = 0$. Since no direction of model error is predicted, inference is two-sided.

Dispersion is assessed through

$$S_P = \frac{1}{n}\sum_{i=1}^{n}\frac{(Y_i - \widehat{m}_i)^2}{\widehat{\tau}_i^2}. \tag{8}$$

If the predictive second moments are correctly specified, S_P has expectation one. Values above one indicate that realized prices are more dispersed around their predicted centers than the fitted distributions imply; values below one indicate less dispersion.

A.4 Calibration results

The interpretation of $\widehat{m}_i$ as a feasible estimate of expected market value requires the predictive centers to be correctly located. This is particularly relevant to the alternative valuation discussed in Section 1: unlike the appraisal, $\widehat{m}_i$ is constructed without conditioning on the realized transaction price. We assess its first-moment calibration using D_P in (7) and test

$$H_0 : \mathbb{E}\left[\frac{Y_i - \widehat{m}_i}{\widehat{m}_i}\right] = 0 \qquad \text{against} \qquad H_1 : \mathbb{E}\left[\frac{Y_i - \widehat{m}_i}{\widehat{m}_i}\right] \neq 0. \tag{9}$$

Across the $n = 694$ dwellings, the mean relative prediction error is $D_P = -0.326\%$, with a standard error of 0.189%. The resulting t statistic is -1.73, with a two-sided p-value of 0.085. A non-parametric bootstrap gives a 95% confidence interval of (-0.688%, 0.048%). We therefore do not reject Hypothesis 9 at the 5% level. The data thus provide no statistically significant evidence of systematic first-moment miscalibration.

Correct calibration of the predictive dispersion is separately important for Hypothesis 1. The estimated second-moment statistic is $S_P = 1.0685$, with a non-parametric bootstrap 95% confidence interval of (0.9625, 1.1861). The calibrated reference value of one lies inside this interval. We therefore find no evidence that realized transaction prices are systematically more or less dispersed around their predicted centers than the working distributions imply. In particular, the data do not support the possibility that excessive predictive dis-

persion is mechanically responsible for the small value of S_A observed in Section 5.

A.5 Predictive performance

Calibration concerns the moments of the predictive distributions. We evaluate the accuracy of m_i and define the absolute relative prediction error as

$$R_i = \left| \frac{Y_i - \widehat{m}_i}{Y_i} \right|. \tag{10}$$

Using the realized transaction price as the denominator expresses the prediction error as a fraction of the value ultimately realized by the dwelling. The median of R_i across the full sample is 3.29%.

The predictive errors can also be related to the statutory accuracy requirements for model-based residential valuations. Under Article 115 of the Besluit Gedragstoezicht financiële ondernemingen Wft, a model-based valuation that is reviewed and approved by an expert and independent appraiser may support a mortgage of up to 90% of the model-implied value when the underlying statistical model deviates by less than 10% from the eventual transaction value in at least 90% of valuations. A second category permits a mortgage of up to 70% of the model-implied value when at least 85% of valuations deviate by less than 15% (Minister van Financiën 2026b).

To compare the model with these numerical accuracy thresholds, define

$$C(\delta) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1} \left\{ \left| \frac{Y_i - \widehat{m}_i}{Y_i} \right| < \delta \right\}, \tag{11}$$

where δ is the permitted relative deviation. We obtain $C(0.10) = 97.41\%$ and $C(0.15) = 99.57\%$. The model therefore exceeds, within our evaluation sample, the numerical accuracy requirement associated with the stricter 90%-of-model-value category by a substantial margin. The sample calculation is not itself the annual assurance assessment required by

Article 115 and should therefore be read as a comparison with the statutory accuracy threshold rather than as a regulatory certification of the model.

B Two forms of rounding

Throughout, $\kappa > 0$ and $[z]_\kappa \coloneqq \kappa \operatorname{round}(z/\kappa)$, where round rounds to the nearest integer with halfway cases to the nearest even integer. The only property of $[\cdot]_\kappa$ used below is that it rounds to a nearest multiple of κ, that is, $[z]_\kappa \in \mathcal{M}_\kappa \coloneqq \{j\kappa : j \in \mathbb{Z}\}$, the set of multiples of κ, and

$$\left|z - [z]_\kappa\right| \le \frac{\kappa}{2} \qquad \text{for all } z \in \mathbb{R}, \tag{12}$$

with strict inequality unless z/κ lies exactly halfway between two integers. The results therefore hold for any tie-breaking convention. For $\kappa = 0$, both forms equal $a_i - y_i$ and the results below are trivial.

Proposition B.1. *For all $a_i, y_i \in \mathbb{R}$ and $\kappa > 0$,*

$$\left|\Delta(a_i, y_i; \kappa) - \tilde{\Delta}(a_i, y_i; \kappa)\right| \le \kappa,$$

and the bound is attained. Moreover, if a_i is a multiple of κ and $(a_i - y_i)/\kappa$ does not lie halfway between two integers, then $\Delta(a_i, y_i; \kappa) = \tilde{\Delta}(a_i, y_i; \kappa)$.

