Moody's ANALYTICS

WHITEPAPER

FEBRUARY 2016

Author

Martin Elliot
martin.elliot@moodys.com

Contact Us

Americas
+1.212.553.165
clientservices@moodys.com

Europe
+44.20.7772.5454
clientservices.emea@moodys.com

Asia (Excluding Japan)
+852.2916.1121
clientservices.asia@moodys.com

Japan
+81.3.5408.4100
clientservices.japan@moodys.com

Making Proxy Functions Work in Practice

Summary

The purpose of this paper is to explore many of the practical issues which can be encountered when developing and implementing a process to generate proxy functions using either the Curve Fitting or Least Squares Monte Carlo (LSMC) techniques.

The paper reviews the stages involved in proxy generation, and identifies the challenges in implementing them, as part of a robust and integrated business as usual (BAU) process. It then demonstrates methods which can be used to overcome these challenges. In particular, this paper highlights the importance of process automation and rigorous validation to ensure that the outputs produced at each stage are fit for purpose.

Finally, the paper addresses the importance of performance in both the scenario generation and function fitting stages of the proxy generation process and identifies a number of methods which can be used to optimize performance in a production environment.

Moody's

Introduction

Recent years have seen growth in the popularity of proxy modeling, with insurers increasingly looking to proxy techniques as a method which allows them to address several important business challenges, including:

» Modeling the full distribution of their balance sheet for use in risk-based capital calculations.

» Fast revaluation of their balance sheet or capital requirement for use in continuous solvency monitoring or hedge revaluation.

» Providing a forward-looking view of risk and capital for use in an Own Risk and Solvency Assessment.

Growth in popularity means that proxy modeling theory has been widely publicized. However, when insurers implement a method for generating proxy functions, there are many practical challenges which they can encounter at the different stages of the process. In this paper, we identify these challenges and show some methods which can be used to address them.

Stages in the proxy generation process

The end-to-end process used to produce proxy functions using Monte Carlo techniques consists of several main stages. For illustration in this paper the process is split into the following four stages:

Fitting Scenario Generation

The purpose of this stage is to generate fitting scenarios that are valued by the Asset Liability Modeling (ALM) cash flow model to generate the asset and liability values which are regressed against to produce the proxy function. These fitting scenarios contain stresses to risk factors which are believed to affect the asset or liability being modeled.

An important part of this process for either curve fitting or Least Squares Monte Carlo (LSMC)[1] involves identifying the risk factors which affect the value of the asset or liability for which a proxy function is required. After the risk factors are identified, the range over which they are to be stressed must be determined. This results in the creation of a multi-dimensional real-world stress space, where each dimension represents a different risk factor.

Each point in this stress space simultaneously stresses multiple risk factors. If the liabilities being modeled include options or guarantees, then these scenarios are produced using a market-consistent economic scenario generator (ESG). Generation of fitting points within this space is a 2-step process:

1. Generating real-world "outer" fitting points within the stress space.

2. Generating associated market-consistent "inner" scenarios.

The real-world "outer" fitting points can either be user-defined or automated using algorithms, for example a quasi-random number generator such as Sobol. Liabilities with options or guarantees are typically valued using stochastic market-consistent scenarios. Therefore, for each real-world "outer" fitting point there is also a requirement to generate an associated set of market-consistent "inner" scenarios, and here lies the

---

1 For more information on the background to these techniques please refer to Appendix A

difference between curve fitting and LSMC approaches. Curve fitting uses a small set of "accurate" fitting points, each valued using many market-consistent scenarios, whereas LSMC uses many "inaccurate" fitting points, each valued using a small number, typically 1 or 2, of market-consistent scenarios as illustrated in figure 5.

Scenario Valuation

The purpose of this stage is to value the fitting scenarios within the existing ALM system to give us the value of the asset or liability being modeled under each real-world scenario. These values become the dependent variable values in the regression which is carried out in the subsequent fitting stage. The corresponding set of simultaneous stresses to each of the risk factors complete the regression space by becoming the independent variables used in the function fitting process.

