BEST PRACTICES IN AUTOMATED VALUATION MODEL (AVM) VALIDATION

January 2009

Presented by:

January 2009

Table of Contents

AVM TESTING & ANALYSIS

Purpose

The mission of the Collateral Assessment & Technologies Committee (CATC) is to promote and coordinate education and awareness of alternative collateral assessment tools and technologies including automated valuation models (AVMs), fraud detection tools, collateral scoring, forecasting applications and derivatives thereof. In these efforts, CATC places a premium, above all others, on the transparent and objective evaluation, implementation and application of these tools.

The AVM validation landscape has changed significantly over the past several years. These changes have been fueled in large part as a result of increased attention placed on the use of AVMs by the regulatory agencies that make up the Federal Financial Institutions Examination Council (FFIEC). Although the regulatory agencies had previously remained largely silent on AVM issues, their attention was another welcomed step in the further acceptance of AVMs throughout the mortgage process. Regulatory guidance first expressly referenced the use of AVMs on both a portfolio and transactional basis in December of 2004 in OCC Bulletin 2004-59.[1] Greater attention was dedicated to AVMs the following year in OCC Bulletin 2005-22[2] addressing broader issues of Credit & Collateral Risk Management practices. In these publications, the regulatory agencies referenced earlier guidance found in OCC Bulletin 2000-16[3] as the framework under which lending institutions were expected to conduct validations of their AVM testing and selection procedures. Although 2000-16 does not directly or indirectly address AVM models, it generally discusses the validation of any "model" employed by a lending institution and is still cited as the primary regulatory source for validations of AVM models.[4]

The growing spotlight on AVMs has helped demonstrate the positive contribution of AVMs to the collateral risk management landscape and further solidified their role as one of the leading real estate valuation and risk products and services. This spotlight has further led to the growth of relatively new markets centered on the testing and validation of AVMs. New market participants include third-party testing consultants, AVM test data providers, and cascade service providers. As with any new enterprise (or new participant in an existing market), it is incumbent upon these individuals, organizations or agencies to establish independence and fully educate themselves on the products, services, policies and procedures to

[1] OCC Bulletin 2004-59, "Retail Lending Examination Procedures".
[2] OCC Bulletin 2005-22, "Credit Risk Management Guidance for Home Equity Lending".
[3] OCC Bulletin 2000-16, "Risk Modeling – Model Validation".
[4] For example, see Proposed Interagency Appraisal and Evaluation Guidelines, November 2008.

which they hold themselves out as experts or are charged with oversight. More importantly, stakeholders new and established must be able to separate fact from fiction, identify potential conflicts of interest, and manage them appropriately. Additionally, as with any type of outsourced service, it is incumbent upon the AVM user to fully understand the processes and results as the responsibility for the outcome remains with the AVM user. This document will further discuss the roles and expectations of third-parties in the AVM validation process.

Background

The use of AVMs for valuation and collateral risk analysis in the mortgage and other real estate related industries has been well established since their introduction more than ten years ago. Prior to the 2007-2008 housing market slow-down, the industry saw increased volume and acceptance of alternative evaluations on first and second liens. This environment proved to be fertile testing ground for stakeholders exploring comfort levels with these relatively new but very effective collateral assessment alternatives.

Clearly, the efficiencies, objectivity and cost effectiveness of these evaluations increased volume as a result of one of the largest real estate booms in history. Nevertheless, AVMs were scarcely used on the subprime loans and first mortgages that led to the current market condition. However, as the housing market continues to undergo significant contraction, the use of these alternative evaluations is essential for prudent risk management as they provide for objective determinations of value and risk.[5]

As AVM users and regulators have grown to accept the use of AVMs as a whole, increased applications have been seen in many new areas of the industry such as prequalification screening, large portfolio valuation and analysis, marketing campaigns and initiatives, first purchase and second lien originations, implementation into automated underwriting systems, loan origination systems and collateral management systems, as well as secondary market acquisition, due diligence, RMBS and ABS securitization and servicing applications. Currently, many lenders and investors alike are using AVMs as a primary quality assurance mechanism to validate all chosen collateral valuation method results. The users of AVMs in the sectors described above include financial institutions such as originators, investors, ratings agencies, servicers, and any other organization or third-party that systematically employs automated valuation tools in the analysis of residential real property ("AVM user").

[5] For a discussion of AVM performance in soft or declining markets, please see the 2004 CATC whitepaper on this topic titled "Systemic Risk in Residential Property Valuation Perceptions and Reality" available from www.catconline.com. *See also*, "Automated Valuation Models Provide Better Protection In Soft Markets," Michael G. Bradley and Mark A. Beardsell, Secondary Marketing Executive, Apr. 2005, p. 38.

AVM usage will continue to expand into new and previously untapped arenas due in part to technological advances and increases in data availability, just as the introduction of analytics such as fitness scoring, fraud scoring and collateral scoring in which the AVM may be a core component has further driven AVM usage in the mortgage lending industry. In addition, specialty AVMs such as default or REO ("real estate owned") AVMs, valuation forecasting AVMs, market scoring, and variations of AVMs based on selective market criteria such as accuracy, hit rate, or lien type have been recently introduced. These products and services will not be fully discussed in this document but must be independently validated in accordance with best practices and current interagency guidelines.

AVMs have significantly evolved over the last several decades. Current hybrid models employ multiple methodologies which create unique and complex modeling that cannot be easily or quickly understood. Therefore, a comprehensive and consistent approach is required to understand and evaluate the modeling results.

There are real costs associated with the processing of test files and the maintenance of duplicative testing environments to support the burgeoning volume. Given the exponential growth in testing volumes and frequency the associated costs have become burdensome for the AVM providers. To support this type of processing going forward these types of costs will need to be offset. Historically the providers have gone uncompensated for their participation in these tests. Moving forward, our recommendation is for a free and open sharing, between all parties, of robust data files, detailed test results, AVM approval processes, and final AVM cascade structure; together with the basis for those decisions. The concept of "transparency" serves to benefit all of the industry. In the absence of this type of information sharing, financial consideration for records processed would be required. Third party consultants and test data providers involved in the process are compensated for the validation process and it is only appropriate that AVM providers be treated the same.

AVM Strengths & Weaknesses

An AVM is one of the most cost effective collateral assessment tools available for fast, objective and accurate valuations of residential property on either an individual or portfolio basis. However, that does not mean AVMs are designed for optimal performance in every situation. To the contrary, no real estate valuation product or service (e.g. AVM, appraisal or BPO ("broker price opinion")) should be arbitrarily considered as viable for use in each and every valuation scenario. AVMs are just one of the tools that any organization interested in collateral valuation should utilize given proper validation procedures and understanding of

expected performance. It is incumbent upon any lender, investor or similarly situated stakeholder to determine the applicability and associated risk of *any* valuation service(s) selected on a given loan or property.

