

The Confidence Band

Automated valuation is being sold badly across our industry. The problem is not inaccuracy. It is a single point value published with no evidence gate, no disclosed band, and no screen on what the comparables actually were.

Robert S. Stewart Jr.

FOUNDER & CEO, CR EQUITY AI, INC.

AUGUST 2026

·

9 MIN READ

6 DEFECT CATEGORIES FOUND AUDITING OUR OWN VALUATION MODULE	5 REQUIREMENTS IMPOSED BY THE AVM QUALITY CONTROL RULE
0 POINT VALUES PUBLISHED WHEN EVIDENCE IS INSUFFICIENT	Band DISCLOSED ON EVERY AUTOMATED ESTIMATE

A valuation model returns \$412,000. Not \$410,000, not "roughly \$400,000 to \$430,000" — \$412,000, rendered in the same confident typeface whether it was derived from eleven tightly clustered arm's-length sales on the same street or from three distressed transfers two miles away in a different school district.

That is the failure mode. It is not that automated valuation is inaccurate; on dense, homogeneous residential stock it is often more consistent than a hurried appraisal. The failure is **false precision** — the presentation of an estimate as though it were a measurement, with the uncertainty stripped out before anyone downstream can price it.

SECTION 01

What the rule actually asks for

The interagency quality control standards for automated valuation models are not, in substance, a technology mandate. They are a set of five obligations that read more like governance than data science: maintain a high level of confidence in accuracy; protect against manipulation of the data; avoid conflicts of interest; require random sample testing and reviews; and comply with applicable nondiscrimination law.

Read those in order and a pattern appears. Only the first is about being right. The other four are about being *accountable* for being right — which is a different engineering problem, and the one most vendors have not solved.

SECTION 02

Six things we found auditing ourselves

We put our own valuation module through a structured audit rather than waiting for a counterparty to do it for us. It found defects. I am publishing the categories because a firm that claims its valuation stack emerged clean from its first audit is either not auditing or not telling you.

EXHIBIT 1

Defect categories and remediation

Findings from internal review of the collateral valuation module, and the control implemented in response.

DEFECT	WHY IT MATTERS	CONTROL IMPLEMENTED
Point value with no band	Downstream users cannot distinguish a well-evidenced estimate from a thin one.	Confidence band computed and published with every estimate.
Assessor anchoring	The model quietly converged toward the tax-assessed figure, importing the assessor's lag and method.	Assessed value demoted to a corroborating input; never a primary driver.
Unscreened comparables	Foreclosure, intrafamily and portfolio transfers were weighted as arm's-length sales.	Transaction-character screen applied before weighting; excluded sales listed.
Stale comparables	Sales outside the recency window were used without time adjustment.	Hard recency window; time adjustment required and disclosed beyond it.
Proxy variables	Inputs correlated with protected characteristics entered the value estimate.	Removed from valuation entirely; see Section 04.
Silent fallback	A weaker data source substituted without the substitution being recorded.	Source class stored on every field; fallback events logged and surfaced.
Summary of internal findings, condensed. Remediation status is tracked in the firm's model risk register and is available to counterparties under diligence.		

The assessor-anchoring finding deserves particular attention, because it is subtle and it is common. A model trained partly on assessment data will tend to reproduce assessment behavior, including its jurisdictional lag and its systematic errors. The output looks independent. It is not.

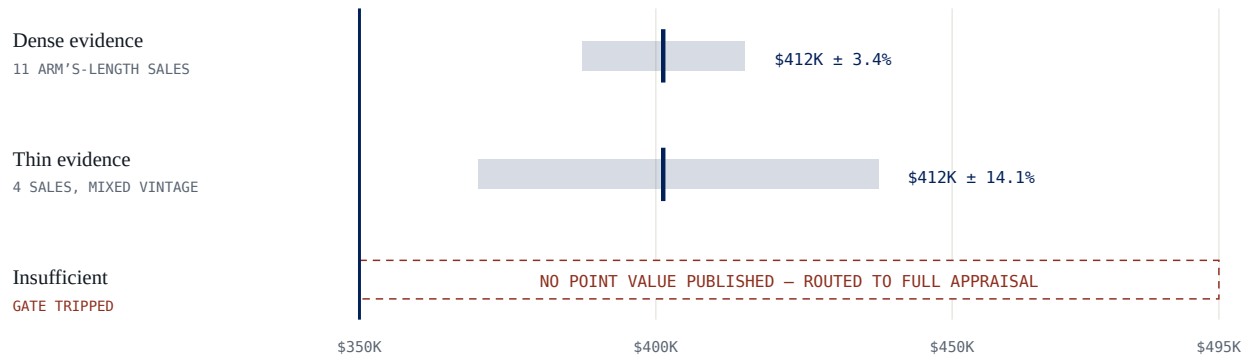
SECTION 02B

What an honest estimate looks like

EXHIBIT 2

Three evidence states, three outputs

The width of the band is the message. When evidence is insufficient, the correct output is no point value at all.



Illustrative. Band widths shown for explanation and do not represent published tolerances for any asset class or market.

SECTION 03

The insufficient-evidence gate

The hardest engineering decision in this module was building a system willing to say nothing.

Every incentive pushes the other way. A product that always returns a number demonstrates better. A pipeline that always returns a number never stalls. But a valuation produced from evidence that cannot support it is worse than no valuation, because it enters the file with the same authority as a good one and gets relied upon by people who never saw the comparables.