Function Fitting

The purpose of this stage is to generate a mathematical proxy function describing the asset or liability value being modeled in terms of the corresponding stresses to the risk factors. This process uses the ALM values and risk factor stresses as inputs to the regression process and uses a statistical regression method to find the functional form and coefficients of the proxy. After it is generated, this function can be used to revalue the asset or liability given changes to the risk factors it contains.

Validation

The purpose of this stage is to validate the proxy functions generated by the previous stage to ensure that they are an accurate representation of the asset or liability being modeled. The validation process uses metrics and charts to assess the quality of fit and appropriateness of the proxy function produced by the regression process.

Fitting Scenario Generation Challenges

During the scenario generation process, insurers sometimes find it challenging to recalibrate their ESG for real-world stresses, and to automate and validate their scenario generation process.

RECALIBRATING ESG FOR REAL-WORLD STRESSES

As discussed in the previous section, if an insurer is using stochastic simulation to value products which contain options or guarantees, it must be able to generate market-consistent scenarios to value each of the real-world fitting points. The ideal process requires the recalibration of the market-consistent ESG for each stressed real-world scenario. However due to time and technology constraints, such a process is not usually possible.

An approximation of the effect of recalibrating under the stress is needed. There are several ways to achieve this recalibration:

» Output stressing.

» Analytical stress functions.

» Calibration and interpolation.

Applying stresses to the outputs from a base ESG run is the easiest method, but it ignores any structure within the ESG and results in inconsistent outputs which do not correctly capture the ESG model dynamics.

This can lead to potential bias in the scenarios or failure to capture important behaviors and correlations. Output stressing is fine for applying an equity index stress, but there is limited ability to correctly capture the effect of risk drivers such as yield curve level, as under non-lognormal models, implied volatility changes significantly if you shift the yield curve.

A more advanced approach is to use analytical stress functions. This approach overcomes many of the limitations of output stressing as it stresses the ESG parameters directly before they are run. This makes use of any structural model within the ESG to ensure that the outputs remain consistent. However, these methods might not provide enough flexible control over the stressed at-the-money swaption or equity implied volatility surface.

One possible solution is a calibration and interpolation approach to stress implied volatility surfaces. Using this method, the stress function has 2 parts – an initialization phase which is carried out before each of the stresses is generated and the stressing phase which is carried out for each real-world fitting or validation point.

The initialization phase starts by generating a coarse grid of points. Equities require a one dimensional array as we only want to vary the volatility. Swaption IV requires a two or three dimensional array as we want at least one component representing the nominal yield curve (NYC) level, usually the first yield curve principal component (and second if necessary), and volatility. The points on the grid represent stresses to each of these risk factors and are generated by distributing points evenly between the minimum and maximum values specified for the relevant risk factor stresses.

For each point on the coarse grid the relevant stresses are applied to the yield curve level and implied volatility surface. A calibration tool is then used to recalibrate the ESG for that point on the coarse grid based on the resulting stressed volatilities and, if applicable, NYC levels.

As a volatility shock does not have a uniform effect across the volatility surface, we want to be able to apply any shock non-uniformly. Because we want to represent the shock to volatilities as a single risk factor value, we choose a suitable point on the volatility surface, for example the 5x5 point. We then use the shock to this point as the risk factor, with non-uniformity being achieved by specifying a set of scaling factors, which specify the relative effect of the stress at other points on the surface.

The second phase, is then applying each real-world stress to the ESG. Using linear interpolation between the nearest points on the coarse grid to calculate the corresponding values of the required ESG parameters from the calibration tool for that stress. As the interpolation process is quicker than the full calibration process, the 10s of thousands of fitting scenarios can all be generated in a fast, robust, and accurate way with more flexibility compared to the analytic stress functions.

SCENARIO AUTOMATION

Generating fitting scenarios involves configuring and running many individual ESG runs. The steps involved include:

» Generating the real-world fitting points.
» Recalibrating the model parameters in the base simulation to reflect the real-world stress.
» Running the stressed ESG.
» Collating and processing the output data.

