

## From Our Readers

### “Principles for Calculating AVM Performance Metrics”

#### To the Editor

“Principles for Calculating AVM Performance Metrics” (Winter 2020), by Hans R. Isakson, PhD, Mark D. Ecker, PhD, and Lee Kennedy, illustrates factors that impact the performance of automated valuation models (AVMs). I am somewhat concerned, however, that the assumptions presented in the article seem to obscure the point that selling price is not always equal to market value. This point is made in the Appraisal Institute textbooks, which talk about adjustments to price, but we may need a reminder.

The fact that selling price is not always market value is the primary assumption in the testing and validation of appraisal valuation models. Experienced appraisers know that sometimes you must adjust a selling price and not take price as a given. My point here is that sale price is not always representative of market value or most probable price. Assuming it is the best representation of market value automatically induces appraisal error, according to Isakson, Ecker, and Kennedy, which then is overstated. Reality suggests that some people pay more than average and some pay less. An appraisal that does not match selling price is not necessarily an error.

Why do we observe such a range of possible selling prices? The derivation of market value is via the most probable price. *The Appraisal of Real*

*Estate*, fifteenth edition, describes conditions required to derive a normative market value. These conditions include

- buyers and sellers that are typically motivated;
- parties that are well informed or well advised;
- a reasonable time allowed for exposure in the open market;
- payment made in cash or comparable financing; and
- price representing the normal consideration for the property sold unaffected by special or creative financing or sales concessions granted by anyone associated with the sale.<sup>1</sup>

These conditions are sometimes not known or not met. Buyers and sellers may not be typically motivated or equally informed. Sometimes sellers are motivated for a quicker sale and intentionally underprice the property to attract more buyers. This can result in a bidding war on occasion or a below-market sale price on other occasions.

Buyers and sellers possess different levels of information, have different tastes and preferences, and have different search costs and motivations. For example, there are several academic papers that show out-of-town buyers, with their higher search costs, tend to pay higher prices than local patient and more-informed buyers.<sup>2</sup> This informational advantage of local buyers has

1. See *The Appraisal of Real Estate*, 15th ed. (Chicago: Appraisal Institute, 2020), 48–50.

2. See, for example, Katrin Kandlbinder, Norman G. Miller, and Michael Sklarz, “Leveling the Playing Field: Out-of-Town Buyer Premiums in US Housing Markets Over Time,” *International Journal of Housing Markets and Analysis* 12, no. 3 (June 2019): 337–404; Val E. Lambson, Grant R. McQueen, and Barrett A. Slade, “Do Out-of-State Buyers Pay More for Real Estate? An Examination of Anchoring-Induced Bias and Search Costs,” *Real Estate Economics* 32, no. 1 (February 2004): 85–126; and Norman Miller, Michael Sklarz, and Nicholas Ordway, “Japanese Purchases, Exchange Rates and Speculation in Residential Real Estate Markets,” *Journal of Real Estate Research* 3, no. 3 (1988): 39–49.

held true even after various websites became the dominant method of home-buying research. In addition, the research indicates that sellers in financial distress or facing essential moves have lower reservation prices, and they list homes at lower prices to reduce time on the market. From an economist's point of view, unique tastes and preferences, high search costs, and high opportunity costs for either buyers or sellers are all rational factors that explain why we see a range of selling prices for very similar homes or properties. Such factors can, especially in thinner markets, result in a large range of potential selling prices that may be several percent above or below the most probable price. Sale prices even ten percent above or below the most probable price can be rational transaction prices. How should an appraiser treat such a selling price or price dispersion in general?

The first point is to recognize that price dispersion exists, and this will always result in some appraisal error based on unknown noise and unknown factors within the system that drives prices. Second, only a few of the observed transaction prices will be at a most probable price; most sales will be above or below. This brings us back to how to judge an automated valuation model when the contract sale price is unknown.

