Market Data & Risk Master

Model Risk Management for a Trading Firm’s Calculations

A Data Centric View of Model Risk Management

Charlie Browne

Contents

Introduction

Regulatory guidance over the last number of years requires that a bank has a robust set of Model Risk management processes in place. Examples of such regulation are:

- **FRTB:** Bank for International Settlements (BIS) Basel Committee on Banking Supervision, Minimum capital requirement for market risk, January 2016, also known as "Fundamental Review of the Trading Book"
- **PRA Model Risk Management:** The Bank of England, Prudential Regulation Authority, Supervisory Statement | SS3/18, Model risk management principles for stress testing, April 2018
- **TRIM:** The European Central Bank (ECB) Guide for the Targeted Review of Internal Models (TRIM) – in support of the European Banking Authority (EBA) Capital Requirements Regulation (CRR)
- **Fed Model Risk Management:** Federal Reserve and Office of the Comptroller of the Currency OCC, SR 11-7: Guidance on Model Risk Management

This paper:

- describes what each of these regulations say about Model Risk Management,
- describes a framework for the different types of models that are used within a trading firm. The framework is referred to as the modeling ecosystem
- gives an overview of the importance of data lineage in a modeling ecosystem
- provides background to some key data lineage concepts
    - instrument lineage
    - market data traceability
    - model validation lineage
    - product taxonomies

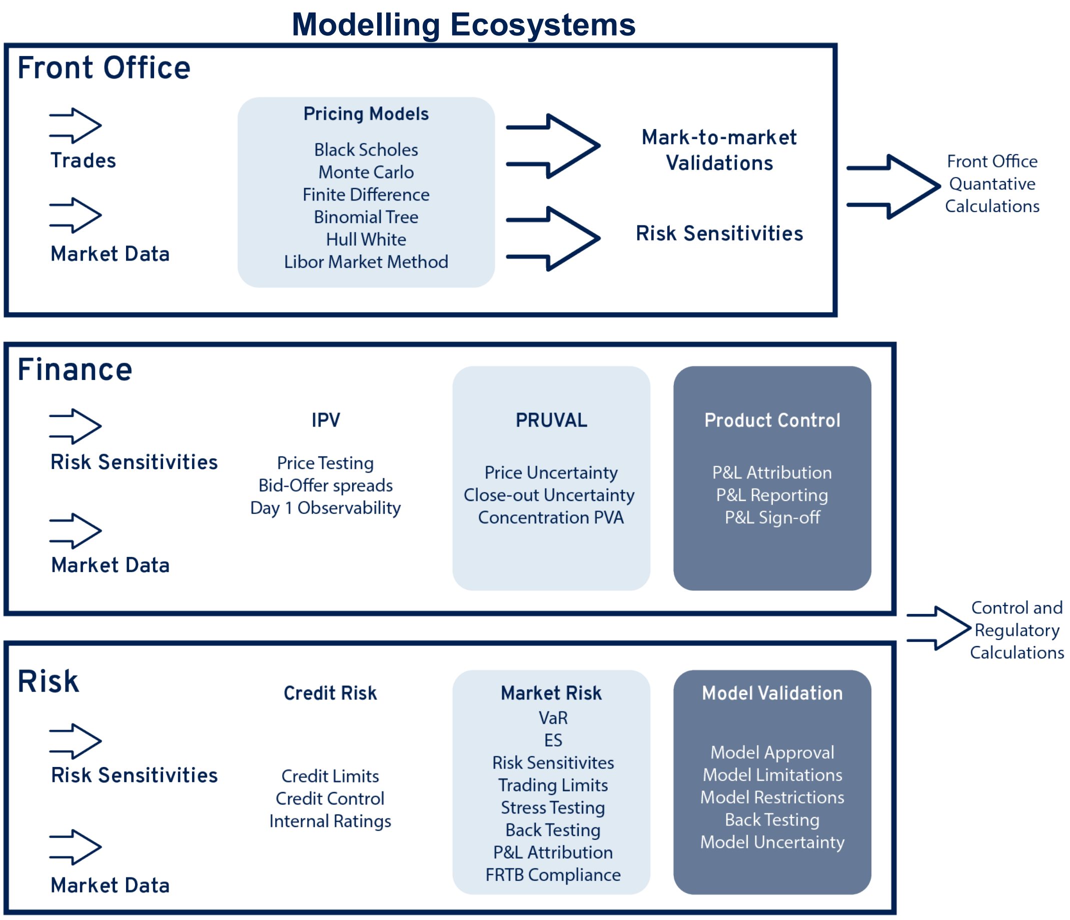
The Modeling Ecosystem

A key theme in much of the regulation that impacts model risk management is the need for transparency in the processes that generate the P&L and capital charges that impact bank balance sheets. In particular, regulators are looking for more visibility of the inputs, processes and outputs generated by quantitative models that are used to both

- generate fair values and associated daily P&L figures and
- generate the P&L vectors used in internal models for market risk management

The inputs, processes and outputs used to generate P&L can be thought of as modeling ecosystems. Modeling ecosystems can in turn be broken into two separate sub-categories:

- Modeling ecosystems for the front office that calculate fair values of derivative and other held-for-trading positions and
- Modeling ecosystems used for risk and finance control and regulatory calculations.