As a result, automation using technology is an enabler to carry out these requirements. Managing fitting scenario generation across multiple portfolios of assets and liabilities involves significant data manipulation and the running of complex processes. In addition, there can be a range of stakeholders and users across different functional areas (for example, Group vs Business Unit), geographies and so on. Thus, robust enterprise technology is increasingly required to meet the demands of insurers to manage these processes more efficiently, maintain appropriate audit trails and help to reduce operational risk.

Real-world fitting points can be auto-generated using a low-discrepancy generator such as a Sobol generator to fill a specified multi-dimensional space.

An integrated ESG, combined with analytical stress functions, enables generation of the inner market-consistent fitting scenarios associated with each real world fitting point in a format suitable for your ALM engines within an automated process, reducing manual effort and operator error.

SCENARIO VALIDATION

After the fitting scenarios have been generated, it is important to be able to validate the scenarios produced to ensure that they are fit for purpose and that no errors have occurred in the generation process. Standard tests used for validation of market-consistent scenarios, such as martingale tests or implied volatility tests, are not feasible when the number of market-consistent trials is small due to significant sampling error. However, it is possible to use the analysis test results in other ways to help with validating the fitting scenarios.

Two methods which can be used are:

» Superset martingale tests.
» Asset pricing tests.

The superset martingale tests assess the risk neutrality of the inner simulations by treating all the scenarios as a whole. The estimated prices and standard errors for each fitting point are averaged to produce an average martingale performance – which can then be presented in the same way as a standard martingale test from a standard ESG run.

Figure 1: Superset Martingale Tests

EUR.Zero Coupon Bond Martingale Test

— Estimated Price — Expected Price — Error (95th Percentile)
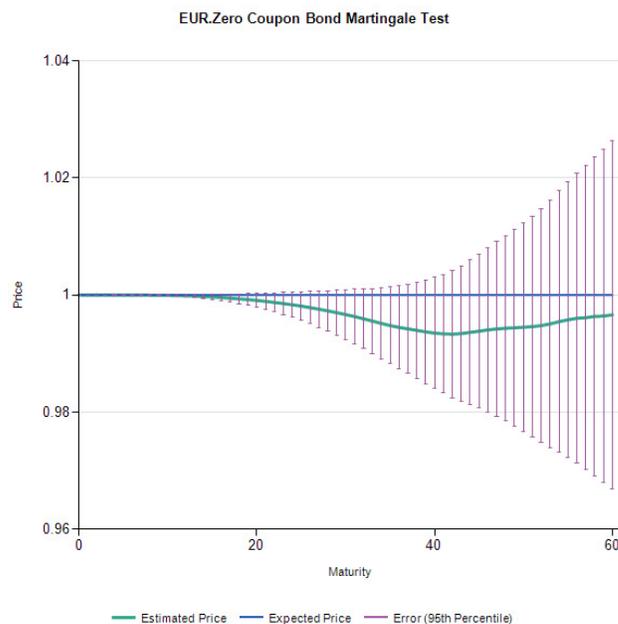
This test can identify systematic bias in the scenario set but the behavior might be different in localized regions of the stress sample space, that is particular stresses might create some violation of risk neutrality. These violations are cancelled out if you look at the entire superset.

Asset pricing tests provide further validation, not just looking at the scenarios, on average, but looking at the results of the analysis tests produced by the ESG across the stress range. This is achieved by making use of the regression functionality used in the function fitting process to fit proxy functions to the prices of simple instruments, such as swaption prices, equity option prices and zero coupon bond prices, calculated by the ESG under each stress and output in the analysis tests. For each asset pricing test, a maturity is specified for a selected analysis test. The prices calculated by the analysis test for the specified maturity, for example price of an equity call option with 10-year maturity, under each fitting and validation scenario are used as the dependent variables in a function fitting proxy, using appropriate risk factors.

The resulting proxy function for the asset pricing test describe the prices for the assets in terms of the shocks to selected risk factors. If stress functions are being applied inconsistently or any discrepancy exists between the modeling of fitting and validation scenarios, it appears as a bias in the proxy function fits. Furthermore, the importance of variance and drift in both the fitting and the validation scenarios is estimated by the quality of fit achieved for these simple assets.