So the gate exists: below a defined evidence threshold — too few qualifying sales, excessive dispersion, an asset class outside the model's competence — the engine declines to publish a point value and routes the file to full appraisal. The file does not stop. It simply stops pretending.

A system that cannot decline to answer has no way to tell you when it is guessing.

SECTION 04

Where fair lending exposure hides

Automated valuation carries a specific and underappreciated fair lending risk, and it does not arrive through the obvious door.

No credible vendor uses protected characteristics as inputs. The exposure comes from proxies — variables that correlate with those characteristics closely enough to reproduce their effects. School ratings. Crime indices. Composite "neighborhood quality" or "desirability" scores sold by third parties, whose construction the buyer cannot inspect. Each is defensible in isolation and each is a candidate for encoding historical patterns into a current value.

Our position is deliberately conservative: those variables do not enter the value estimate. Collateral value is derived from the physical asset and from qualifying transactions of comparable physical assets. Location enters through the comparables themselves — the sales that actually occurred — not through a purchased score describing who lives there. This costs some explanatory power. It is worth it, and the nondiscrimination requirement in the AVM rule makes the trade increasingly non-optional.

SECTION 04B

Who is allowed to touch the number

Accuracy controls address whether a valuation is right. Independence controls address whether it was allowed to be right — and they are the requirement most often satisfied on paper and violated in practice.

The failure pattern is well documented and does not require anyone to act in bad faith. A file is short on value. A production person, whose compensation depends on closings, contacts a valuation provider, whose future assignments depend on that production person. Nothing improper is said. The number moves anyway. Multiply across an origination platform and you have systematically biased collateral, discovered years later by whoever bought the paper.

The controls that prevent it are structural rather than cultural:

- ◆ **Separation of ordering from production.** Valuation assignments are placed through a function with no compensation tied to whether the loan closes. Loan officers and brokers do not select the appraiser.
- ◆ **No coaching, no target.** No communication to a valuation provider may include the loan amount sought, the contract price where avoidable, an estimated value, or any indication of the value needed for the file to work.
- ◆ **Reconsideration of value on evidence only.** A value may be challenged — through a documented process, with specific additional comparables or factual corrections, routed through the independent function. It may never be challenged on the grounds that the number is inconvenient.
- ◆ **Vendor rotation and performance monitoring.** Concentration with a single provider in a market is itself a risk indicator, and variance between provider estimates and subsequent sale outcomes is tracked.
- ◆ **Automated estimates are subject to the same rules.** Nobody re-runs the model with adjusted inputs until it produces a workable figure. Re-runs are logged, and the first result is retained.

That last control is the one most specific to automated valuation and the one least often implemented. A model that can be quietly re-run is a model with a dial on it, and a dial in the hands of production is exactly what the independence requirements exist to prevent.

SECTION 05

Appraisal, model, or both

The useful question is not which method is better. It is which method is competent for a given asset, and the answer varies more than the marketing suggests.

EXHIBIT 3 Valuation approach by asset class Where an automated estimate is sufficient, where it is supporting evidence, and where it should not be relied upon.		
ASSET CLASS	EVIDENCE DENSITY	APPROACH
Single-family, homogeneous tract	High	Automated estimate with disclosed band; appraisal by exception.
Two to four unit residential	Moderate	Automated estimate supported by income approach; appraisal at higher leverage.
Small multifamily, five to twenty units	Moderate to low	Income approach primary; automated estimate as corroboration only.
Mixed-use and neighborhood retail	Low	Full appraisal. Automated estimate not relied upon for sizing.
Special purpose and hospitality	Very low	Full appraisal with operating analysis. Model gate closed.
Land and entitled development	Very low	Full appraisal; entitlement and absorption analysis required.
General framework. Specific requirements are set by program, leverage, jurisdiction and file condition.		

The through-line across all of it: a valuation is a claim about an uncertain quantity, and the honest way to publish a claim about an uncertain quantity is to publish the uncertainty with it. Every counterparty who buys, warehouses or insures this paper is entitled to know not only what we think the collateral is worth, but how confidently, on what evidence, and where we declined to guess.

Robert S. Stewart Jr. is Founder and Chief Executive Officer of CR Equity AI, Inc., an AI-native specialty real estate private credit and commercial lending platform. A U.S. military veteran and licensed real estate professional in Florida and Virginia, he founded CR Equity AI in 2021 and leads development of AIVAA™, the firm's proprietary underwriting and valuation engine. [Learn more at crequity.ai/aivaa](https://crequity.ai/aivaa).

SOURCES & FURTHER READING

1. Interagency final rule on quality control standards for automated valuation models.
2. Uniform Standards of Professional Appraisal Practice; federal appraisal independence requirements.
3. MISMO — collateral data standards.
4. Bloomberg — appraisal bias litigation and valuation integrity coverage.
5. Forbes and Yahoo Finance — proptech valuation accuracy claims and home price index divergence.

FOR INSTITUTIONS

See how AIVAA scores collateral confidence.

Asset-class gating, arm's-length comparable screening, disclosed confidence bands and an explicit insufficient-evidence state — documented for credit committees, warehouse counterparties and diligence teams.

[CLICK FOR MORE INFORMATION](#)

[REQUEST THE VALUATION POLICY](#)

Business-purpose lending only. Worked examples are illustrative and do not represent an offer of terms. Figures cited reflect data available as of the publication date and are drawn from the sources listed. Third-party publications and organizations referenced are not affiliated with, and do not endorse, CR Equity AI, Inc.