

CAT Error Processing

Questions for SROs and FIF Recommendations

October 20, 2015

Prepared by the Financial Information Forum Working Group

Contact: marylou.vonkaenel@fif.com

Questions to SROs

Error Rates

1. Can you please verify that the following scenario would NOT be counted as an error for a CAT Reporter? A CAT report is submitted to the Plan Processor prior to 8 AM on T+1 (where T = business day of reportable event) and an error in that CAT report is identified and fixed via a correction submitted to the Plan Processor before the 8 AM cut-off time on T+1.
2. Can the SROs explain how individual firms will be assessed regarding CAT regulatory reporting?
 - a. How will each firm's compliance threshold be calculated?
 - b. Over what period of time will the calculation be made (e.g., daily, weekly, monthly)?
3. Can the SROs explain how the overall group (e.g., large industry members) will be assessed regarding regulatory reporting?
 - a. If the maximum error rate for large industry members exceeds 5%, what actions will be taken by the SROs?
 - b. Will higher than expected error rates cause the SROs to assess what is causing the higher error rates and look for solutions to reduce the error rate, or to adjust the error rate itself?
4. The maximum error rate and the error correction timeframes do not address the scenario when the Plan Processor encounters an error (e.g., the Plan Processor creates an incorrect linkage between events; the Plan Processor identifies an error which is not an error; the Plan Processor misses an error validation/processing deadline, etc.). This is especially likely during the first year of operations of a new CAT system.
 - a. How do the SROs propose to handle these scenarios? How will the error rate be adjusted?
 - b. Will the error correction deadlines be shifted for CAT Reporters if the Plan Processors misses an error validation deadline?

Error Correction Timeframes

5. The CAT NMS Plan has established that corrected CAT data must be available to regulators by T+5. The CAT NMS Plan also established interim checkpoints (e.g., submission of original CAT Report, identification by Plan Processor of errors in original submission, re-submission of corrected CAT Reports by CAT Reporters) to enable the T+5 objective. The interim checkpoints that have been established give more time to the Plan Processor for data ingestion and error identification and less time to CAT Reporters for CAT report submission and error correction.
 - a. Was one of the criteria for evaluating the CAT bidders the overall efficiency in data ingestion and validation, including reduced time periods needed by the bidder for these activities?
 - b. Can the interim checkpoints remain flexible in the CAT NMS Plan until after Plan Processor selection, so that additional time could be allocated to CAT Reporter processing, if the selected Plan Processor does not need as much time for data ingestion and validation, and can allow more time for correction while still maintaining the T+5 objective?

Error Validation of Customer Data

6. Clarification is needed on CAT Reporter responsibility regarding the accuracy of customer information provided by a customer to a broker/dealer.
 - a. If contradictory information is provided by a customer to two broker/dealers, how will the Plan Processor determine which broker dealer has incorrect customer data and is therefore responsible for correcting the data?
 - b. If the customer is not responsive or timely in providing corrected data to the broker/dealer, will the broker/dealer be penalized if the error correction timeframe is exceeded?
 - c. If inaccurate customer data is not central to the regulatory responsibilities of market reconstruction and market surveillance (e.g., incorrect zip code), will the CAT Reporter be penalized if the customer does not provide corrected data to the broker/dealer within the error correction timeframe?
 - d. What is the required frequency that a broker/dealer must verify these important customer information data points?
7. Do the SROs have set parameters for the Plan Processor on what is deemed a “material inconsistency”¹ or “reason identified as the source of the issue” that would need to be communicated to CAT Reporters as an error in customer data, beyond the examples² provided in the Plan?

CAT Reporter Report Cards

8. Can the SROs provide additional description of the types of report cards (for CAT Reporter performance) that will be supported?