Isakson, Ecker, and Kennedy suggest using sale prices as the appropriate metric for judging the performance of a value estimate. Compare 10,000 appraisals to 10,000 selling prices, and the closer you are to the sale price the better, or so they presume. In aggregate, this is true, and all appraisal models are geared towards explaining selling price. On average, if you group together many appraisals, they will be close (within a few percentage points) to the market aggregate aver-

age selling price over the entire market. But at the individual level, it is wrong to presume that sale price is the best measure of the most probable price. Sale prices will deviate from most probable price. If a home sells for more or less because the seller is in no hurry or the buyer is desperate, that does not mean the appraisal is wrong. Some people pay more than market value and some less. Yet, according to Isakson, Ecker, and Kennedy, it is an error and the valuer or model will be faulted for under appraising the property using the kinds of statistics suggested. They are wrong. The best appraisers and appraisal models will often conclude market value is above or below the sale price.<sup>3</sup>

Isakson, Ecker, and Kennedy state, "Due to the proprietary intellectual property contained within an AVM, assessing AVM credibility, i.e., its validity, accuracy, and precision, is accomplished through an examination of the AVM's performance metrics." They go on to describe several metrics, including the forecast standard deviation (FSD) of the error, again based on sale price not based on most probable price, which they do not know. They suggest that the FSD of appraisal error is likely underreported and that is why we all need someone like a third-party vendor to calculate them for us. They also suggest buckets of "error" to look for extreme errors or compare buckets to price tiers or geography, but none of this is relevant unless you assume that sale price is the only possible correct value conclusion. It ignores inherent price dispersion within the market and the real purpose behind most home appraisals, which is to support financing risk analysis.<sup>4</sup>

With respect to collateral risk evaluation as a primary purpose for most residential appraisals,

---

3. Unfortunately, for financing purposes there may be strong lender pressure to hit the mark by concluding a value at or above the sale price; those who instead choose to protect lenders and mortgage investors may find themselves with fewer appraisal assignments.

4. Other appraisal purposes can be for investors or corporate resellers, but in such cases, there is no sales contract and no known sale price.

the metrics suggested by Isakson, Ecker, and Kennedy are mostly aimed at deriving work for third-party vendors. A better metric is how well do the appraised values serve the needs of the clients. For example, how much is the typical loss in foreclosure relative to the appraised value or model used? Do the sale prices or the appraised values predict selling prices better in the event of a foreclosure? These would be metrics worthy of consideration.

The appraisal industry has been under great pressure to speed up efficiency, lower costs, and improve accuracy. Many appraisal firms use automated data sorting and filtering and hybrid automated models that now allow their appraisal experience to be brought into the process. But, we should not fall back on using evaluation tools that assume the selling price is always the best estimate of market value. In fact, often it is higher or lower than the most probable price, and buyers may pay more or less than average. Not everyone pays the average price and to assume whatever they pay is market value oversimplifies reality and is a false premise. Instead of discussing trials and errors, confidence levels, and FSD by price range, all measured from this invented index of market value (sale price), the best measure of accuracy is how well does the value estimate perform in explaining losses after foreclosures. That would tell us much more about how good a value estimate is, rather than how well it conforms to selling prices.

Norman G. Miller, PhD  
San Diego, California

### Authors' Response

Thank you to Dr. Norman Miller for his letter to the editor regarding our article, "Principles for Calculating AVM Performance Metrics," published in *The Appraisal Journal* (Winter 2020).

Our article describes the state-of-the-art performance metrics that automated valuation model (AVM) vendors and others currently use to assess the accuracy, precision, and reliability of the AVM valuations. These metrics are ubiquitous within the AVM industry.<sup>5</sup> The purpose of our article is to (1) illustrate that the calculations of AVM performance metrics are not standardized among AVM vendors, and (2) present a set of calculation principles that, if followed by AVM vendors, would result in directly comparable, standardized, AVM performance metrics. The four calculation principles that we present are in concordance with established appraisal principles. We argue that AVM vendors would benefit by following standard appraisal principles. Given the increasing use of AVM valuations as potential substitutes for in-person appraisals, we believe appraisers would also benefit by having a working knowledge of the AVM industry and, in particular, how AVM providers assess the credibility of their estimates of value.

In his letter, Dr. Miller argues that the proposed index of market value (sale price) should not be used to evaluate valuation models, because selling prices are tainted by price dispersion. He explains that, sometimes, buyers might pay more, and sellers might accept less, than market value for a property—thus creating a transaction in which selling price deviates from market value.

5. See Miriam Steurer and Robert J. Hill, "Metrics for Measuring the Performance of Machine Learning Prediction Models: An Application to the Housing Market" (Graz Economics Paper, February 2020), available at <https://bit.ly/38lwGJz>; and Mark D. Ecker, Hans R. Isakson, and Lee Kennedy, "An Exposition of AVM Performance Metrics," *Journal of Real Estate Practice and Education* (forthcoming), available at <https://bit.ly/3lslnIG>.