Proof. Write $d_i = a_i - y_i$. By (12) applied to a_i and y_i separately,

$$\left|\tilde{\Delta}(a_i, y_i; \kappa) - d_i\right| = \left|\left([a_i]_\kappa - a_i\right) - \left([y_i]_\kappa - y_i\right)\right| \le \frac{\kappa}{2} + \frac{\kappa}{2} = \kappa,$$

and by (12) applied to d_i,

$$\left|\Delta(a_i, y_i; \kappa) - d_i\right| \le \frac{\kappa}{2}.$$

By the triangle inequality,

$$\left|\Delta(a_i, y_i; \kappa) - \tilde{\Delta}(a_i, y_i; \kappa)\right| \leq \frac{3\kappa}{2}.$$

Both $\Delta(a_i, y_i; \kappa)$ and $\tilde{\Delta}(a_i, y_i; \kappa)$ are multiples of κ, hence so is their difference. The only multiples of κ with absolute value at most $3\kappa/2$ are $-\kappa$, 0 and κ, which proves the bound.

For the second statement, let $a_i \in \kappa\mathbb{Z}$, so that $[a_i]_\kappa = a_i$ and $\tilde{\Delta}(a_i, y_i; \kappa) = a_i - [y_i]_\kappa$. This is a multiple of κ, and

$$\left|\tilde{\Delta}(a_i, y_i; \kappa) - d_i\right| = \left|y_i - [y_i]_\kappa\right| \leq \frac{\kappa}{2},$$

so that $\tilde{\Delta}(a_i, y_i; \kappa)$ is a nearest multiple of κ to d_i, as is $\Delta(a_i, y_i; \kappa)$. If d_i/κ does not lie halfway between two integers, the nearest multiple of κ to d_i is unique, and the two coincide.

To see that the bound is attained, let $\kappa = €5{,}000$, $a_i = €552{,}600$ and $y_i = €552{,}400$. Then $\Delta(a_i, y_i; \kappa) = [200]_\kappa = 0$, whereas $[a_i]_\kappa = €555{,}000$ and $[y_i]_\kappa = €550{,}000$, so that $\tilde{\Delta}(a_i, y_i; \kappa) = €5{,}000 = \kappa$. □

Neither form of rounding is stronger than the other, in the sense that neither $\Delta(a_i, y_i; \kappa) = 0$ implies $\tilde{\Delta}(a_i, y_i; \kappa) = 0$, nor the converse. The example in the proof gives $\Delta(a_i, y_i; \kappa) = 0$ with $\tilde{\Delta}(a_i, y_i; \kappa) \neq 0$, because the two amounts lie on opposite sides of the halfway point €552,500. Conversely, the same appraisal with $y_i = €557{,}400$ gives $[a_i]_\kappa = [y_i]_\kappa = €555{,}000$, so that $\tilde{\Delta}(a_i, y_i; \kappa) = 0$, whereas $\Delta(a_i, y_i; \kappa) = [-4{,}800]_\kappa = -€5{,}000$. Both examples involve an appraisal that is not a multiple of κ; by Proposition B.1, the two forms can otherwise differ only at halfway cases.

B.1 Summary statistics for both rounding methods

Table 6 provides the summary statistics of both rounding methods introduced in 3.3.

Table 6: Agreement under the two forms of rounding

	$\kappa = €2{,}500$	$\kappa = €5{,}000$	$\kappa = €10{,}000$
Panel A: Rounding the difference, $\Delta_i(\kappa) = [a_i - y_i]_\kappa$			
$\Delta_i(\kappa) = 0$	31.1%	41.5%	60.8%
$0 \leq \Delta_i(\kappa) \leq \kappa$	40.6%	56.3%	71.0%
$-\kappa \leq \Delta_i(\kappa) \leq 0$	38.2%	50.6%	69.0%
Panel B: Rounding both amounts, $\tilde{\Delta}_i(\kappa) = [a_i]_\kappa - [y_i]_\kappa$			
$\tilde{\Delta}_i(\kappa) = 0$	31.0%	39.9%	51.0%
$0 \leq \tilde{\Delta}_i(\kappa) \leq \kappa$	40.6%	56.3%	67.0%
$-\kappa \leq \tilde{\Delta}_i(\kappa) \leq 0$	38.0%	49.3%	63.1%

NOTE: Shares of all $n = 694$ dwellings. $[z]_\kappa = \kappa\,\mathrm{round}(z/\kappa)$ rounds to the nearest multiple of κ, with halfway cases to the nearest even multiple. The second and third rows of each panel count the appraisals that equal the price up to rounding or lie one step of κ above, respectively below, it. Panel A repeats the shares of Table 2. The two forms differ by at most one step, $|\Delta_i(\kappa) - \tilde{\Delta}_i(\kappa)| \leq \kappa$ (Proposition B.1), and coincide whenever a_i is a multiple of κ, except at halfway cases.