FIF Recommendations

FIF understands and respects the regulatory requirement to have accurate data available to the regulators as soon as possible to fulfill its market surveillance and market reconstruction responsibilities. And, in fact, all data will be available to regulators as soon as ingested by the Plan Processor, whether correct or not. The goal of all CAT Reporters is to submit accurate reporting of its customer and market transactions on initial submission to the CAT, because reporting errors are costly and time consuming to correct. It is in the CAT Reporters’ self-interest to have as low a CAT error reporting rate as possible. In addition, CAT Reporters have the added incentive to drive quickly to a low error rate, because OATS retirement is dependent on the accuracy and reliability of CAT reporting.

¹ CAT NMS Plan, Appendix D-35, “Material inconsistencies such as two different people with the same SSN must be communicated to the submitting CAT Reporters and resolved within the established error correction timeframe as detailed in Section 8.”

² CAT NMS Plan, Appendix D-35, “Reason identified as the source of the issue, such as: duplicate SSN, significantly different Name; duplicate SSN, different DOB; discrepancies in LTID; or others as determined by the Plan Processor”

FIF has consistently voiced its members' concerns that the CAT NMS Plan's stated error rates and error correction timeframes are excessively aggressive, given the current experience with OATS:

- CAT represents a more complex regulatory reporting system than OATS, with the addition of options and other asset classes, customer information and new reporting events. It will be a newly developed system, created on an aggressive development and test schedule. This normally would indicate a higher than normal initial error rate for the CAT system itself, as well as reporting by the CAT Reporters.
- The current OATS error correction timeframe permits an error correction cycle 2 days longer than the proposed CAT error correction timeframe. The CAT NMS Plan has cited nothing in its proposed support structure that can explain how CAT Reporters can meet these error correction cycles.
- The CAT NMS Plan does not include any description of reporting, testing or validation tools that will be made available to the CAT Reporter that might justify an improved reporting experience.
- The CAT NMS Plan specifically excludes CAT Reporter access to its reported data, in bulk data access format, that would facilitate error validation and correction.

FIF is concerned that the stated aggressive error rates and error correction timeframes are not allowing a reasonable CAT system introduction whereby both the Plan Processor and CAT Reporters can learn and adapt to the new system in a collaborative environment, with a focus on rapid improvement and performance rather than on penalties and failure.

The following FIF recommendations related to error rates and corrections were previously included in FIF's response to the CAT NMS Plan³ and Optimal Solution Recommendations⁴:

- The initial error rates and error correction timeframes are targets for which penalties should not be assessed until it has been demonstrated that the CAT system, and the CAT Reporters can reasonably achieve the targets.
- If CAT Reporters achieve reasonably low error rates, their obligation for OATS reporting should be suspended until OATS can be retired.
- CAT Reporters should have bulk data access to their reported data, in downloadable format, to allow ease of error validation and resubmission of corrected data.
- The Plan Processor should provide adequate access to secure testing environments and sophisticated testing tools, especially to support processes associated with linkage testing, to assist CAT Reporters in accurate submission of CAT Reports and rapid correction of errors.

Additional recommendations, not included in previous FIF recommendations, follow.

Customer Data

1. FIF recommends that customer information fields be categorized by the Plan Processor and the SROs in degree of importance for market surveillance and market reconstruction, so that focus can be concentrated on ensuring accuracy of the most important fields from a compliance viewpoint. Different criteria could be established based on the customer data categorization for

³ FIF Response to CAT NMS Plan, November 19, 2014

⁴ FIF CAT WG – CAT Processor Proposed Optimal Solution Recommendation, September 15, 2014

correction turn-around time; e.g., customer unique identifier (LTID or social security number) would be of highest priority and must be accurate. Any errors in unique customer identifier must be corrected within a 24 to 48 hour time period. Zip code may be of lesser importance and not as critical; an error with zip codes could be corrected within a 30 day period and not impact regulators' ability to surveil the marketplace.