Due to this price dispersion, Dr. Miller suggests that a better benchmark for measuring valuation accuracy, supplanting selling price, is how well a valuation explains losses after foreclosures. Unfortunately, Dr. Miller does not provide any details or supporting references describing the practical implementation of his suggested measure for assessing valuation accuracy.

In economics, price dispersion refers to the situation in which two identical commodities sell for different prices at the same point in time and at the same location,<sup>6</sup> a concept related to the law of one price and arbitrage pricing.<sup>7</sup> In housing markets, price dispersion is more difficult to measure because two houses are never identical. Moreover, a statistical model that identifies and measures price dispersion also produces an estimate of market value. Then, according to Dr. Miller, this price dispersion invalidates the selling price determined by the parties with intimate knowledge of the transaction—the buyer and the seller. That is, the output from Dr. Miller’s statistical model is used to suggest that selling prices from arm’s-length transactions are not estimates of market value. In contrast, AVM valuations are compared to selling prices; they do not invalidate selling prices.

In his letter, Dr. Miller does not mention that buyers and sellers have incentives to reduce price dispersion. The out-of-town buyer who hires a

local buyer’s agent can avoid overpaying for a house. Likewise, the desperate seller can hire a seller’s agent to avoid accepting a below-market-value price for his or her house. Sellers’ and buyers’ agents in housing markets work to minimize price dispersion effects.

In the AVM industry, it is widely recognized that the price of a property negotiated under the conditions requisite to an arm’s-length transaction<sup>8</sup> represents the best indicator of the market value for a property. CoreLogic, a leading provider of AVM products to financial and other institutions, recognizes the importance of an arm’s-length selling price when it states, “a purchase price is really the best (and only) benchmark for a property’s true value.”<sup>9</sup> Consequentially, AVMs are routinely tested and calibrated using selling prices. That is, an AVM valuation of a subject property is compared to the selling price of the property, as of the same date. In a 2019 report, the Mortgage Bankers Association states, “AVMs are designed and calibrated based on data from millions of purchase transactions across various geographic markets and over time.”<sup>10</sup> Further evidence of the importance of selling prices has been expressed by the International Association of Assessing Officers (IAAO), which takes the position that “all sales are candidates as valid sales [representing market values] for the ratio study unless sufficient and

6. Bev Dahlby and Douglas West, “Price Dispersion in an Automobile Market,” *Journal of Political Economy* 94, no. 2 (1986): 418–438.

7. See Salim Rashid, “The ‘Law’ of One Price: Implausible, Yet Consequential,” *Quarterly Journal of Austrian Economics* 10, no. 1 (Spring 2007): 79, available at <https://bit.ly/32GHDqW>; and N. Gregory Mankiw, *Principles of Economics*, 6th ed. (Mason, OH: South-Western Cengage Learning, 2011), 686.

8. We undertook and reported the usual measures to “scrub” the sales data used to ensure that it included only arm’s-length transactions representative of market values; see footnotes 45 and 49 of our article.

9. CoreLogic, “MLS Data in AVMs: The Case For—and Against” (unpublished white paper, 2010), 2, available at <https://bit.ly/2K2Va5T>.

10. Valuation Analytics Workgroup, *The State of Automated Valuation Models in the Age of Big Data* (Washington, DC: Mortgage Bankers Association, January 2019), 14, available at <https://bit.ly/35nJ7Z6>.

compelling information can be documented to show otherwise.”<sup>11</sup> Despite Dr. Miller’s comments regarding price dispersion, these authorities (and others) accept that selling price in an arm’s-length transaction represents the best indicator of market value of a property.

In his letter, Dr. Miller offers a substitute for selling price to use as a benchmark in measuring valuation accuracy. He argues that “the best measure of accuracy is how well does the value estimate perform in explaining losses after foreclosures. That would tell us much more about how good a value estimate is, rather than how well it conforms to selling prices.” Here, Dr. Miller proposes a novel measure of AVM valuation (or appraised value) accuracy—one that would require the practitioner to wait months, if not years, for a mortgage to default before assessing the foreclosure losses associated with any valuation obtained at the time of the origination of the loan. Only then could one potentially assess how well or poorly the valuation explained any losses.

A fundamental question that Dr. Miller does not address is why any value estimate would be related to a mortgage subsequently going into default. That is, just because the lender suffered a foreclosure loss does not mean that the original selling price (or appraised value or AVM valuation) months or years beforehand was inaccurate. Attributing foreclosure losses to valuation errors is muddled with the true cause(s) of loan defaults, very few of which are valuation related.<sup>12</sup> A review of the academic literature

and industry press regarding the root causes of residential mortgage defaults can be found in the *Report to Congress on the Root Causes of the Foreclosure Crisis*.<sup>13</sup> None of the studies examined in that report advocates using foreclosure losses to supplant selling prices to measure valuation accuracy.