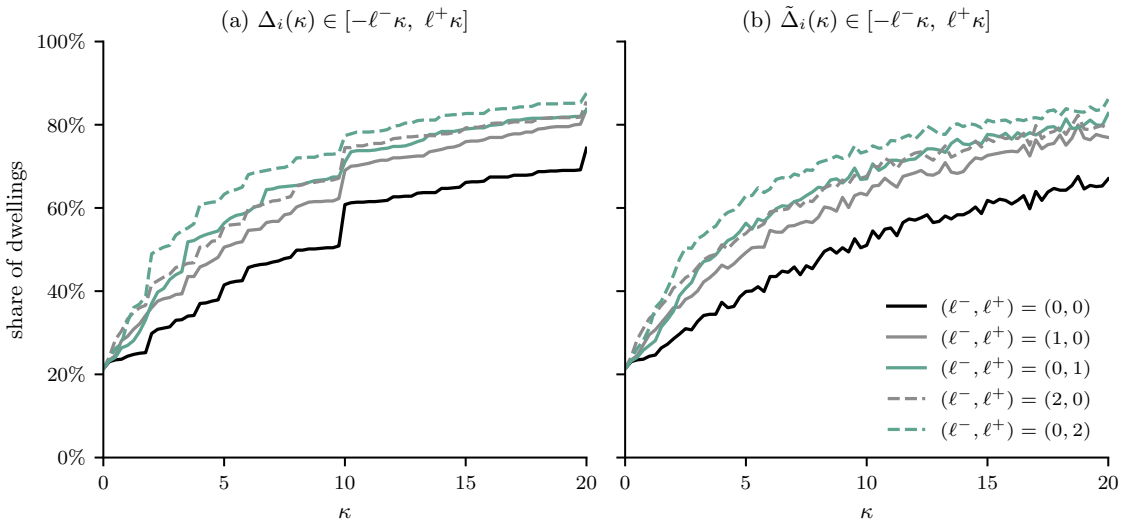
Figure 6 shows these shares for a fine grid of κ. For nonnegative integers ℓ^- and ℓ^+, it plots the share of dwellings whose rounded difference lies at most ℓ^- steps of κ below and at most ℓ^+ steps above the transaction price,

$$W(\kappa; \ell^-, \ell^+) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left\{-\ell^- \kappa \leq \Delta_i(\kappa) \leq \ell^+ \kappa\right\},$$

and the same share with $\tilde{\Delta}_i(\kappa)$ in place of $\Delta_i(\kappa)$. The window $(\ell^-, \ell^+) = (0, 0)$ gives the share that equals the price up to rounding, and the windows $(0, 1)$ and $(1, 0)$ give the shares at most one step above, respectively below, the price, as in the last two rows of Table 2. Comparing $(0, \ell)$ with $(\ell, 0)$ for $\ell = 1, 2$ thus contrasts appraisals above the price with appraisals below it at the same distance. For both forms of rounding, the windows that extend above the price lie above those that extend below it at every κ, so that the upward asymmetry of Section 4.2 does not depend on the choice of rounding step.

Figure 7 shows the appraisal–price differences that are exactly a multiple of €5,000, by step size. These are the 148 exact copies at $k = 0$ and the 174 departures by a multiple of €5,000 in Panel B of Table 3. At every step size, more appraisals lie above the transaction

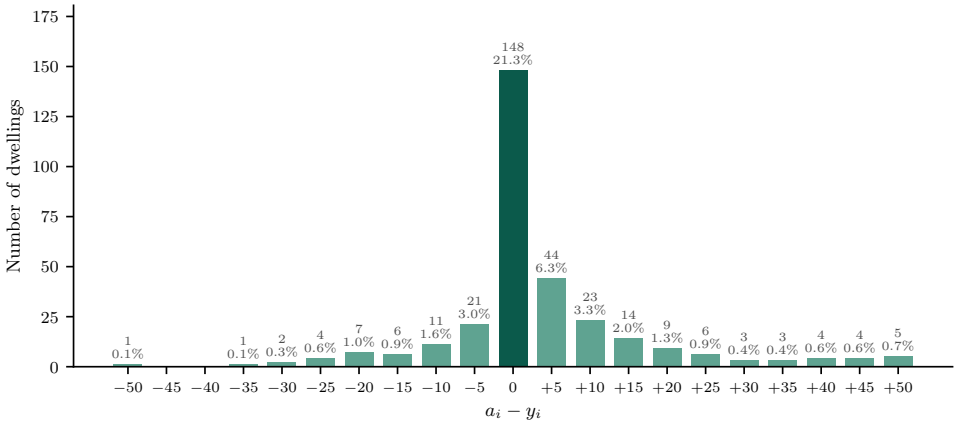
Figure 6: Agreement between appraisals and transaction prices across rounding steps



NOTE: Share of dwellings whose rounded appraisal–price difference lies within $[-\ell^-\kappa, \ell^+\kappa]$, for rounding steps κ up to €20,000, in steps of €250. Panel (a) rounds the difference, $\Delta_i(\kappa) = [a_i - y_i]_\kappa$; Panel (b) rounds both amounts, $\tilde{\Delta}_i(\kappa) = [a_i]_\kappa - [y_i]_\kappa$. The black line is the window $(\ell^-, \ell^+) = (0, 0)$; green lines extend the window above the price and gray lines below it, solid by one step of κ and dashed by two. At $\kappa = 0$, every line equals the share of exact copies, 21.3%. Table 2 reports the shares for $\kappa =$ €2,500, €5,000 and €10,000. The horizontal axis shows κ in thousands of euros; shares are of all 694 dwellings.

price than below it: 44 against 21 at €5,000, 23 against 11 at €10,000, and 53 against 22 at €15,000 or more.