Objective Determinations of Value

In contrast to the appraisal process where lenders and brokers may exert pressure on appraisers to meet a target value, an AVM is objective and provides a market estimate of value without any outside influence. According to a major industry study released in 2007, over 90% of appraisers responded that they have felt lender pressure to "meet value". Just four years prior, only 55% of appraisers felt pressured by lenders.[6] This "pressure" or influence does not exist with AVMs. As a direct result of years of lenders pressuring and influencing appraisers to "hit a value", the NY Attorney General created of the Home Valuation Code of Conduct to address the "bias" in appraisals.[7]

Accuracy & Loan Performance

No appraisal or other evaluation product has undergone the rigorous level of testing and scrutiny as AVMs. The performance of AVMs in both appreciating and declining markets has demonstrated their level of accuracy and the unbiased nature of their results when compared to other valuation techniques. Contrary to conventional wisdom this makes AVMs the most understood tool, in terms of valuation accuracy performance, available to lenders today. Experience has shown that, with all other performance factors held equal, loans underwritten with AVMs have outperformed similar loans underwritten utilizing appraisals or other evaluation methods.[8]

Rapid Fulfillment

An AVM report, including the estimated market value of the subject property, is returned within seconds, whether ordered directly or as a part of an overall valuations cascade. This permits AVM users to meet their requirements and satisfy their customer requests in a matter of moments, as compared to the days that it takes other valuation product requests to be fulfilled.

[6] October Research National Appraisal Survey, 2004 and 2007.
[7] Home Valuation Code of Conduct
[8] For a discussion of AVM performance in soft or declining markets, please see the 2004 CATC whitepaper on this topic titled "Systemic Risk in Residential Property Valuation Perceptions and Reality" available from www.catconline.com.

Cost Effectiveness

The cost of individual AVM reports has declined rapidly over the past several years because of the rapid expansion of computing power, the escalated usage of the AVM reports, and competition within the industry. AVMs are often the most cost effective valuation tool that an AVM user can employ, resulting in substantial savings in the valuation process, including the risk component.

Limitations

Generally, AVMs are designed to perform optimally in the valuation of predominantly homogeneous properties in urban and suburban areas that have large amounts of historical and current data. These properties account for a very high percentage of the total U.S. residential real estate market. In much the same way as appraisals, BPOs, and other forms of valuation, some AVMs may experience performance degradation as a function of price (e.g. at the value tails for both very low-end and high-end price levels), geography, property type (e.g. single family residence vs. condo) or property age (e.g. due to the lack of available data in relatively new developments). Fortunately for the AVM user, better AVMs will be significantly less susceptible to these influences, and further, these issues can be effectively tested employing unbiased and appropriate validation procedures.

AVM Model Standardization

There are standard or common *elements* (i.e. analytics or data output) of industry-accepted AVMs that can ease the process of comparing seemingly disparate models. These fields include: Estimated Market Value, Value Range (reasonable high and low values), Confidence Score, and a standardized subject address.

Many industry participants have raised the question of whether model "standardization" should be imposed. This is predominantly discussed relative to AVM confidence scoring. The discussion of model "standardization" can be misleading and is generally cited by those unfamiliar with AVM development or use. AVM models represent unique approaches that may have fundamental differences in design and objectives by definition. No two AVMs were developed the same way; each has its own focus, strengths, sources of data, and valuation techniques. However, analysis, selection, and use of these unique models can be accomplished through prudent validation procedures that would eliminate the concerns raised by those who incorrectly cite "standardization" as the answer. It matters less how you get to the answer, but that the answer is accurate and that the model will remain so as market dynamics change. Consistent, timely,

frequent, and robust testing of the accuracy and performance under varying market conditions will answer these questions for an individual AVM or cascade.

Varying expectations of performance based on an organization's appetite for risk and intended uses further argues against calls for standardization. Risk tolerances may appropriately vary from one organization to the next as well as within an organization across different applications. This legitimate reality has led, in part, to the increased adoption of AVM cascades discussed more fully below.

The various strengths of each model can be leveraged through sound AVM testing and analysis, permitting AVM users to establish appropriate criteria for the use of each model to generate optimal performance. The key for any organization is to first and foremost decide what is most important in its use of an AVM as a collateral assessment alternative. *If an organization cannot first decide on the goal or expected outcome, it's impossible to establish any effective plan*.

To assist lenders in assessing the various commercially available AVMs, the regulatory agencies that make up FFIEC have written guidelines to provide a basic framework around sound model validation policy.[9] By performing well constructed tests, understanding the model logic and continually validating the outcome of the testing/validation process, AVM users can utilize these valuation tools with confidence.

Challenges of AVM Analysis

As the mortgage industry continues to expand the use of AVMs as an effective collateral assessment alternative, the testing and analysis of this decision-support tool has become increasingly important for its successful integration into any corporate environment. The evaluation of any enterprise analytic solution should ultimately take into consideration both the stated purpose of the application (e.g. accurate residential property valuation) and its particular use in the organization where it will be implemented (e.g. prequalification, origination, QC/Audit, marketing, or servicing). However, organizations that seek to evaluate an AVM for its potential effectiveness all too often ignore the means needed to achieve their end. The testing procedures used to accurately solicit results that will be reviewed against stated performance metrics, minimum threshold requirements or other performance objectives are critical to achieving realistic expectations of performance in a production setting.

This paper attempts to identify and educate the industry on "best practices" for AVM testing and evaluation. Although not all-inclusive, this paper attempts to set

[9] Proposed Interagency Appraisal and Evaluation Guidelines, November 2008.

forth a broad range of methodologies and procedures that may be selected individually or in concert depending on an organization's implementation strategy. This document is not intended to be overly prescriptive, but instead establishes a framework in which an organization can effectively validate AVM models to meet identified risk tolerances and particular applications in an objective fashion.

AVM Testing Considerations

Although pre-testing evaluation and due diligence are important to achieve an understanding of each product, an organization's decision of which AVMs to use and in which circumstances will be in large part due to the results of testing and analysis. Therefore, an AVM user must conduct a valid and accurate evaluation to provide the best possible opportunity to select the optimal AVM(s) for its needs. An understanding of testing procedures, particularly the frequency of AVM testing, and the selection of input data and benchmarks will be vital to the AVM user throughout this process.

Third-Party Testing Consultants

The task of AVM testing and due diligence can be both time consuming and resource intensive. For some AVM users, this means utilizing external testing sources for some or all phases of the testing process. Several third-party testing consultants have emerged to gather data, perform and review AVM test results, create comparative analyses, and make recommendations as to the best AVM selection and cascades based on individual client circumstances. While the use of these third-party consultants has become a commonly accepted practice industry-wide, selecting the right company to perform these analyses is as important as selecting an appropriate AVM.

An appropriate third-party is one that is completely independent and transparent with all participants, including the AVM providers. It is essential that third-parties provide the same detailed information back to all participants (AVM users and AVM providers) including the test design and procedures, benchmarks, resulting conclusions and blind comparisons, preferably at both the loan level and in aggregate. This *critical* transparency establishes an even playing field allowing all AVM providers an equal opportunity to succeed. It also mitigates any conflicts of interests and permits all participants to effectively validate results and benchmarks. Open communication during the validation process ensures that the assumptions and decisions have been properly vetted.