The use of selling prices to calculate sales errors, from which nearly all AVM performance metrics are derived, is ubiquitous in the AVM industry, in spite of Dr. Miller’s concern that these metrics ignore price dispersion within the market. The discussion of price dispersion is an academic exercise that has not been established to provide any real-world usefulness to today’s valuation practitioner. In addition, the proposed foreclosure-based metric for measuring the accuracy of valuations lacks a coherent definition and is unproven in practical application. We know of no mortgage originators, appraisers, AVM providers, or third-party AVM evaluators who make use of the foreclosure-based metric proposed by Dr. Miller, and it offers little to advance the readers’ understanding of currently used AVM performance metrics.

*Hans R. Isakson, PhD*  
Cedar Falls, Iowa

*Mark D. Ecker, PhD*  
Cedar Falls, Iowa

*Lee Kennedy*  
Simi Valley, California

11. IAAO, *Standard on Ratio Studies*, Assessment Standards of the International Association of Assessing Officers (Kansas City, MO: IAAO, 2013), 48, available <https://bit.ly/3po7HRT>.

12. The Home Mortgage Disclosure Act (<https://bit.ly/35odtKP>) lists 32 reasons for mortgage default, and none of them is related to valuation accuracy. See US Department of Housing and Urban Development, Appendix 2—Delinquency/Default Reason (DDR) Codes, available at <https://bit.ly/3acE68w>.

13. Office of Policy Development and Research, US Department of Housing and Urban Development (January 2010), available at <https://bit.ly/3EzDOZ>.

## “Revisiting Market Value and Market Rent”

---

### To the Editor

I read with interest the comments by Neil F. McDonald, MAI, AI-GRS, in his Summer 2020 letter to the editor related to *The Appraisal Journal* article “Revisiting Market Value and Market Rent” (Winter 2020). In his letter, Mr. McDonald states that an appraisal can be only “retrospective” or “prospective,” defined as before or after the “moment of the report.” How to properly apply the label “retrospective” to an appraisal is something that I have been wrestling with for quite some time.

The term *retrospective appraisal* is not found in the definitions section of the Uniform Standards of Professional Appraisal Practice (USPAP). However, USPAP Advisory Opinion 34 (AO-34) describes a *retrospective appraisal* as an appraisal where the “effective date of the appraisal [is] prior to the date of the report.” (Line 18) In the next paragraph, *prospective appraisal* is defined as an appraisal where the “effective date of the appraisal [is] subsequent to the date of the report.” (Line 21) This leaves the final category of *current appraisals*, which “occur when the effective date of the appraisal is contemporaneous with the date of the report.” (Line 15) This definitional framework, however, results in the label “retrospective” being applied to almost all appraisals, as the percentage of appraisals where the effective date is the same as, or after, the report date would be very small.

In the USPAP Frequently Asked Questions, FAQ 155, “Effective Date and Date of Report,” loosens the definitions considerably. It states that in a current appraisal “contemporaneous means arising, existing or occurring during the same time period,” and “contemporaneous is not intended to mean simultaneous.” The vagueness is defended by the assertion that the time period

could differ from assignment to assignment. FAQ 155 does mention one strict rule however; that rule is there cannot have been any significant change in property characteristics or market conditions between the effective date of the appraisal and the report date. This still leaves uncertainty in the characterizations of reports.

For example, say I appraise a barn, with an effective date on Friday, and a report date the following Monday. Is this a “current appraisal”? Do I need to know if something happened to the barn over the weekend. If the barn burned down on Saturday, the report suddenly becomes retrospective. Is the appraiser required to check on the property on the report date? Is this definition reasonable? Could we do better?

In my discussions, I have found little agreement about proper interpretation. One common response seems to be that no further precision is needed. While this may possibly be adequate in general, it is surely inadequate in other settings, such as for students being tested on the definitions in appraisal exams, litigation, etc. Therefore, in the interests of rigor, I offer the following definitions for consideration:

- *Retrospective appraisal*—the effective date is prior to the date of agreement to perform the assignment.
- *Current appraisal*—the effective date is between the date of agreement to perform the assignment and the report date (inclusive).
- *Prospective appraisal*—effective date is after the report date.

Stephen P. Reynolds, MAI, AI-GRS  
Lexington, Massachusetts