Figure 7: Appraisal–price differences that are exact multiples of €5,000



NOTE: Number of dwellings for which the appraisal–price difference $a_i - y_i$ equals exactly $k\times$ €5,000 for an integer k, including the exact copies at $k = 0$. Not shown are the 372 differences that are not a multiple of €5,000: 188 dwellings (27.1%) lie below the transaction price and 184 (26.5%) above. Also not shown is 1 dwelling with $a_i - y_i < -52{,}500$ and 5 with $a_i - y_i > 52{,}500$. The horizontal axis is in thousands of euros; percentages are shares of all 694 dwellings.

C Details on Section 5.2

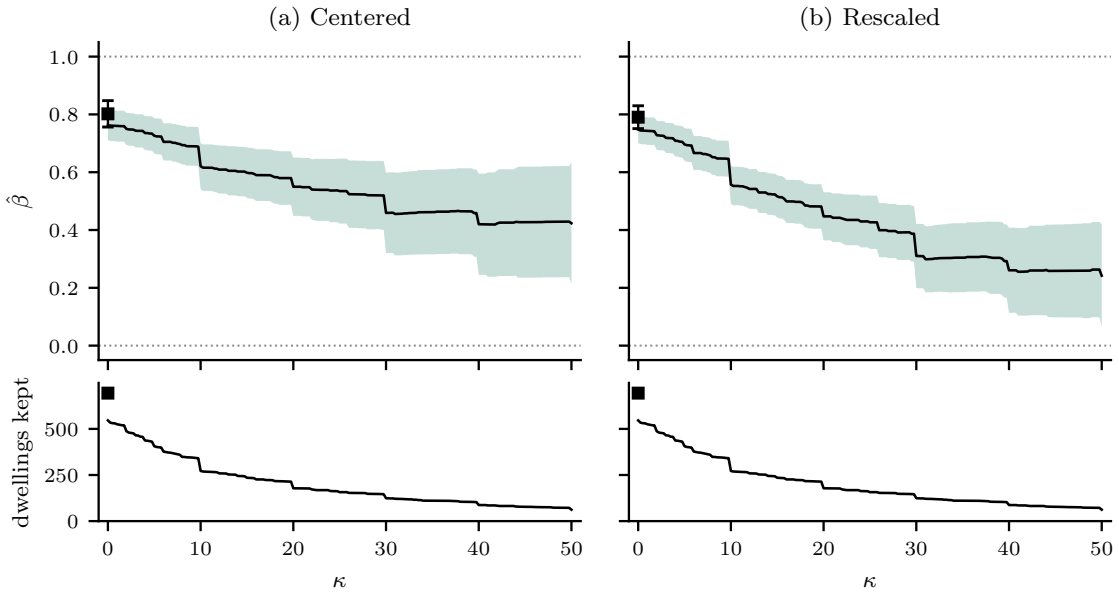
Table 7 reports two checks on the estimates of Table 5. The first rounds both amounts to multiples of κ before centering, (5), so that appraisals that copy the price up to rounding are treated as exact copies; the slope is essentially unchanged for every κ. The second excludes the dwellings with $\Delta_i(\kappa) = 0$, that is, $|a_i - y_i| \leq \kappa/2$. Figure 8 shows this exclusion for a fine grid of κ up to €50,000.

Table 7: Anchoring regressions under rounding and exclusion

Panel A: Centered (4)	n	$\hat{\alpha}$	$\hat{\beta}$	R^2
All dwellings	694	814 (538)	0.802 (0.023)	0.778
Both amounts rounded to multiples of κ				
$\kappa =$ €2,500	694	870 (538)	0.803 (0.023)	0.779
$\kappa =$ €5,000	694	970 (540)	0.803 (0.023)	0.778
$\kappa =$ €10,000	694	829 (555)	0.795 (0.024)	0.764
Excluding $\Delta_i(\kappa) = 0$				
Exact copies, $a_i = y_i$	546	1,388 (659)	0.763 (0.027)	0.739
$\kappa =$ €2,500	478	1,670 (745)	0.748 (0.029)	0.722
$\kappa =$ €5,000	406	2,034 (863)	0.726 (0.031)	0.699
$\kappa =$ €10,000	272	2,062 (1,214)	0.621 (0.041)	0.587
Panel B: Rescaled (6)	n	$\hat{\alpha}$	$\hat{\beta}$	R^2
All dwellings	694	0.024 (0.017)	0.790 (0.020)	0.761
Both amounts rounded to multiples of κ				
$\kappa =$ €2,500	694	0.026 (0.017)	0.792 (0.020)	0.762
$\kappa =$ €5,000	694	0.030 (0.018)	0.791 (0.020)	0.759
$\kappa =$ €10,000	694	0.024 (0.018)	0.778 (0.021)	0.743
Excluding $\Delta_i(\kappa) = 0$				
Exact copies, $a_i = y_i$	546	0.046 (0.021)	0.747 (0.024)	0.717
$\kappa =$ €2,500	478	0.054 (0.024)	0.726 (0.026)	0.694
$\kappa =$ €5,000	406	0.068 (0.028)	0.696 (0.029)	0.660
$\kappa =$ €10,000	272	0.065 (0.038)	0.557 (0.034)	0.508