Prior to testing, AVM users should execute proper testing agreements. In the event an AVM user engages a third-party testing consultant, tri-party contracts with both the provider and the consultant must adequately address the disclosure

of confidential information between the parties, the duties and responsibilities of each party, the data that is required to be shared, and permitted uses of the test data by all participants in the test process. AVM users that chose to use a third-party should still understand all aspects of the third party's process including: data selection and acquisition, data validation, analytic process, analytic benchmarks, result reporting, and recommendations.

A list of topics related to the selection of these third-party providers can be found in Appendix A.

Cascade Validation

As a result of the unique differences in AVMs as well as the varying and legitimate applications discussed above, technologies have been created to extract the best performance among multiple products. These technologies have become known as AVM cascades or waterfalls and have become the rule rather than the exception over recent years. In the case of an AVM cascade, however, the whole does not equal the sum of the parts. When an AVM user or cascade service provider develops an AVM cascade, an entirely new system is created that performs independently of the individual AVM components. As a result, any AVM cascade *itself* must be validated independently from the individual AVM models which will be discussed more fully below.

Validation of the Production Environment

The primary reason for AVM validations is improved production performance in the form of objectives (e.g. more accurate valuations, reduced costs, lower default rates, and decreased loss severity). It is critical that the AVM user not detach AVM test results from real-time production performance. The AVM validation process does not stop with the AVM test. To avoid being surprised by a significantly different production experience than one would have expected based on AVM test results, AVM users should take steps necessary to ensure that the AVM testing results, performance/accuracy data and cascade recommendations being presented to the AVM user are validated in production and meet their internal goals regarding use of these automated tools.

If an AVM's performance in production is materially different or worse than expectations from the testing process, the AVM user should contact the third party or AVM provider directly to determine why there is a difference and attempt to resolve the cause in future testing efforts. The AVM user may also want to change the cascade and remove an AVM from their cascade if the AVM is performing worse in production than it did in testing. In the case of AVM platforms, production validation would include the reporting of transactional AVM usage by each AVM provider. CATC will provide sample audit and other reports online at www.catconline.com.

THE AVM TEST PROCESS

The AVM validation process involves several steps that may include most or all of the following: (a) design of a test to solicit results consistent with risk tolerances and expected use; (b) preparation of an appropriate test sample and benchmarks; (c) pre-test due diligence of AVM providers and any third-party participants; (d) selection of test participants and third-parties; (e) completion of necessary documentation between the parties; (f) execution of the test within required parameters; (g) determination of analysis subset(s); (h) individual AVM model analysis; (i) determination of ranking functions for cascade development; (j) cascade simulation, testing and validation; (k) selection of appropriate AVMs and cascade platform providers; (l) AVM or cascade implementation; (m) AVM or cascade production testing and launch; (n) procedures to validate production experience against both cascade and individual test results; and, (o) presentation and discussion of results with the test participants.

Pre-Testing Evaluation & Due Diligence

Any AVM provider should already be performing its own significant due diligence evaluation on a regular basis. The goal of this internal testing and evaluation should include (a) model development, modification or calibration, (b) data integrity analysis, cleaning and matching where needed, (c) the determination of coverage areas that meet minimum performance thresholds, (d) analysis of performance degradation over multiple categories, and (e) determination of confidence score correlation to various accuracy metrics. AVM providers should be able to share high level information about their own procedures and approaches to internal model validation however, given the proprietary nature of this AVM technology, the disclosure of this type of information will be limited.

Although most AVM providers will never divulge the algorithms and other ingredients that make up their proprietary methodologies, any AVM provider should “open the box” to the greatest extent possible without threatening the proprietary nature of its intellectual property. It is the responsibility of the AVM user to expect and the AVM provider to deliver the information that will allow for an informed decision. While some AVM users and other stakeholders have demonstrated an appropriate level of testing, documentation and control surrounding the use of AVMs, not all AVM users have the same level of scrutiny, experience and knowledge which may create issues during regulatory examinations or performance degradation in a production environment.

For those evaluating AVM providers and other third-parties associated with the validation process, it is important to determine what due diligence procedures, if

any, are being employed by each provider and address some basic information involving each model or third-party *prior to any testing*. In Appendix A, a list of topics can be found that AVM providers or third-parties should be willing and able to provide. The response to these questions will provide better insight into the capabilities of the providers chosen for evaluation. This review of the logical and conceptual soundness of the model(s) and its developers is essential.

Frequency of Testing

Individual AVM model performance may change over time due to: (a) data availability and integration, (b) model development and architecture, (c) backend data infrastructure, (d) in-house intellectual knowledge and personnel, (e) hardware systems, (f) corporate vision and management, (g) software development and functionality, and (h) as AVM providers merge, become acquired, or cease operations. Existing AVM providers may neglect to update and maintain their systems, or newcomers may introduce alternative methodologies. In the end, the AVM landscape is constantly evolving – as a whole – in a very positive direction. However, changes in individual AVM provider performance over time can be significant for the better or worse. Without appropriate periodic testing and analysis, AVM users may be failing to receive the highest quality results for which they are paying.

Further, an AVM provider's performance may vary from one test to another over time. AVM users should measure and track performance volatility over time and determine their own tolerance. Those AVM providers that exceed expected volatility thresholds should be more deeply scrutinized.

This fact places the burden on those responsible for AVM evaluation to continuously re-evaluate their organization's testing procedures and repeat testing on a periodic and regular basis. Ideally, AVM tests should be conducted at least semi-annually. Some AVM users are beginning to test on a quarterly or monthly basis. However, at a minimum, an AVM user should re-test both its current and outside AVM providers on an annual basis. As referred to above, each separate test event may involve the periodic collection of test results from AVM providers to obtain a statistically significant sample.

Test Samples

The most critical concern of any AVM validation process is to establish a process that ensures AVM providers do not have access to any benchmark or other information that results in anything less than an objective determination of AVM performance.

As a result of broadened technology, improvements in computing power and expansion of digitized residential real property data availability, AVM providers have become increasingly faster in loading property information to their databases. Historically, arms length sales records were deemed to be relatively unknown to AVM providers if they were less than 60 days old from recordation. However, most of this sales data is now readily available within 30 days. While this improvement in the currency and depth of AVM provider databases is a desired outcome, it makes the construction of an appropriate benchmark file even more challenging.

Currently, even pre-sales loan information (a.k.a. pipeline information on loans scheduled but not yet closed) can potentially become available to AVM providers before the pending sales have occurred. AVM users should consider the possibility that certain types of pre-sales arms-length transactions may already exist in the AVM provider's database, directly or indirectly via related business units, at the time of a validation test:

- Arms-length transactions with an agreed upon sales amount and closing date, that have not yet closed, and the AVM user has provider work that is in-process.
    - Preliminary title commitment has not yet been requested
    - Appraisal report has been ordered using in-house staff appraisers and/or proprietary closed systems
    - Appraisal report has not yet been ordered through an Appraisal Management Company (AMC) or 3rd party appraiser who may be using portal services
- Arms-length transactions with an agreed upon sales amount and closing date, not yet closed, with all external provider work completed.
- Arms-length transactions closed within 1 – 3 days, For Sale By Owner
- Arms-length sales transaction activity that may potentially be listed in MLS'
- Many county deed offices are now automated and make sales recordation available within days of the closing date
- Arms-length transactions closed within 30 days

When selecting benchmark values for AVM validation testing it is optimal to choose transactions that are the least likely to be known to the AVM providers. This capability may not be available to all AVM users. As such, AVM validation results should be reviewed to assess whether any pre-sales benchmarks may have been known to the AVM providers. It may be prudent to inquire with the providers to determine if such data is included. If this is found to be the case then these transactions should be backed out of the validation test for all AVM providers.

