ANSWERMINE
WE LET THE DATA SPEAK

Introduction

AnswerMine Group LLC ("AMG") submits the following comments in response to the comment request published in the Federal Register titled:

> FEDERAL RESERVE SYSTEM
> 12 CFR Part 225
> Capital Plans; Proposed Agency
> Information Collection Activities:
> Comment Request

Specifically, AMG will address the collection of data required by FR Y-14Q defined as granular data covering target BHCs' various asset classes within the Retail Risk Schedule for the reporting period. This data would be used to support supervisory stress test models and for continuous monitoring efforts, on a quarterly basis.

AMG Background

AMG is an Illinois based data-mining and analytical consulting company specializing in the discovery of subtle patterns in large data sets which can provide the basis for predicting future outcomes. It has 15 years of experience in appending individual credit bureau data to financial institution loan level data to uncover patterns in borrower behavior under varying scenarios. The company has developed the capability to process large data sets quickly at this granular level (for example, monthly and on a national scale) and provide a range of analytical outcomes for both tactical and strategic portfolio risk management strategies. This capability provides us with the ability to use granular record level data to expose the interaction of both weak and strong variables which dictate borrower reaction and, in turn, loan performance. This "bottom up" approach provides significant improvement over the limitations incumbent to a "top down" traditional approach which requires categories and segments to be set based on the analyst's bias and prior to comprehensive understanding of the issue being addressed. It is our experience that model performance is improved significantly working with individual credit data and loan level records.

[redacted]

ANSWERMINE
WE LET THE DATA SPEAK

****

ISSUE:

"Whether the proposed collection of information is necessary for the proper performance of the Federal Reserve's functions; including whether the information has practical utility"

COMMENT:

It is AMG's experience and opinion that information outlined in the sample exhibits provided, while providing a sound starting point for analysis, is insufficient for the performance of the Federal Reserve's functions as we understand them. First, the arbitrary segmentation of the data rolling up "like assets" into buckets will obfuscate the real drivers of performance both in a static scenario as well as an alternative "stress" scenario. By using segment data, the "consumer" will remain opaque to supervisors. In today's environment, where asset values do not provide the same level of support in retail transactions, the loan performance risk lies in the individual borrower. Without the ability to analyze the data at the individual level, and in turn aggregating individual loan performance expectations, the function of evaluating capital adequacy is unnecessarily limited. Secondly, the reliance on FICO scores as a proxy for an individual's credit evaluation has proven to be woefully inadequate throughout this current financial crisis, both at the outset and now ongoing. On the other hand, individual credit data (as opposed to single score) is very predictive of consumer behavior. By basing the data collection at the individual loan level, in combination with large data analytical engines, the FED will be able to more accurately capture the actual loan performance and better quantify the capital adequacy underpinning a target portfolio. The ability to collect and analyze the data at the individual level significantly enhances the utility of the data accumulated. At the loan level, the supervisor can adapt its focus as economic and business scenarios change without having to disrupt the regulated entity. By developing this ability the regulator can provide the oversight necessary to gain a forward looking analysis without imposing any undue burden on the market.

[redacted]

****

ISSUE:

"The accuracy of the Federal Reserve's estimate of the burden of the proposed information collection, including the validity of the methodology and assumptions used"

COMMENT:

While the burden of proposed information collection will invariably differ from BHC to BHC, there is one common aspect for every institution. They have to aggregate the individual loan data to meet the segmentation requirements indicated in the examples. Not only is individual loan data simpler to produce, but any aggregation creates an additional burden on the individual institutions. It also creates an unnecessary limitation on the future utility of the data received by the supervisor. If well formatted loan level data is received, the ultimate burden to the respondent will be reduced in the event a different "look" is deemed necessary by the supervisor or analyst. The effort should be put into making the loan level data formatting consistent across the target BHC's. Once this is accomplished the data collection process would be optimized. Additionally, each bank has contracts with the credit bureaus for credit data other than FICO scores. Individual level data can be acquired under existing master service agreements, thereby reducing reliance on the FICO scores which may not support the assumptions

****

ISSUE:

"Ways to enhance the quality, utility, and clarity of the information to be collected"

COMMENT:

Currently, there is high availability of data, low cost of storage and low cost /high speed computing for analytics. Restricting data collection to arbitrary segments made up of individual records limits the potential interpretive value of the information. By limiting the input data in this way, using segments established by traditional industry bias unnecessarily limits the utility of the analysis, especially in the situation represented by this current financial crisis. The bottom line is that rolled up data (ostensibly for the purpose of ease of data management or to use yesterday's statistical tools) will reduce the effectiveness of the current analysis by reducing the ability to consider nuanced changes in the behavior of the borrowers. The aggregation of these changes will ultimately result in significant impact on an institution's overall current capital adequacy analysis and the anticipated capital needs in alternative stressed situations. Having data aggregated to meet today's needs will likely mean that the supervisor will not have the right data to anticipate

[redacted]

tomorrow's needs. Having loan level data allows the supervisor the flexibility to meet tomorrow's needs.

****

<u>ISSUE:</u>

"Ways to minimize the burden of information collection on respondents, including through the use of automated collection techniques or other forms of information technology"

COMMENT:

Information collection and ultimate utility must accommodate the needs of legacy systems in the target BHC community. It must be "light" on the existing IT resources to limit the impact of periodically providing usable data from the respondents. Complicating this issue is the need to accumulate data in a consistent format across the respondents so the supervisory authorities can get a systemic view of the data set and run alternative scenarios widely across regions and markets. This will be critical to meeting any meaning forward looking objectives of the new regulations.

Two approaches can be used to achieve these objectives: require the respondents to provide the data in a designated flat file format or accept an output file from a "standard" IT platform which can be transformed into a consistent usable flat file for further analysis. The second approach both reduces the burden on the respondents and puts the quality of the data in the purview of the analyst. Data control on input is critical to getting good results especially when manipulating large data sets from disparate sources. The use of granular credit data is not difficult because the BHC's all have data transfer relationships with the credit bureaus and protocols both in and out in their normal business. Piggy-backing on these contracts should not be operationally difficult or financially burdensome.

Needless to say the use of existing communication protocols and the use of generic file formats (txt, csv) which can be easily interpreted by inexpensive yet powerful computer applications are both available and recommended.

[redacted]

ANSWERMINE
WE LET THE DATA SPEAK

****

ISSUE:

"The effects on burden should the Federal Reserve decide to move from collecting segment-level data to collecting loan-level data for a select number of Retail Risk portfolios."

COMMENT:

Once a data collection template is decided upon most BHC's will have to spend resources to meet these requirement. Most likely this will require reversion back to loan-level data to aggregate up to the required segment level format. This process may or may not be difficult depending on the flexibility and capability of the institution's legacy systems architecture. Requiring loan-level data (in a ubiquitous flat file format) would certainly reduce the IT burden on the respondent. Using existing data transmission processes and tested FCRA compliant procedures the respondent will not be required to establish a FED specific task to meet unique requirements. Certain considerations may need to be made to accommodate file sizes and security issues. Given the current capabilities of computing power, cost effective storage and industry accepted data manipulation tools the FED analysts or third party contractors can efficiently adapt this record level data to provide both input in approved capital adequacy analyses such as stress test projects as well as to provide ad hoc analyses to address issues yet to be determined as more detailed supervision challenges evolve. This approach will permit the use of the data set by other analysts using other software applications and mathematical approaches other than regression from which significant benefits can be gained exploiting large data sets.

Respectfully submitted

S Cornew
[redacted]

E C Buchanan
[redacted]

[redacted]