NOTE: Heteroskedasticity-robust (HC1) standard errors in parentheses; $\hat{\alpha}$ is in euros in Panel A and in units of $\widehat{\tau}_i$ in Panel B. The rounded rows estimate (5) and its rescaled counterpart, with both amounts rounded to multiples of κ before centering. The excluded rows estimate (4) and (6) on the dwellings with $\Delta_i(\kappa) \neq 0$, that is, $|a_i - y_i| > \kappa/2$; the first of these excludes the exact copies only. In every row, $\beta = 0$ and $\beta = 1$ are rejected with two-sided $p < 0.001$.

Figure 8: Anchoring slope excluding appraisals close to the price



NOTE: Estimates of the centered regression (4) in Panel (a) and the rescaled regression (6) in Panel (b) on the dwellings with $\Delta_i(\kappa) \neq 0$, that is, $|a_i - y_i| > \kappa/2$, with 95% heteroskedasticity-robust confidence bands; at $\kappa = 0$ the line starts from the sample without exact copies. The black square is the full sample, with its 95% confidence interval. The lower panels show the number of dwellings kept. The horizontal axes show κ in thousands of euros. Table 7 reports the estimates for selected values of κ.

D Validity under a biased appraiser

Section 2 defines $\xi_i = A_i - \mu_i$ without imposing that it is conditionally centered; whether it is, is the content of Hypothesis 2, whose null is rejected in Section 5. This appendix shows that the coupling test does not depend on that restriction, so the two findings do not conflict.

Write

$$b_i = \mathbb{E}[\xi_i], \qquad \eta_i = \xi_i - b_i, \qquad \mathbb{E}[\eta_i] = 0,$$

so that $A_i = \mu_i + b_i + \eta_i$. The displacement b_i may vary freely with the characteristics.

Proposition D.1. *If $\eta_i \perp \varepsilon_i \mid X_i = x_i$, then*

$$\mathbb{E}\big[\Delta_i^2\big] = b_i^2 + \varsigma_i^2 + \sigma_i^2 \;\geq\; \sigma_i^2,$$

with equality if and only if $\xi_i = 0$ almost surely.

Proof. Since $\Delta_i = b_i + \eta_i - \varepsilon_i$ and $\mathbb{E}[\eta_i] = \mathbb{E}[\varepsilon_i] = 0$, the cross terms involving b_i vanish, and the remaining cross term vanishes by conditional independence. Hence $\mathbb{E}[\Delta_i^2] = b_i^2 + \mathbb{E}[\eta_i^2] + \sigma_i^2$, and $\mathbb{E}[\eta_i^2] = \varsigma_i^2$. The first two terms are non-negative, which gives the bound; equality requires both to vanish. □

Displacement therefore enters through $b_i^2 \geq 0$: a biased appraiser produces a larger appraisal–price second moment, never a smaller one, so bias cannot generate the violation the test looks for.

The randomization is unaffected for a separate reason. Under Hypothesis 1 the conditional law of Y_i given $A_i = a_i$ and $X_i = x_i$ is P_i whatever the value of b_i, so the null distribution of the price does not depend on where the appraisal sits. Conditioning on the observed a_i holds any displacement fixed across replications, where it enters only through the benchmark (3). The coupling test therefore requires conditional independence together with a working predictive law for transaction prices, and not a centered appraiser.

E Proofs of and further details on Section 6

E.1 Proof of Proposition 1

Proof. For clarity, we index each expectation in this proof by the law over which it integrates. Each price has law P, so that by independence $(Y_{\text{old}}, Y_{\text{new}})$ has law $P \otimes P$, and K_y denotes the conditional law of η given $Y_{\text{old}} = y$. Define

$$h(y) \coloneqq \mathbb{E}_P\big[(\ell y - Y_{\text{new}})^+\big] = \int (\ell y - t)^+ \, \mathrm{d}P(t),$$

which is finite since $(\ell y - t)^+ \leq \ell|y| + |t|$. For every t, the map $y \mapsto (\ell y - t)^+$ is convex, and integration preserves convexity, so h is convex. Integrating first over Y_{new}, which is

independent of (Y_{old}, η), Fubini's theorem gives

$$E_{P\otimes P}\big[(\ell Y_{\text{old}} - Y_{\text{new}})^+\big] = \mathbb{E}_P\big[h(Y_{\text{old}})\big],$$

and additionally

$$E\big[(\ell(Y_{\text{old}} + \eta) - Y_{\text{new}})^+\big] = \mathbb{E}_P\Big[\mathbb{E}_{K_{Y_{\text{old}}}}\big[h(Y_{\text{old}} + \eta)\big]\Big],$$

where the expectation on the left of the second identity integrates over Y_{old}, η and Y_{new} jointly, and on the right the outer expectation integrates over Y_{old} and the inner over η given Y_{old}. For the first inequality, Jensen's inequality gives $\mathbb{E}_P[h(Y_{\text{old}})] \geq h\big(\mathbb{E}_P[Y_{\text{old}}]\big) = h(\mu_Y)$. For the second, Jensen's inequality under K_y gives, for every y,

$$\mathbb{E}_{K_y}\big[h(y + \eta)\big] \geq h\big(y + \mathbb{E}_{K_y}[\eta]\big) = h(y),$$

since $\mathbb{E}_{K_y}[\eta] = 0$. Integrating over $Y_{\text{old}} \sim P$ gives $\mathbb{E}\big[(\ell(Y_{\text{old}} + \eta) - Y_{\text{new}})^+\big] \geq \mathbb{E}_P[h(Y_{\text{old}})]$, which is the second inequality. □