The preparation of the test sample is the single most important aspect of an AVM validation yet the most often overlooked. Considerations should include which benchmark or baseline values to use, property types, price ranges, volume levels (county or state level), an AVM user's specific business footprint, if any, and may even include properties to test for specific model type capabilities or lack thereof. Once again, preparation of the test should always take into consideration the planned use of AVMs within a particular organization. Nevertheless, there are several general guidelines that should be followed, if at all possible, to get a complete picture of the AVM's capabilities.

There are other considerations beyond benchmark values that judge the effectiveness of an AVM or an evaluation in determining asset performance. These may include loan performance, default rates, loan types, lien positions, credit scores, loan to value, and other determinants of creditworthiness. Although not the direct subject of this document, a complete risk management regimen needs to consider these factors.

Selecting the Records & Benchmarks

When creating an input file for AVM analysis, the AVM user or third party should provide a sample set of properties representative of its business footprint. In compiling a sample for an AVM test, it should be representative, unbiased, and a statistically significant number of records at a sufficient rate of oversampling to allow for exclusion of records. Where there are contiguous geographies, the aggregation of records across multiple jurisdictions is appropriate for the design of a valid sample.[10] Some records will be thrown out due to: (a) an AVM provider having the benchmark (the "answer") in its database prior to the test; (b) the record was determined to be an outlier or otherwise excluded as discussed below; or, (c) the record is determined to be invalid.

There has been debate regarding the use of different types of benchmarks. While CATC recommends arms-length purchase transactions as the most reliable, CATC recognizes that the use of alternative benchmarks may provide additional insight into AVM performance considering the particular AVM application and data availability. In any case, the analysis of AVM performance should be segmented by benchmark type in order to identify and consider any inconsistencies.

Arms-length Purchase Money Transactions

Arms-length purchase money transactions provide the best indicator of value since there is a willing buyer and a willing seller in the open market (a.k.a. arms-

[10] "On-Target Or Off-Base? How To Ensure AVM Accuracy," Michael G. Bradley & Philip Nuetzel, Secondary Marketing Executive, October 2007, p. 32.

length transaction). Higher credit grade loan transactions seem to have less variance than subprime loans.

- Sales within the past 1 to 30 days are typically not present in an AVM database that utilizes public record data. This is important because some AVMs may return the sales price as the AVM value given in public records if there is one, leading the tester to believe that the AVM tested is far more accurate than it would be in a production environment where the AVM definitely would not be privy to the prior recent sales information for a given property.

- Due to the lack of current purchase transaction data, some AVM users have provided older purchase transactions (e.g. 6-12 months old) as part of an AVM validation. The AVM providers are asked to value these properties retrospectively to a predefined date either on the aggregate or property level basis. CATC generally does not recommend this type of "retro" AVM validation.

- Transactions that are scheduled to close but have not funded are likely not to be in public records and are considered the best records to test. As stated previously, diligence should be performed to assess whether any pre-sales benchmarks may have been known to the AVM providers.

- Generally, the selection of non-arms-length transactions should be considered unreliable.

- If REO properties are to be included in the sample, these properties should be aggregated and analyzed separately. A separate test using REO AVMs may also be appropriate.

Refinance Transactions

Refinance transactions do not provide a "market-tested" value and tend to result in overvalued collateral. Streamlined programs often allow the use of an original appraisal. Therefore, loan origination date and collateral valuation date can vary by many months causing either a high or low benchmark value dependent on the change in market condition during the interim period. Refinance appraisals done for cash out and mortgage insurance removal tend to test the higher end of the value curve resulting in a high benchmark value. If refinance appraisals are included in a test sample, these properties should be aggregated and analyzed separately.

Appraisals

While arm's length transaction purchase data is preferable, not every AVM user has sufficient data of this type for testing purposes. In this case, appraised values may be given consideration. If appraised values are to be used as the benchmark for the AVM comparison, it is important that only appraisals performed on purchase loan transactions are used as benchmarks and that both the type of appraisal and appraisal date be provided. It should be identified when preparing the AVM analysis that appraisal values were used, since there are variances between the appraised value and the actual sale price for any given property, some of which may be significant. Considering these differences one should place less weight on these benchmarks compared to an arms length transaction. Again, these properties should be analyzed separately from other benchmark types.

However, as results from AVMs are replacing the need for appraisals in certain cases, the appropriate usage of appraisal values and appraisal dates as benchmarks can measure the potential incremental risk, if any, of using an AVM value in making an underwriting decision to that of the appraisal. Note that appraisals are currently the most accepted valuation approach for underwriting many types of loans. This approach should permit ranking of the tested AVMs versus the appraisal benchmarks to measure each AVMs performance in this regard.

Aggregated Test Results & Purchased Test Data

If an AVM user cannot obtain a sufficient number of good records internally for use as benchmarks, the next best option may be to aggregate test results over a reasonable time period in order to generate a more statistically significant test sample. For example, this type of testing may be conducted on monthly or even greater frequency.

In some cases, the AVM user may not be able to generate even periodic test files for aggregation. In other cases, objectives may exist which would require otherwise unavailable data (e.g. the AVM user wishes to expand beyond their existing footprint). In such events, the alternative may be to purchase records from other sources such as the public record data aggregators. In these scenarios, significant conflicts of interest may arise that require immediate disclosure to all participants including the AVM user, third-party testing consultants and the AVM providers. For example, many public records data aggregators also develop or distribute their own AVM models. In these scenarios, CATC recommends that another data source be used for benchmarks in the AVM testing or a process created that permits the AVM user access to keyed, yet unposted recorder sales. Specifically, if records are purchased from such aggregators and the aggregator also has proprietary AVMs, CATC

recommends that the aggregators AVMs be excluded from that particular test. If this is not practical, strict rules should be established prior to testing, including contractual verbiage and an auditable process that is transparent to all parties to ensure that no "conflict of interest" exists. Further, test results and reports should disclose all sources of test data.

Other Benchmark Values

Customer estimated values, MLS list prices and other valuation results (BPOs) offer little consistency and should not be used as benchmarks.

Non-disclosure states: Obtaining benchmark data for non-disclosure states presents another problem in AVM testing. Non-disclosure states are those where sales prices are not made public at the recordation of the deed. While definitions of non-disclosure status may vary, these jurisdictions include states such as Alaska, Texas, Utah, New Mexico, Kansas, Mississippi, Missouri, Idaho, Iowa, Utah, and Wyoming. It is important that an accurate "market value" is used as a benchmark when analyzing AVM performance in non-disclosure states. The benchmark in non-disclosure states should not be a derived value or a "best guess". Appraisal data and the AVM user's own purchase data should be used to quantify performance. It is especially important to perform post production testing on AVM performance in non-disclosure states to make sure that the test performance is in fact what the AVM user is experiencing in their day to day use of AVMs in production.