As you fit a function across the range, the asset pricing test functionality can give information about scenario validation in localized regions of the stress range. The test can serve as a valuable check of the scenarios before they are valued and used to fit proxy functions. There are still limitations to this functionality, in particular the use of fitted functions means that the sample error present in each observation is averaged out.

However, carrying out checks using tests like superset martingale and asset pricing tests can provide early warning of any issues or errors which might have occurred. The early warning allows you to rectify them before carrying out the subsequent steps – hence reducing the risk of errors passing through the process and requiring costly reruns of ALM models or regression algorithms.

Scenario Valuation Challenges

INPUT FORMAT

As fitting scenarios need to be uploaded to the ALM system, work might be required to configure the scenarios in the correct format, concatenating the individual ESG results and possibly batching them into smaller sets if required by the ALM system.

The scenario generation process can be updated to automate these post-processing modifications

» Concatenating the outputs that are generated for all the individual stresses.
» Updating the numbering of trials to ensure that the concatenated files contain sequentially numbered trials.
» Allowing multiple output formats and batching of outputs.

Taking these steps ensures that the resulting processed output files are in a format that can be readily imported into ALM systems, reducing the operational risk associated with manual post-processing.

INTERFACE BETWEEN ALM AND REGRESSION PROCESS

The ALM process commonly produces results under each of the market-consistent scenarios. However, the regression space used by the subsequent function fitting process require the results for each real-world

outer scenario, which are obtained by averaging the results for the associated market-consistent scenarios. Updating the regression process so that it can accept ALM results for every market-consistent scenario, and automating the calculation of the result for each real-world stress for use in the regression process, reduces operational risk.

Function Fitting

CREATING CANDIDATE TERMS

The set of candidate terms from which the proxy function is selected is either a predefined set of terms or created using a combinatorics algorithm. The latter option is preferred, as it allows the process to remain flexible and ensure that all models are considered, especially when there are changes in the materiality of individual risk factors. A further benefit is the reduction of the user effort that is required to construct and update the candidate term set manually, as risk factors change.

A combinatorics routine uses a set of polynomial order limits to construct the table of all possible terms. Extra flexibility to give greater or lesser emphasis to particular risk factors can be included by adding functionality to specify a maximum term order override value for each risk factor being included in the function.

CHOOSING THE CANDIDATE TERMS

There are many selection methods and configuration options that could be used to choose the functional form when performing a function fitting process. The choices made have implications for the quality of fit of the function, the speed at which the fitting process completes and other subtler considerations such as term marginality[2].

In the early days of curve fitting and LSMC, the identification of the functional form was a manual and iterative process, as users tried to identify which terms (combinations of risk factors) in the polynomial were important to the fit. This is a time-consuming process and does not guarantee an optimal fit. It can also lead to terms with little or no explanatory value being included in the function or the function fitting to the random error or noise instead of the underlying relationship, resulting in over-fitting.

This risk can be reduced by automating the process of identifying the optimal functional form. A stepwise algorithm optimizes the function produced by the regression process by incrementally adding terms. The algorithm looks at all or a subset of the candidate term set at each stage, depending on the algorithm used, and assesses which term most improves a selection criterion, for example Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC).

---

2 This requires that any candidate term can only be added to a given model if the model already contains all terms which are marginal to it. The marginal terms to a given candidate term are its various algebraic factors. For example, $X_2^z X_z$ has terms X_2, X_z, X_2^z and X_2, X_z as marginal terms. Were these terms not present, then the model would not be invariant to an arbitrary change in scale/origin.