E.2 Proof of Proposition 2

Proof. As in the proof of Proposition 1, we index each expectation by the law over which it integrates; by independence, (A, Y_{new}) has law $Q \otimes P$. Define

$$h(y) \coloneqq \mathbb{E}_P\big[(\ell y - Y_{\text{new}})^+\big] = \int (\ell y - t)^+ \, dP(t),$$

which is finite since $(\ell y - t)^+ \leq \ell|y| + |t|$. For every t, the map $y \mapsto (\ell y - t)^+$ is convex and non-decreasing, and integration preserves both properties, so h is convex and non-

decreasing. Integrating first over Y_{new}, Fubini's theorem gives

$$\mathbb{E}_{Q\otimes P}\big[(\ell A - Y_{\text{new}})^+\big] = \mathbb{E}_Q\big[h(A)\big].$$

Jensen's inequality gives $\mathbb{E}_Q[h(A)] \geq h\big(\mathbb{E}_Q[A]\big)$, and since h is non-decreasing and $\mathbb{E}_Q[A] \geq \mu_Y$, also $h\big(\mathbb{E}_Q[A]\big) \geq h(\mu_Y) = \mathbb{E}_P[(\ell\mu_Y - Y_{\text{new}})^+]$. □

E.3 Time effects

Let $h = t_1 - t_0$ denote the number of years between origination and default. The transaction price at origination is $Y_{\text{old}} \sim P_i^{(t_0)} = N(\mu, \sigma^2)$, and the resale price at default is $Y_{\text{new}} \sim P_i^{(t_1)} = N(\mu_h, \sigma^2)$, independent of Y_{old}, with $\mu_h = (1+g)^h\mu$ for a constant annual change g in market value. The loan L_0 is repaid as a 30-year annuity at interest rate r, so that the outstanding balance after h years is $L_h = \rho_h L_0$, with

$$\rho_h = \frac{(1+r)^{30} - (1+r)^h}{(1+r)^{30} - 1}.$$

If $L_h - Y_{\text{new}} = \delta + sZ$ with $Z \sim N(0,1)$, where δ is the expected gap between loan and resale price and s the uncertainty about it, then

$$\mathbb{E}\big[(L_h - Y_{\text{new}})^+\big] = \delta\,\mathbb{P}(Z > -\delta/s) + s\,\mathbb{E}\big[Z\,\mathbb{1}\{Z > -\delta/s\}\big] = \delta\,\Phi(\delta/s) + s\,\phi(\delta/s),$$

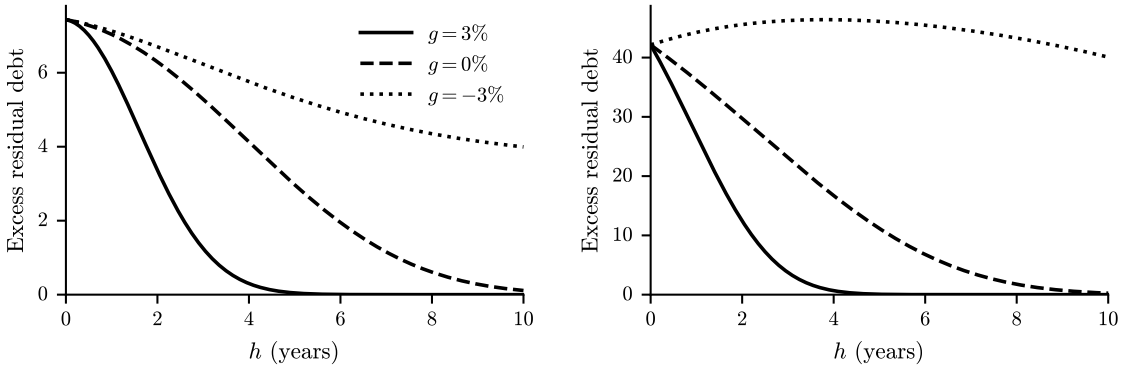
where Φ and ϕ denote the standard normal distribution and density functions. The expected residual debt increases in both δ and s.

We compare two valuations with a loan based on the model valuation, $L_0 = \mu$, for which $\delta = \rho_h\mu - \mu_h$ and $s^2 = \sigma^2$. The first is the strategy of copying the transaction price, $L_0 = Y_{\text{old}}$, evaluated over all transaction prices it may meet. The expectation is then taken over both Y_{old} and Y_{new}, so that δ is unchanged while $s^2 = (1+\rho_h^2)\sigma^2$: copying adds the

deviation of the transaction price to the loan, and this deviation is repaid only at the rate ρ_h. By Proposition 1, this strategy raises the expected residual debt at origination, and the same argument applies at every horizon. The second is one particular outcome of such a strategy, Laila's loan in Example 2. Her price is known, so that $L_0 = €721{,}000$ is fixed and the expectation is taken over Y_{new} only. Then $s^2 = \sigma^2$, while her bid premium raises the expected gap to $\delta = \rho_h L_0 - \mu_h$.