Although not readily available at this time, alternative benchmarks may be pursued in an effort to reduce the likelihood of any AVM provider from being privy to benchmark sales data. These alternatives include capturing purchase transactions that have not yet closed (e.g. escrow records), running these properties against AVM models, and then subsequently identifying closed sales prices. However, it is important to note these alternatives assume that no AVM provider has access to databases that contain pending sales transactions such as MLS or Title data as access to such data would introduce significant bias.

Address Standardization

Most AVM providers employ address standardization or property address matching algorithms in various attempts to standardize the property address prior to valuation of the property. Typically, in a production environment AVM providers are not supplied "pre-scrubbed" property addresses by the AVM user. In an effort to realize a true production experience with respect to hit rate and accuracy, CATC does not recommend the pre-scrubbing of addresses by an AVM user or third party prior to sending the addresses to AVM providers.

Input Fields

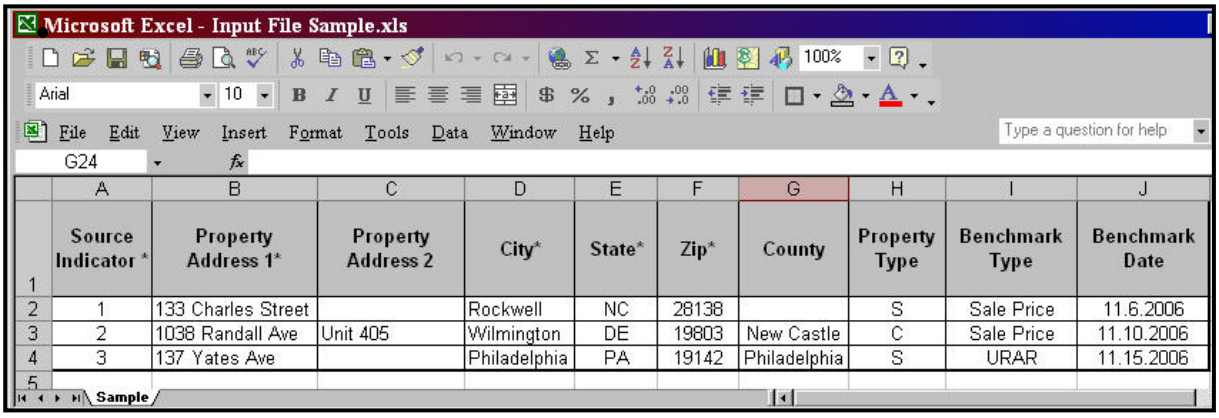
(supplied to the AVM providers)

To perform a comprehensive AVM analysis, provide as many of the following data elements for each loan/property as possible. The more data elements passed, the greater the AVM provider's ability to segment and stratify the results on the AVM user's behalf. The data elements marked with an ***** are required.

Test Sample

- Unique Record Identifier (Record number for matching the results)
- Property Address (not scrubbed)
- Street Address*
- City*
- State*
- Zip*
- County
- Property Type
- Benchmark Type
- If current purchase money transactions are unavailable and appraisals are used for the benchmark values, the appraisal type along with the loan type must be provided.
- Benchmark Date

It is important to remember that AVM providers have specific formats that they use to process test files. AVM users must make certain that their files are comparable. AVM providers currently accept Microsoft Excel files where each Input Filed is listed in a separate column as displayed in the following sample.



	A	B	C	D	E	F	G	H	I	J
1	Source Indicator *	Property Address 1*	Property Address 2	City*	State*	Zip*	County	Property Type	Benchmark Type	Benchmark Date
2	1	133 Charles Street		Rockwell	NC	28138		S	Sale Price	11.6.2006
3	2	1038 Randall Ave	Unit 405	Wilmington	DE	19803	New Castle	C	Sale Price	11.10.2006
4	3	137 Yates Ave		Philadelphia	PA	19142	Philadelphia	S	URAR	11.15.2006
5										

Output Fields

(from the AVM provider)

Each AVM provider involved in the validation testing should be requested to return the input file with the new data appended to each record. The data elements marked with an ***** are required. Unless otherwise prohibited by law, appended data elements should include:

- Unique Record Identifier (Record number for matching the results)
- Standardized property address
- Estimate of property value*
- Confidence Score*
- Low Price
- High Price
- Property Type
- Last previous known or pending transaction date*
- Last previous known or pending transaction amount*
- Last previous known or pending transaction type*
- AVM provider unique fields

Post-Processing Data & Analytical Results

(returned to AVM provider and, if applicable, the AVM user subsequent to processing of test sample)

To reduce any appearance of testing inaccuracy or bias, AVM users or test consultants should provide information back to AVM provider test participants and, if applicable, AVM users that include record level benchmarks, basic findings and summary (blind) conclusions of each test. This should also include the final recommendations (e.g. cascade placements on all granularity levels) being made by the AVM user and/or test consultant. The following list includes the types of information that should be shared with each AVM provider (and, if applicable, the AVM user) in an effort to promote full transparency.

Test Sample Information

- Identification of analysis data set(s)
- Identification of excluded property records and reasons for exclusion

Loan Level Data

- Benchmark values for all properties (analysis set and excluded property records)
- Benchmark source
- Credit group tier
- Customer's estimated value
- LTV/CLTV

- Loan amount
- Loan type (e.g. purchase money, cash-out refi, HELOC)
- Last known or pending transaction date
- Last known or pending transaction amount

Reporting

- Blind aggregate reporting by AVM provider by State and County:
  - Accuracy metrics used by AVM user or third party consultant (see below)
  - Hit rates
- AVM user's cascade position at each cross-section within the cascade (e.g. county or price tier)

Analysis Data Sets

The total universe of all results received from AVM providers is the starting point in creating the various data sets that may be used for analysis. The AVM user or third party consultant should identify subsets of AVM results that are compiled based on: (a) test results that are removed or excluded from the sample; and, (b) test results that are grouped and analyzed according to identified common characteristics such as benchmark type. Inappropriate selection of analysis subsets will introduce significant bias to any AVM validation. As a result, the process employed for determining all analysis subsets must be fully transparent and disclosed to all participants. The following list is not exhaustive and there exists debate as to the pros and cons of each method.

"All-In" Analysis Sets

In this case, no results from any AVM being tested are excluded. This analysis should only be done in support of other, more meaningful analysis of AVM performance results. However, it is an informative view in potentially determining the breadth, currency and depth of an AVM provider's database, all of which are important factors in considering AVM performance in production. Each AVM provider has the opportunity to gather and load data; some do it much better and faster than others. This type of analysis may lead to better decision making when used in support of the global removal of test properties by identifying the AVMs that perform on a consistent relative basis.