Figure 2: Stepwise Regression Process
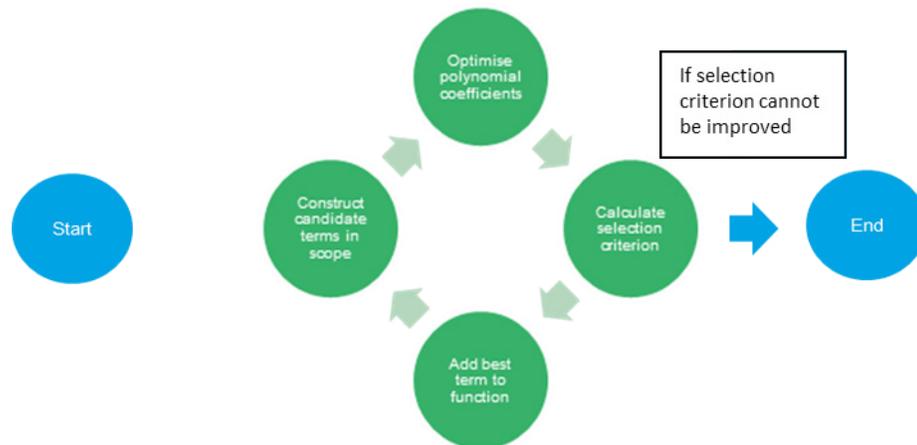
Validation

After a proxy function has been fitted, it is important to validate the function which is produced to ensure that it is an accurate representation of the asset or liability being modeled. There are several different methods which can be used:

1. **Out of sample** – when the proxy function has been created, it is necessary to assess how accurately the proxy function calculates scenarios that are outside the fitting sample. Achieved by comparing in-sample and out-of-sample R^2 values and confirming that the latter does not display a significant reduction in the R^2 metric.

2. **Full stochastic validation** – if liabilities are being valued stochastically using market-consistent scenarios, valuing a small set of multivariate real-world stressed ESG simulations with many market-consistent inner scenarios give accurately valued points in the real-world space due to the reduction of sampling error as the number of market-consistent scenarios increase. The values produced for the scenarios using the ALM model and the corresponding proxy function predictions are ideally close. The differences between the two values for each scenario can be used to calculate error metrics to assess the goodness of fit for the proxy function.

3. **Residual analysis** – residual plot, partial residual plot, QQ plots, are tools that can help assess the goodness of fit from a statistical perspective. They are useful to check for statistical dispersion such as heteroscedasticity which could bias the resultant proxy function.

4. **Risk drivers fit analysis** – is valuable to look at how the function behaves as a single risk factor is stressed across its economic range, and can help assess if the risk driver is affecting the value being modeled and whether the economic range is correct.

5. **Confidence intervals** – around validation points allow further interpretation of goodness of fit and help define drift due to sampling error.

Figure 3: Examples of Validation Charts (Source: RiskIntegrity™ Proxy Generator)
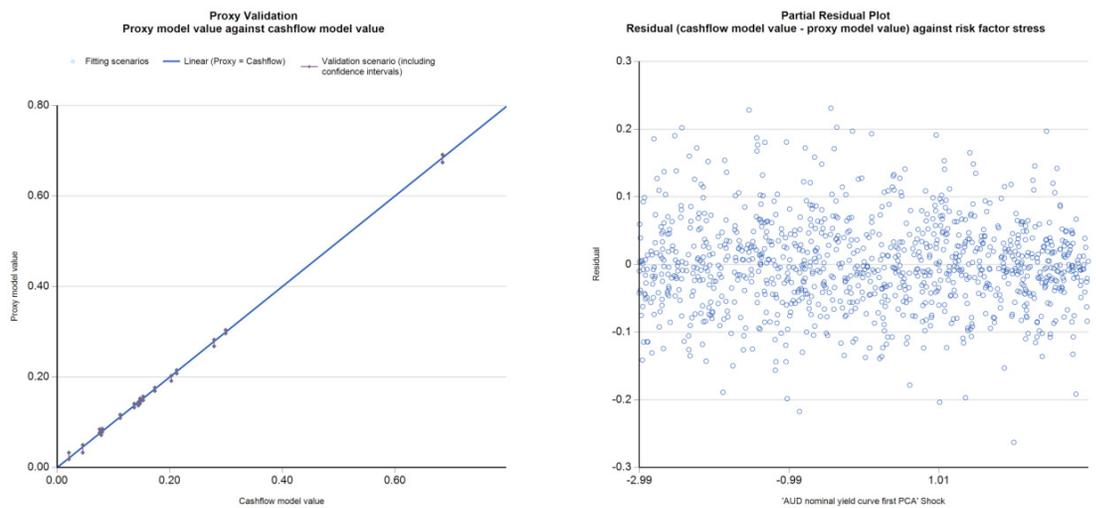
Performance

Under either the Curve Fitting or LSMC methods, performance is an important consideration to allow often aggressive production timelines to be met.