Figure 9 shows the excess expected residual debt over the model valuation, $\mathbb{E}[(L_h - Y_{\text{new}})^+] - \mathbb{E}[(\rho_h \mu - Y_{\text{new}})^+]$, for $\mu = €663{,}000$ and $\sigma = €45{,}000$ as in Example 2, $r = 4\%$, and $g \in \{-3\%, 0\%, 3\%\}$, without sale costs or a forced-sale discount. For the copying strategy (Panel (a)), the excess is approximately €7,400 at origination; it has almost disappeared after four years if prices rise, but approximately €4,000 remains after ten years if they fall. For Laila (Panel (b)), the excess of approximately €42,000 at origination vanishes within four years if prices rise, but grows to approximately €46,000 if they fall.

Figure 9: Excess expected residual debt over time



(a) copied transaction price

(b) Laila's loan

NOTE: Excess expected residual debt relative to a loan based on the model valuation, $\mathbb{E}[(L_h - Y_{\text{new}})^+] - \mathbb{E}[(\rho_h \mu - Y_{\text{new}})^+]$, by the number of years h until default, for annual changes in market value of $g = 3\%$, 0% and -3%. Panel (a) shows the strategy of copying the transaction price, averaged over all transaction prices; Panel (b) shows Laila's loan of €721,000 in Example 2. The market value at origination is normal with mean €663,000 and standard deviation €45,000, and the loan is repaid as a 30-year annuity at 4%. The vertical axes are in thousands of euros. The scenarios are illustrative.

E.4 Accuracy of the predictive mean

This appendix illustrates the variance reduction from replacing the copied price $\hat{\mu}_1 = Y_{\text{old}}$ by the predictive mean $\hat{\mu}_3 = m_Y$ of Section 6, using a linear regression as the predictive model. Suppose, for illustration, that market values follow $Y = X'\beta + u$ with $\mathbb{E}[u \mid X] = 0$ and $\mathbb{V}(u \mid X = x) = \sigma^2$, so that $\mu_Y = x'\beta$, and that Y_{old} is a draw from this conditional law. Let $\hat{m} \coloneqq x'\hat{\beta}$, where $\hat{\beta}$ is the least-squares estimator fitted on n_0 independent transactions other than the dwelling itself, with design matrix $\mathbf{X}$ of full column rank. Both estimators are unbiased for μ_Y, but

$$\mathbb{V}(\hat{\mu}_1) = \sigma^2, \qquad \mathbb{V}(\hat{m}) = x'\mathbb{V}(\hat{\beta})\,x = \sigma^2\, v_{n_0}, \qquad v_{n_0} \coloneqq x'(\mathbf{X}'\mathbf{X})^{-1}x.$$

The variance of the copied price is fixed at the full market variance, whereas v_{n_0} is the leverage of the characteristics x in the design, of order k/n_0 for a dwelling with typical characteristics and k regressors. The condition $v_{n_0} < 1$ therefore holds for any dwelling within the range of the fitting sample once n_0 is moderately large, and the predictive mean then dominates the copied price in mean squared error.

E.5 Details on regression misspecification

The regressions (4) and (6) centre the appraisal and the transaction price on the model centre $\widehat{m}_i$ rather than on the conditional mean $\mu_i = \mathbb{E}[Y_i]$, which is unknown. This appendix shows how the error in $\widehat{m}_i$ affects the slope, and how large it would have to be to explain the slopes we find.

We write $\varepsilon_i = y_i - \mu_i$ for the deviation of the transaction price from its conditional mean, $\xi_i = a_i - \mu_i$ for the deviation of the appraisal, and $d_i = \widehat{m}_i - \mu_i$ for the error of the model, and assume that d_i is uncorrelated with ε_i and with ξ_i. The first holds because the model is estimated on transactions outside the sample, so that, given the training data, d_i

depends on X_i only, while $\mathbb{E}[\varepsilon_i \mid X_i] = 0$. The second is an assumption: it requires that appraisal errors and errors in the predictive mean are uncorrelated across dwellings. For (6) the same assumptions are needed for the deviations divided by $\widehat{\tau}_i$; the first again holds automatically, since $\widehat{\tau}_i$ depends on X_i only.