Global Removal of Test Properties (Recommended)

If any AVM provider reports having a current sales price for a test sample property, that property is removed from the test sample universe for all AVMs in order to keep the test results on an "apples to apples" basis across all AVMs. As

mentioned earlier, this is the best method of test sample property removal. This method may not be employed by some third party testers or AVM users because it results in reducing the sample size of the test and, quite possibly, added expense in obtaining additional input data. If at least one AVM provider had prior access to the specific benchmark sale transaction before returning the test result (i.e. the final estimate of value) that record should be removed from the test sample set for *all* AVM providers.

Limited Removal of Test Properties

If an AVM provider reports having knowledge of a current sales price for a subject property, it is removed from the test results of that specific AVM provider or AVM model only. This results in an "apples to oranges" comparison that demonstrates poor scientific testing practices. The bias can be compounded when a test property is removed for one AVM model but not another where both AVM models belong to the same AVM provider. This approach also often eliminates the AVM provider that has done the best in accumulating data in an area from being considered for usage. In the absence of extreme circumstances (e.g. one AVM provider has hit the exact benchmark for 30% of the file without having provided last known or pending transaction data), the use of "limited removal" analysis sets is highly discouraged.

INTERPRETING THE RESULTS

There are a number of different views of what constitutes "good" or optimal AVM performance. As CATC has attempted to convey throughout this document, these differences may certainly be appropriate given AVM users' appetite for risk or application of an AVM process or system. The objectives of the AVM user in implementing an AVM system should dictate the metrics to be used in determining performance results for individual AVM models and, in a separate effort, for the development of an AVM cascade. Although CATC does not endorse any one metric over another, set forth below is the identification and discussion of the more commonly used performance metrics.

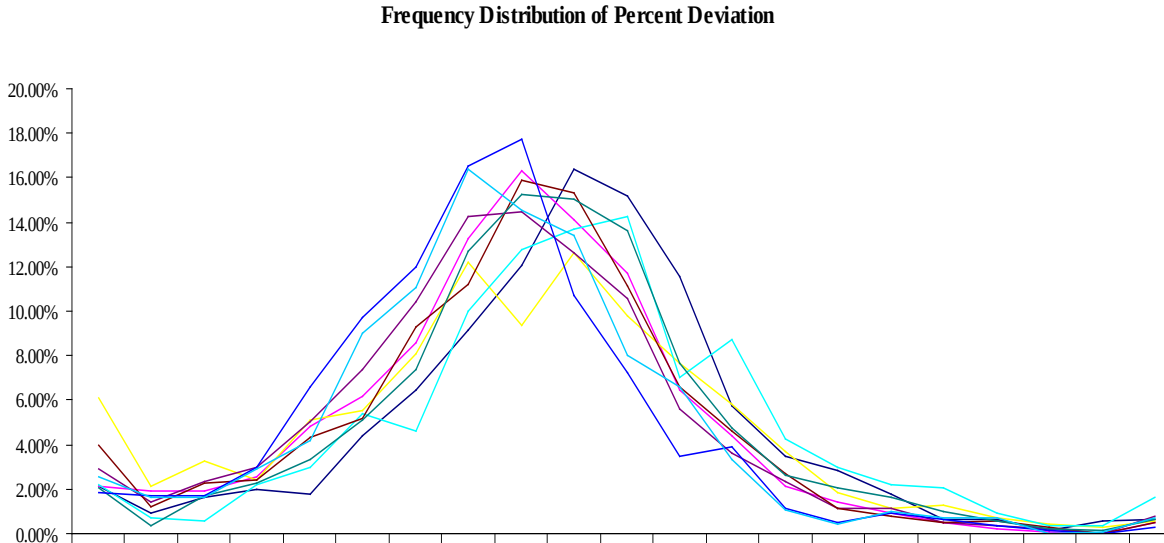
Accuracy

There is no single measure that will indicate overall performance; rather it is a set of measures that will describe the behavior of the data set against expected behavior. CATC would suggest that, at a minimum, these measures are appropriate for consideration in performing this analysis.

Frequency Distribution - % Error

Although this form of analysis is one of the most basic, it can also be the most useful. This statistic is found by subtracting the benchmark value from the AVM value, and dividing the result by the benchmark value. [(AVM Value – Benchmark Value)/Benchmark Value]. The variance of the AVM value to the benchmark value is displayed as a percentage.

The closer this percentage is to zero, the closer the AVM value is to the benchmark value. If the percentage is less than zero, then the AVM value is less than the benchmark value. If the percentage is greater than zero, the AVM value is higher than the benchmark value. This calculation can be used on a number of levels, overall, by state, by county, by price tier, etc.



	-50%	-45%	-40%	-35%	-30%	-25%	-20%	-15%	-10%	-5%	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
AVM 1	2.14%	0.94%	1.61%	2.01%	1.74%	4.42%	6.43%	9.12%	12.06%	16.35%	15.15%	11.53%	5.76%	3.49%	2.82%	1.74%	0.67%	0.67%	0.13%	0.54%	0.67%
AVM 2	2.10%	1.90%	1.90%	2.57%	4.86%	6.19%	8.57%	13.24%	16.29%	14.10%	11.71%	6.48%	4.38%	2.10%	1.43%	0.95%	0.48%	0.19%	0.10%	0.00%	0.48%
AVM 3	6.10%	2.13%	3.26%	2.41%	5.11%	5.53%	8.09%	12.20%	9.36%	12.62%	9.79%	7.66%	5.82%	3.69%	1.84%	1.13%	1.28%	0.71%	0.43%	0.28%	0.57%
AVM 4	2.22%	0.74%	0.56%	2.22%	2.96%	5.37%	4.63%	10.00%	12.78%	13.70%	14.26%	7.04%	8.70%	4.26%	2.96%	2.22%	2.04%	0.93%	0.37%	0.37%	1.67%
AVM 5	2.89%	1.40%	2.33%	2.99%	5.04%	7.38%	10.46%	14.29%	14.47%	12.61%	10.55%	5.60%	3.64%	2.33%	1.12%	1.12%	0.47%	0.37%	0.19%	0.00%	0.75%
AVM 6	4.00%	1.20%	2.30%	2.40%	4.30%	5.21%	9.31%	11.21%	15.92%	15.32%	11.11%	6.61%	4.60%	2.70%	1.10%	0.80%	0.50%	0.60%	0.30%	0.00%	0.50%
AVM 7	2.06%	0.32%	1.73%	2.27%	3.35%	5.09%	7.36%	12.66%	15.26%	15.04%	13.64%	7.68%	4.76%	2.60%	2.06%	1.62%	0.97%	0.54%	0.22%	0.11%	0.65%
AVM 8	1.81%	1.68%	1.68%	2.97%	6.59%	9.69%	12.02%	16.54%	17.70%	10.72%	7.24%	3.49%	3.88%	1.16%	0.52%	0.90%	0.65%	0.39%	0.13%	0.00%	0.26%
AVM 9	2.55%	1.67%	1.67%	2.94%	4.22%	9.02%	11.08%	16.37%	14.51%	13.43%	8.04%	6.57%	3.33%	1.08%	0.39%	0.98%	0.69%	0.69%	0.00%	0.10%	0.69%

Absolute % Error

The Absolute % Error calculates the *magnitude* of an error without regard to whether it is an over prediction or an under prediction (i.e. the "+" or "-" sign is removed). The Absolute % Error Rate between the AVM value and the benchmark value is given by:

Absolute ((AVM Value - Benchmark Value) / Benchmark Value)) if the AVM Value > Benchmark Value

((Benchmark Value - AVM Value) / Benchmark Value)) if the AVM Value < Benchmark Value

Mean (Average) Absolute % Error

To calculate the *Mean* Absolute % Error, the Absolute % Errors are summed and divided by the number of records being summed. The Mean Absolute % Error gives an average error magnitude in the sample.