When fitting proxy functions for many assets or liabilities, it is important to optimize the function fitting process to minimize run time and make most efficient use of the computing resources assigned to the proxy generation process. Similarly, a significant number of market-consistent scenarios are run to generate the fitting scenario set which takes a significant amount of time, even if automated.

Possible methods of optimizing performance include:

» Parallelism of stress operations – allows multiple fitting scenarios to be run in parallel on a single computer, using the individual cores available on the processor.

» Parallelism of regression operations – when running a stepwise algorithm multiple regressions are required at each stage to identify which term to add to the function. Allowing multiple regressions to be run in parallel on a single processor reduces overall run time significantly.

» Scalability and Task Distribution – the capability to support scenario generation or function fitting across multiple processors or servers greatly increases the ability to run tasks in parallel. The scenarios or proxies to be run can be split up and distributed between all available processors, offering significant reductions in overall run time.

» In memory scenario generation – the capability to run ESG simulations in memory and read the ESG outputs in memory reduces the time spent carrying out input/output tasks and the risk of errors occurring during these tasks.

Conclusion

The use of proxy functions has increased, and continues to increase, as the techniques involved have become widely recognized as credible solutions to key business challenges. Insurers are increasingly looking to proxy techniques as a way of helping them understand their risks and improve their decision-making process. The theory behind these techniques is becoming better understood, but implementing them in a robust and integrated way as part of business as usual (BAU) processes still involves challenges.

This paper has looked at many of the practical issues which can be encountered when implementing the Curve Fitting or Least squares Monte Carlo proxy techniques. It has reviewed some of the main challenges encountered at the various stages of the process, particularly during the generation of scenarios for use in the fitting of proxy functions, and shown some ways in which these challenges can be addressed.

For many of the challenges covered in this paper, automation is a major component in implementing proxy methods as part of BAU processes. Implementation often requires robust, enterprise technology to manage the scenario generation and function fitting processes efficiently, while maintaining auditability and transparency. Integrating components such as ESG, calibration tools, and regression algorithms into the proxy generation process also facilitates greater automation and reduces operational risk.

Performance is also an important issue. Scalable architecture and parallelization are two possible methods which can make more effective use of computational resources for both scenario generation and function fitting, and help to meet production timetables.

The methods discussed in this paper continue to evolve, and the requirement for rapid risk-based modeling increases. The power of proxy techniques is becoming clear, as is the value to insurers of investing resources in finding ways to resolve the implementation challenges.

Appendix A - Background

Under a traditional Monte Carlo approach, the ability to value complex liabilities using many scenarios, for example in the calculation of economic capital or during an assessment of hedge effectiveness, can be limited by the considerable run-times associated with asset-liability management (ALM) models.