The regressand of (4) is $a_i - \widehat{m}_i = \xi_i - d_i$ and the regressor is $y_i - \widehat{m}_i = \varepsilon_i - d_i$, so the population slope is

$$\beta = \frac{\mathbb{C}\mathrm{ov}(\xi_i, \varepsilon_i) + \mathbb{V}(d_i)}{\mathbb{V}(\varepsilon_i) + \mathbb{V}(d_i)} = (1 - s)\,\tilde{\beta} + s \cdot 1, \qquad s = \frac{\mathbb{V}(d_i)}{\mathbb{V}(\varepsilon_i) + \mathbb{V}(d_i)}, \tag{13}$$

where $\tilde{\beta} = \mathbb{C}\mathrm{ov}(\xi_i, \varepsilon_i)/\mathbb{V}(\varepsilon_i)$ is the slope we would obtain with μ_i known and s is the share of the variance of $y_i - \widehat{m}_i$ that is due to model error. Without the assumption in Remark 1, information shared by buyers and appraisers but not contained in x_i acts as additional model error. Here we make explicit the distributions that act on the operators used. The variances and covariances in (13) are taken under the joint law of $(X_i, \varepsilon_i, \xi_i)$ across dwellings, with the training sample held fixed. Hence $\mathbb{V}(d_i)$ measures how the model error varies across dwellings, and $\mathbb{V}(\varepsilon_i)$ the average conditional variance of the transaction price.

The model error thus pulls the slope towards one. In particular, if appraisals were uncorrelated with the price deviation, $\tilde{\beta} = 0$, the slope would equal the share of model error, $\beta = s$. Solving (13) for the slope of interest gives

$$\tilde{\beta} = \frac{\beta - s}{1 - s}. \tag{14}$$

Without the second assumption, the right-hand side of (13) has the additional term $-\mathbb{C}\mathrm{ov}(\xi_i, d_i)/\mathbb{V}(y_i - \widehat{m}_i)$. If appraisals and the model share blind spots, so that both deviate in the same direction, this covariance is positive and the slope of interest exceeds the value implied by (14). The analysis below is conditional on the absence of this term.

Since the variance of $y_i - \widehat{m}_i$ contains both the variance of the price deviation and that of the model error, the share s is not identified from transaction prices and predictive means alone. We therefore perform a sensitivity analysis to assess its impact on the estimated slope. Table 8 reports the slope of interest implied by (14) for a range of hypothetical values of s, on the full sample. The second column expresses the model error in units of the price deviation, $\sqrt{\mathbb{V}(d_i)/\mathbb{V}(\varepsilon_i)} = \sqrt{s/(1-s)}$. At $s = 0$ the slope of interest coincides with the estimated slope.

To judge whether these shares are large, recall what the model centre estimates. The model centre $\widehat{m}_i$ estimates the conditional mean μ_i by pooling the information in n_0 comparable transactions. The transaction price, used as an estimator of μ_i, does not pool at all: its error is the price deviation ε_i itself, so that its variance is the full $\mathbb{V}(\varepsilon_i)$, however many other transactions are available. A ratio of one, $s = 0.5$, therefore means that the model estimates the expected price of a dwelling no more accurately than copying the price of a single transaction would. Pooling reduces the error far below this level. In the OLS model of Appendix E.4, the variance of d_i is of order $\mathbb{V}(\varepsilon_i)\, k/n_0$, where k is the number of parameters and n_0 the number of training transactions. The predictive model in this paper is trained on transactions that precede the 694 transactions in our sample, and n_0 is many times larger than k, so that the variance of d_i due to estimation error is likely to be orders of magnitude smaller than that of ε_i, corresponding to values of s close to zero. The shares in the lower half of Table 8 cannot arise from estimation error; they require that the model systematically fails to capture how the conditional mean varies with the characteristics it uses, to the point of being less accurate than the transaction price it is meant to improve upon.

Even such errors leave much of the slope intact. On the full sample, a model error as large as the price deviation, $s = 0.5$, still leaves an implied slope of 0.60 for (4) and 0.58 for (6). Attributing half of the estimated slope of (4) to model error requires $s = 0.67$, a

Table 8: Anchoring slope for a given share of model error

		Slope of interest $\tilde{\beta}$	
s	$\sqrt{s/(1-s)}$	Centered (4)	Rescaled (6)
0	0.00	0.802	0.790
0.2	0.50	0.753	0.738
0.4	0.82	0.670	0.650
0.5	1.00	0.604	0.581
0.6	1.22	0.505	0.476
0.7	1.53	0.340	0.301
0.8	2.00	0.010	−0.049

NOTE: s is the share of the variance of $y_i - \widehat{m}_i$ due to model error, and the second column is the corresponding ratio of the standard deviation of the model error to that of the price deviation, $\sqrt{s/(1-s)}$. The last two columns give the slope of interest $\tilde{\beta} = (\hat{\beta} - s)/(1-s)$ implied by (13), with $\hat{\beta}$ the estimated slope on all dwellings from Table 5. For the rescaled regression, s and the ratio refer to the deviations divided by $\widehat{\tau}_i$.

model error with 1.4 times the standard deviation of the price deviation. Reconciling the estimated slope of 0.80 with a slope of interest of zero requires $s = 0.80$: model error must then account for 80 percent of the variance of the regressor, so that its standard deviation is twice that of the price deviation, and the model would be twice as inaccurate as the copied transaction price.

These calculations quantify the sensitivity of the estimated slope to errors in the predictive mean under the maintained orthogonality assumption. They do not identify the actual share s, and therefore cannot rule out an explanation based on misspecification of the predictive mean. They do show how large such misspecification must be to reconcile the observed slope with the absence of linear dependence between the appraisal and the idiosyncratic price deviation.