The smaller the mean absolute % error is, the closer the AVM values are to the benchmark values.

Mean (Average) % Error:

The Mean % Error is the average error rate of the sample. If an AVM tends to over value properties the Mean Error will be positive and if it tends to under value properties the Mean Error will be negative.

Mean Error Rate % = average ((AVM Value - Benchmark Value) / Benchmark Value).

Median Absolute % Error

The Median Absolute % Error is calculated by arranging the Absolute % Errors in order from smallest to largest and then selecting the middle value, or the 50th percentile. The Median Absolute % Error is an accurate indication of an AVM's central tendency without being strongly influenced by extremely large or small % Error outliers.

Median % Error

The Median % Error is calculated by arranging the % Errors in order from smallest to largest and then selecting the middle value, or the 50th percentile. The Median % Error is an accurate indication of an AVM's central tendency without being strongly influenced by extremely large or small % Error outliers.

Standard Deviation of the % Error

The standard deviation is a measure of how widely values are dispersed from the average value (the mean). For example, the Standard Deviation of the % Errors would be given by:

$$\sqrt{\frac{\sum_{i=1}^{n}\left((\%\ Error)_i - Mean\ \%\ Error\right)^2}{n-1}}$$

where n is the sample size.

Confidence Scores

Depending on the application, AVM test results (estimates of property value) can be useful in evaluating confidence scores, a measure of uncertainty of the point estimate of value given by the AVM. This will allow the AVM user to set up business model criteria based on overall accuracy of that estimate of value. One approach to test the reliability of an AVM provider's confidence score is to determine the relationship between confidence scores and the accuracy of underlying predicted market values. For example, the point estimates of value can be stratified by major confidence score segments or "buckets" to determine the Median Absolute % Error for all properties in that specific segment (e.g., 90-100, 80-89, 70-79, H, M, L, etc.). Typically, the AVM % Error becomes larger as the confidence score degrades. Experience has shown that the use of correlation coefficients (r-squared) is not indicative of confidence score reliability.

Model-to-Model Comparisons

When comparing results from different models, an understanding of the unique characteristics of different methods and features of the model being tested is extremely helpful when interpreting the results. Some points to consider:

- How does the model treat the subject's last recorded sale?
- What external value benchmarks are available and appropriate?
- What types of properties are in the sample set?
- What is the average age of the last sale date in the sample set?
- Is address validity and locality represented in the sample set?
- Is the model being tested in a fashion representative of normal business usage?

The following exhibit displays a summary of how the AVMs' results could be compared to one another.

	Sample Size	Category	AVM 1	AVM 2	AVM 3	AVM 4	AVM 5	Cascade
		Hit Rate	56%	60%	62%	54%	67%	86%
	20000	< -20%	10%	7%	4%	10%	21%	13%
		> 20%	10%	7%	20%	15%	5%	10%
Total Sample $50000 - $9200000		+ / - 10%	55%	56%	50%	45%	53%	53%
		+ / - 15%	70%	67%	69%	60%	64%	47%
		+ / - 20%	80%	86%	76%	75%	74%	77%
		Mean ABS Error	16%	15%	17%	19%	21%	19%
		Mean Error	7%	6%	8%	-9%	-18%	1%
		Median Error	1%	0%	3%	-4%	-3%	-1%
		Standard Deviation	17%	16%	19%	23%	26%	18%

Cross Sections of Evaluation & Performance Degradation

Depending on the level of resources, risk tolerance and other issues, analysis of AVM test results should be segmented over any number of categories to determine whether significant performance degradation exists that would affect AVM usage in production.

AVM hit rates, accuracy and confidence score analysis can easily be broken down by price range, property type, and geography (e.g. state, county, FIPS). This is helpful to define if, when and how the AVM will be used. If multiple AVMs are being considered for use in a cascading AVM system, segmentation and analysis by geography should be conducted at a minimum to determine AVM ranking on a zip code, county or state basis. For an analysis at any level of granularity, there should be a statistically significant number of observations to perform a meaningful analysis.

CASCADE DEVELOPMENT & ANALYSIS

Over the past several years sophisticated AVM cascades have been developed and implemented using rules based decision systems that permit an AVM user to rank order in which AVMs are used under specific circumstances (e.g. by product type, geography, price tier, property type, etc.). A discussion of AVM cascade design, development and application could encompass a full whitepaper in its own right and is not the focus of this document. Nevertheless, AVM users must understand the basic relationship of AVM testing and validation to AVM cascade development.

Cascades often contain “bump” logic for ordering the next AVM if the first AVM is a “no hit” or if it does not meet the AVM user acceptance criteria (e.g. confidence score threshold). Accordingly, these valuation cascades process the returned

AVM information against the pre-determined risk criteria of the AVM user. Such risk criteria may include a specific minimum confidence score by AVM that is acceptable for use in underwriting of a specific product type, for specific credit score ranges, by estimated value ranges, and property type.

There are several approaches when building a cascade. One approach uses a rank ordering of the AVMs based on their individual accuracy results from the validation of the test analysis, described above. An alternative is to use an iterative approach where AVMs are tested in different rank orders within the cascade to determine optimal performance of the overall cascade. The objective of these iterative rankings is to determine the appropriate placement of the approved AVMs within the cascade positions by specific geography. It is possible, for example for AVM "A" to have very high accuracy within a given geography, but a limited hit rate. While that AVM may be the most accurate, AVM "B" may have a higher hit rate and very good accuracy and include every property for which AVM "A" reported an estimated value. In that case the AVM user may decide to move AVM "B" into the first position in that geographic area as it would return all the hits of AVM "A" with comparable accuracy, but would also return other value estimates for additional properties. The position of each approved AVM within each geographic area needs to be determined by such analysis, with the results documented at the detail sample property level and in aggregate, then distributed to all parties. These are by no means the best or only approaches to cascade development.

Once the AVM cascade order has been established the cascade service provider needs to implement the cascade decisions into its system. The AVM user should verify that the cascade has been implemented correctly according to the AVM user's criteria. The AVM user needs to conduct periodic tests and audits to determine that the cascade order of the AVMs remains consistent.

POST-TESTING EVALUATION & DUE DILIGENCE

Regardless of the level of due diligence performed when constructing and analyzing AVM tests, it is imperative that results from a test are compared to the AVM user's experience in production. Despite the best efforts to obtain "good records" for testing, some recently closed records or scheduled to close may have already made it into an AVM's database. In these instances, an AVM's performance in a test may overstate that AVM's performance in production. For example, if an "arms length" purchase money transaction that was closed within the last 1-30 days was known to an AVM prior to the test, that AVM will appear to have a very accurate value for that property. However, in production that AVM will not have the luxury of knowing the most recent sale price and the accuracy/performance would not be as good as in a test.