2. The CAT NMS Plan stipulates that the Plan Processor must communicate CAT customer errors to CAT Reporters by 5 PM on T+1. This may require staffing adjustments by the CAT Reporters, having some compliance, operations and IT staff work 2nd shift to analyze/correct the errors. While the CAT Reporters may then be staffed, any corrections requiring client input may not be possible until the next day, at the earliest. Practically speaking, a 5 PM, T+1 checkpoint is equivalent to a T+2, 8 AM checkpoint from a CAT Reporter's perspective. Also, communicating errors in CAT transactions and customer information errors to CAT Reporters on different timeframes (T+1, noon and T+1, 5 PM respectively) is a disadvantage for CAT Reporters because there may be linkages between transaction and customer information errors. The error correction window is already too narrow; this further shortens the timeframe for correcting linked transaction and customer information errors. FIF recommends that the Plan Processor identify errors with customer information data by T+1, noon to coincide with the deadline indicated in the CAT NMS Plan for the Plan Processor to identify errors in transaction reports.

Error Rates

3. Currently, the CAT NMS Plan defines two groups, large industry members and small industry members, for which error rates will be calculated. FIF members are in favor of this categorization for purposes of assessing the CAT error rate; however, more granular analysis is needed for determining trends and identifying the root cause of reporting problems, whether they are CAT Reporters' issues or due to the Plan Processor. For example: additional categories defined for the purpose of trend analysis, such as by asset class, business model, desk type or number of desks, distribution of event types or destination, or parameters within the CAT report, will help identify where reporting problems tend to arise. This should enable the Plan Processor to better determine if there are patterns of errors that can be easily or quickly corrected. Reporting of this trend analysis data to CAT Reporters would also assist them in better understanding and improving their reporting.
4. There is a current body of evidence and experience with the processing of equity order reporting. There is little or no experience with regulatory reporting of customer and options data. For the first six months of CAT regulatory reporting, it is recommended that the CAT maximum error rate be applied only to equity order reporting. Data on customer and options error rates also should be collected during the first six months of CAT reporting, to provide a basis upon which the SROs may later establish an appropriate maximum error rate for these classes of data as well.
5. Error rates should be calculated on post-correction (T+3) not pre-correction (T+1) data. Post-correction error rates should be used for calculations when determining eligibility for OATS retirement.

Support for Error Processing:

6. An OATS-like system should be made available to CAT Reporters to facilitate sorting by error code through a CAT portal and permit CAT Reporters to correct the erroneous fields, as opposed to re-submitting the CAT entire set of reports.
7. Additionally, data extracted from the CAT data base by the Plan Processor and provided to the CAT Reporters should be in a downloadable format for easy correction/modification by a CAT Reporter and allow easy upload to provide corrections to the Plan Processor.

Penalties

8. To ensure the Plan Processor is stable and accurately processing reported data, FIF recommends that error rate reporting commences from day one of CAT reporting but that enforcement of penalties begin twelve months following the start of CAT reporting, coincident with small industry member reporting. This provides a reasonable balance between a stabilization period for both Plan Processor and CAT Reporters to optimize their reporting systems and encourages rapid improvement of CAT reporting to enable as early retirement of OATS as possible. It also puts large and small CAT reporters on equal footing in regards to penalty processing.
9. FIF recommends that penalties should only be assessed by one system (either CAT or OATS) during the period of dual reporting to both CAT and OATS.

Test Tools

10. To assist CAT Reporters meeting the required error rate, a set of test tools are needed that would allow CAT Reporters to submit CAT reports for testing/validation prior to submission to the Plan Processor. The validation should include not just syntactic and semantic validation, but also validation of the linkage logic of the CAT reports. Although linkage logic validation is much more complicated, it would have the most value to CAT Reporters in ensuring their CAT report submissions are accurate. Use of T-1 production data or use of a subset of volunteer industry firms submitting reports so that matching logic could be tested are two possible methods to enable linkage logic validation.