Figure 4: Nested Stochastic Simulation
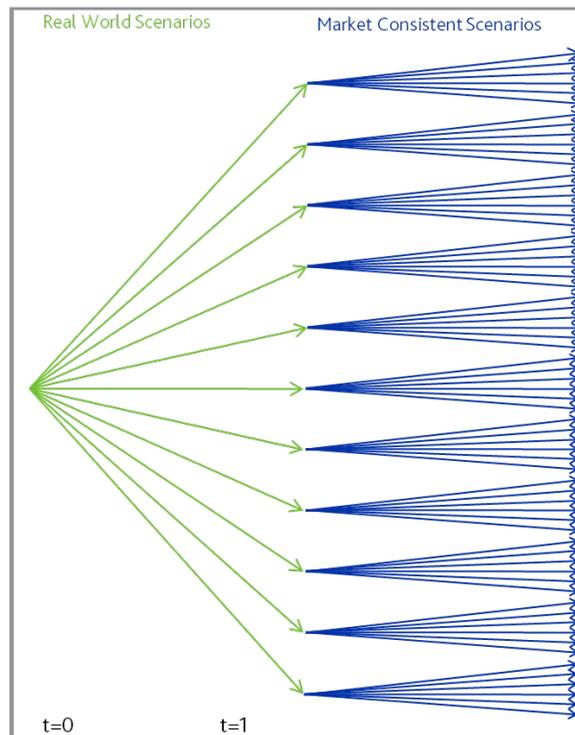
The nested stochastic process required to obtain the liability values involves running many real-world (outer) simulations, with each outer simulation valued by many market-consistent (inner) simulations. This process, as illustrated in Figure 4, can rapidly become very time consuming and require a large amount of technical resources. For example, if 100,000 real-world simulations are used, each with 1000 market-consistent simulations, then the total number of simulations run is 100,000,000, which for most models requires a long time to calculate and creates a huge computational burden.

Proxy functions have been developed as a more feasible method for calculating Monte Carlo risk-based capital. These techniques have developed rapidly in recent years. Initially, some firms developed replicating portfolios to model their liabilities. The usefulness of the technique was limited as it proved difficult to find assets with the same risk characteristics as more complex liabilities, and the technique typically cannot deal with non-market risks. As a result, curve fitting and, more recently, the Least Squares Monte Carlo (LSMC) technique, originally developed by Longstaff and Schwartz as a method for valuing American options, have gained popularity. These techniques use more general mathematical functions rather than a replicating portfolio to model liabilities (or assets).

Figure 5: Curve Fitting and Least Squares Monte Carlo Methods
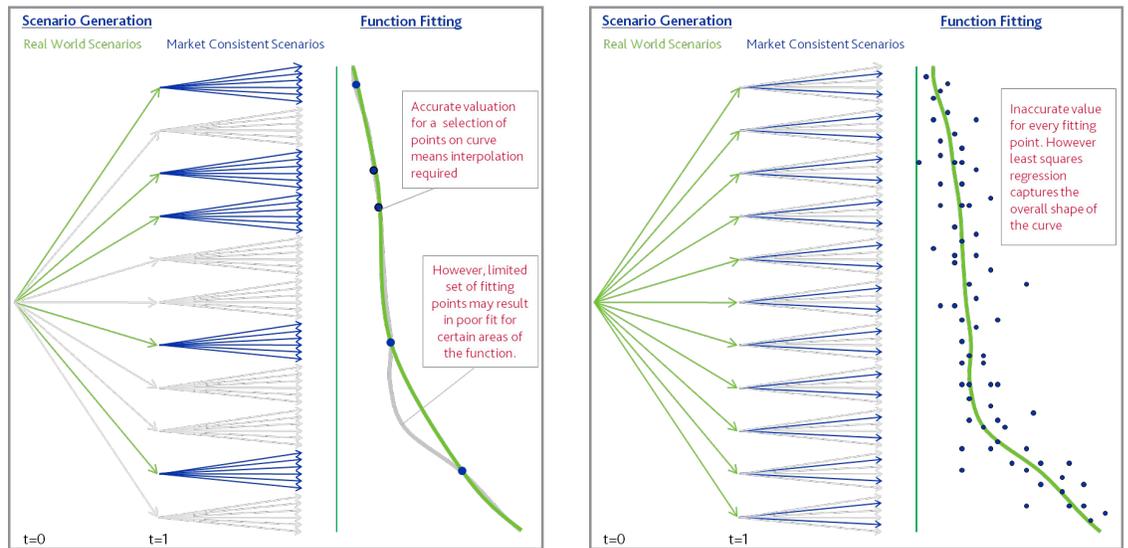
The proxy functions produced by these techniques can be used to value the asset or liability being modeled under many scenarios, greatly reducing the time and computational power required, for example to calculate the 1-year Value at Risk (VaR).