This scenario can be quantified by monitoring the performance of the AVM cascade by comparing production performance against expected performance based upon test results. If actual production performance does not meet the expectations of the test results, further research should be performed to diagnose why the discrepancy exists. The AVM user should review production reports at least monthly to verify the cascade's effectiveness by geography and AVM for hit rate, usable hit rate, and cost per usable AVM.

Summary Analysis should be shared with all AVM provider participants and any "Cascade Recommendation", whether generic or created specifically for a given AVM user should be shared with AVM provider participants in a blind fashion. This blind cascade information should include high level decisions that went into the recommended cascade including precision, price (if applicable), hit rate, etc. In addition, the end user should be encouraged to provide to all AVM provider

CONTRACTUAL AGREEMENTS

All AVM Providers participating in a test should be required to execute an agreement, which states the AVM results returned by a given provider were derived using the exact same methodology that would be used in a real world production environment. For example, no additional steps were taken by the AVM Provider to "enhance" their performance for the test that doesn't exist in their day to day production methodology when returning an AVM Model result.

CATC maintains a Website at www.catconline.com where continuing discussion of this and related topics, sample forms and other materials shall be posted and updated from time to time.

SUMMARY

This paper has set forth many possible aspects of AVM testing and analysis an organization should consider in any evaluation. It would be unrealistic to attempt to incorporate every procedure or analysis in any one test. If nothing else is taken from this paper, an organization must first decide on the goals and expected outcomes of its AVM usage, then design and implement a plan for evaluating multiple AVM providers in an attempt to effectively realize those outcomes. The procedures set forth herein serve as a reference to the critical elements for evaluating and selecting AVM providers best suited to the user's goals. The most standard of the elements for testing an AVM provider direct the validation to:

- Use recent sales price benchmark test data within 1 – 30 days.

- Diversify samples across different transactions (e.g. purchase or refinancing) and segments depending on the AVMs intended use (e.g. property type, price range, loan type or geography).
- Prepare test samples in a manner most representative of a production environment.
- Exclude test data at the extremes of the spectrum
- Require early return of test results back from AVM providers (a maximum of 3-5 days) in order to prevent contamination or manipulation of data.
- Determine correlation of confidence scores to accuracy and the distribution of values across levels of confidence.
- Use either or both an all-in or global removal approach to determining the final sample size and results.
- Calculate "usable" hit rates to determine optimal results.
- Utilize a transparent process that is understood by all participants prior to conducting a test.
- Protect confidential information by securing the proper non-disclosure agreements from all participants, including third party consultants.
- Provide full transparency to all parties in a test (AVM provider, AVM end user and third party consultant) by providing the following information to all in t timely manner:
    - Loan level blind test results for all participants
    - Cascade recommendation and cascade implementation that was used after analyzing the results
    - Identification of specific records that were removed prior to analysis and the reason for removal
    - Blind high level aggregate results

For more information, please contact:
catc@catconline.org

APPENDIX A: Testing Evaluation & Due Diligence Topics

AVM Users should conduct diligence around any service provider, internal or external to the institution, before selecting a provider related to any AVM use, validation or systems support. As referenced throughout the CATC best practices documents, these participants include AVM providers, third-party test consultants, third-party party test data providers, and cascade platform providers. It is important to understand the capabilities of each provider, what that represents, and how it is appropriate to the AVM User and their credit risk policies. The outline below provides a set of topics, which may not be all-inclusive, that should be considered in designing and conducting that diligence review. CATC intends on developing and publishing Model Due Diligence Questionnaires on its website located at www.catconline.com.

AVM Providers

I. Vendor Background
   a. Company
   b. Expertise
   c. Systems
   d. Disaster Recovery

II. Model Development
   a. How developed
   b. By who
   c. Internally or acquired
   d. Continuity

III. Modeling Technique(s)
   a. Index
   b. Hedonic
   c. Other
   d. Hybrid

IV. Data
   a. Sources & Availability
   b. Coverage
   c. Refresh & Update
   d. Data Management & Procedures

V. Accuracy
   a. Expected Error
   b. Confidence Score
      i. Definition

ii. Distribution
c. Internal due diligence experience

VI. Coverage
a. Documentation
b. Reporting of changes
i. Additions
ii. Deletions
c. Identification/Determination

VII. Performance
a. Property types
b. Price tiers
c. Geography
d. Hit Rate
e. Volatility/consistency

3rd Party Test Data Providers and Test Consultants

I. Vendor Background
a. Company & Personnel
b. Expertise
c. Systems
d. Disaster Recovery

II. Data Sources & Relationships

III. Property Records
a. Transaction Types
i. Pending Sales
ii. Escrow
iii. Purchase
iv. Appraisal
b. Property Types
c. Age of records

IV. Record preparation
a. Record format
b. Record layout
c. Address standardization

V. Potential conflicts of interests and established procedures to deal with them
a. Data relationships
b. AVM or AVM Provider relationships

c. Cascade platform relationships

VI. Testing process
   a. Overview
   b. Transparency
   c. Contractual Agreements (where applicable)
      i. AVM user (between the AVM user and third party testing consultant)
      ii. Test data provider
      iii. AVM provider Test and Confidentiality
   d. Sample design
      i. Sample definition
      ii. Bias control
      iii. Statistical significance
      iv. Over sampling
      v. Record inclusion, exclusion criteria
   e. Analysis
      i. What is to be done
      ii. AVM Provider selection
      iii. AVM Provider test agreements
      iv. Analysis
      v. Rationale for analytical model
      vi. Metrics
      vii. AVM Provider feedback
   f. Reporting
      i. Results format
      ii. Scope
      iii. Updates
      iv. Corrections and comments
      v. Disclosure
         1. Results
         2. Benchmarks
      vi. Peer Review

VII. Audit
   a. Internal
   b. Independent

AVM Cascade Service Providers

I. Vendor Background
   a. Company & Personnel
   b. Expertise
   c. Systems
   d. Disaster Recovery

II. Platform design
    a. Functionality
        i. AVMs offered
        ii. AVM Criteria
            1. Geography
            2. Confidence Score
            3. Property type
        iii. Other Criteria
            1. Price tier
            2. Loan type or program, Program
            3. Imputed LTV
            4. Credit Score threshold
            5. Other
    b. Hardware, software
    c. Transparency
    d. Change capability and control
    e. Service levels
    f. Disaster recovery

III. Integration
    a. Method
    b. Support
    c. Process
    d. Input
    e. Output
        i. Line reports
        ii. Full reports
        iii. AVM provider messaging

IV. Potential conflicts of interests and established procedures to deal with them
    a. Relationships with or ownership of AVMs or AVM Providers
    b. Relationships or ownership of third-party test consultants or test data providers

V. Reporting
    a. All AVM Provider results captured and reported transitionally
    b. Cascade results
        i. no hit and reason
        ii. AVM exclusion criteria
        iii. Other exclusion criteria
        iv. Actual versus expected Cascade result
    c. Usage and Volume

VI. Audit
   a. Internal
   b. Independent