Following are a few CAT Reporter test scenarios and functionality that could be provided by the Plan Processor to assist CAT Reporters in reporting accurately within the aggressive timeframes:

- a. In the case of a firm receiving a large number of rejects, the firm will determine and correct the root cause of the errors, possibly creating new code or rolling back to a previous version. As the next step, the firm would submit corrected reports initially in a CAT test environment. The Plan Processor could run the submission against yesterday's data for linkages and/or mismatches. If this type of testing support is offered on-demand, the CAT Reporter will be able to verify its changes in response to the errors quickly and without impact to other CAT Reports. If the new submissions in the test environment are successful, the CAT Reporter could immediately, and in bulk, make its corrections. This ability will significantly decrease the time for a large set of corrections to be made. Additionally, it will increase the accuracy of the corrections. Otherwise, if the error processing was similar to the current OATS processing, the CAT Reporter would need to submit just a small set of corrections in production to verify that the

change has had the desired effect. Then the CAT Reporter would have to wait a day (or two) to have the small set of records processed; and if successful, submit the majority of the corrections on the following day(s).

- b. Another scenario would be to test a firm's large internal release of its CAT Reporting software. In this case, a firm might want to submit yesterday's reports created with the firm's new candidate release, in a test environment. The firm would request the Plan Processor to run these new reports against yesterday's data for linkages and/or mismatches. This provides an excellent, additional layer of validation to help eliminate any production errors in reporting, thus driving down error rates.
- c. A third scenario would be the case of a new Plan Processor release being rolled out. It is inefficient and costly to wait for industry-wide testing for the opportunity to have linkages and mismatches identified. It would be ideal to have a test environment in which a subset of CAT Reporters could submit newly formatted records. Then, the Plan Processor could run linkages on this subset of data, reporting any errors. This provides an excellent, additional layer of validation to help eliminate any production errors in reporting, earlier in the test cycle, thus driving down error rates.
- d. Another variation of test functionality is that the Plan Processor could essentially create an internal "shim" such that a CAT Reporter could report yesterday's data with the new format and run this against "mocked-up" production records from yesterday. Then the Plan Processor could run linkages in a meaningful way on the new format of the participating CAT Reporter. This would provide flexibility for each CAT Reporter to do testing repeatedly, on-demand.

Customer Support

11. The CAT NMS Plan specifies that the expected call volume is 2500 calls/month. Assuming 2000 broker/dealers,⁵ this would translate to one call/month/CAT Reporter and Data Submitter. Even in "steady state" mode, one call/month/CAT Reporter and Data Submitter seems "light", but during the test period and first six months of the newly introduced CAT system, 2500 calls/month is very inadequate. A more realistic call volume for the introductory phase of the CAT system would be, at a minimum, two-three calls/week/CAT Reporter and Data Submitter during the test period and then one call/week during first 6 months of production. This would translate to a call volume of 20,000 calls/month during the test phase and 8,000 calls/month during the early production phase. It is important to get this estimate correct because the cost differences/staffing requirements based on this estimate is significant. It would be difficult for any bidder to "contain" this cost difference, meaning that customer service would be severely compromised and curtailed during the CAT system introductory phase. This will make it even harder for CAT Reporters to meet target error rate and correction timeframes, without the assistance of timely customer service.

⁵ CAT Request for Proposal, January 30, 2014. "Note: While there are approximately 5,000 broker-dealers, the anticipated number of broker-dealers that will be engaging in CAT-reportable activity (i.e., trading in NMS securities) is approximately 2,000. The SROs anticipate that some broker-dealers will not directly report to the CAT but will rely on other organizations to report on their behalf. However, the CAT will need to have the flexibility to adapt on a timely basis to changes in the number of entities that report information". RFP also included data sources of 25 exchanges, 2 TRFs, 1 ADF, 3 SIPs/OPRA, 1 OCC.

12. Using the model similar to OATS today, a CAT Reporter who is experiencing a systemic issue that will require extensive corrections and testing, and therefore will not make the error correction window, may submit an “incident” report to CAT and receive a case number, so it can be tracked and referenced when calculating/analyzing a firm’s Compliance Threshold. FIF recommends that CAT institute a comprehensive incident/problem/question reporting and tracking system. This tracking system can be used by both the CAT Reporters and the Plan Processor to record status updates on the problem, and track progress through to closure. Reporting functionality should include the problem tracking system for use by both the Plan Processor and CAT Reporters to view all open and closed problems (CAT Reporters can only see their problems, of course).