Per the diagram above,

- **Front Office Modeling Ecosystems** use quantitative models to calculate the fair values of derivative and other held-for-trading positions. They are either developed in-house by the bank's quantitative research development team or they are routines or models that come with 3rd-party libraries that have been installed and configured in the front office. Examples are SABR, Black-Scholes and the Hull-White interest rate model. The primary purpose of front office quantitative models is the calculation of fair values (also known as mark-to-market values (MTMs), or present values (PVs)) for the bank's held-for-trading positions. Daily P&L is in turn derived from the daily changes in these fair values.

Front Office modeling ecosystems are also used to calibrate the market parameters or risk factors that are needed as inputs to quantitative models. See Market Rates versus Risk Factors below for more detail on this.

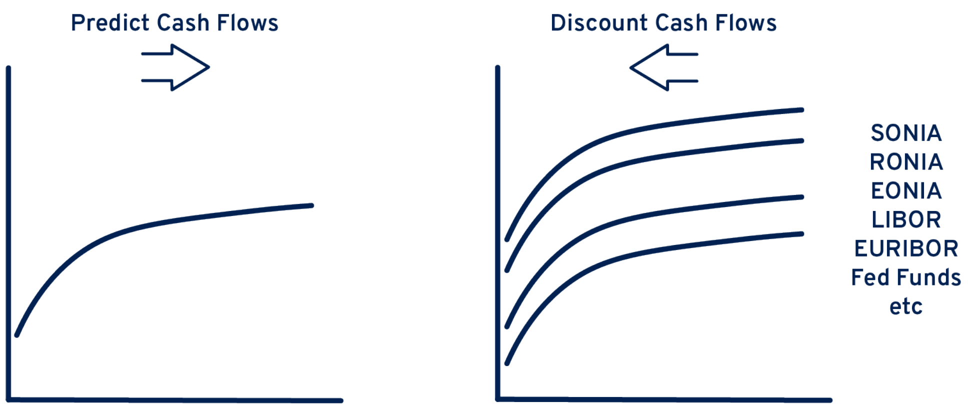
- **Finance and Risk Modeling Ecosystems.** The control regulatory calculations used by Finance and Risk modeling ecosystems are a separate set of calculations that typically sit downstream of front office quantitative calculations. Control and regulatory calculations use both the outputs generated by front office calculations (fair values, risk sensitivities, calibrated market parameters) and other inputs (market data, static data, position data) in order to compute the required calculations. Expected Shortfall, the aggregation of risk sensitivities across correlated risk factors, Value-At-Risk, P&L Attribution, Bid-Offer Reserves calculations, CVA, FVA, Market Price Uncertainty, Close-out Costs Uncertainty, and IPV variances are all examples of control and regulatory calculations that are owned typically by either Risk or Finance within the trading firm.

Front Office Quantitative Calculations

Theoretical Background

The valuation of a derivative or any financial instrument that is held in the trading book can be considered a two-step process:

- Predicting future cash flows: calculation of the expected value of future cash flows and
- Discounting cash flows: calculation of the present value of those predicted cash flows



Predicting future cash flows

There are various well-established techniques for predicting the size of uncertain future cash flows associated with derivatives and other financial instruments held in the trading book. Some categories of these techniques are described below.

- Swap Valuation: the value of a swap is the difference between the present values of the future cash flows of both legs of the swap. Calculating the future cash flows of the fixed leg of the swap is straight-forward. Where the leg of the swap is based on a floating interest rate (e.g. SONIA) or on an equity index, then the swap valuation model needs to be able to estimate the future values of those indices at different future dates.

- Black-Scholes option valuation: Black & Scholes (1973) and later Merton demonstrated that an option's value is the difference between the strike price of the option and the expected value of the asset at expiry of the option adjusting for the fact that the asset price may not turn out to be greater than (in the case of a call option) the strike price of the option. The solution to the Black-Scholes equation is closed-form for European-style options. The path that the asset price follows in the time up to expiry is assumed 1) to exhibit geometric Brownian motion and 2) to grow at the risk-free rate of return.

- Binomial Tree methods: Schwartz (1977), Brennan and Schwartz (1979), and Courtadon (1982) were the pioneers of this method. It involves the calculation of a range of potential future prices of the underlying asset at expiry (and for path-dependent derivatives, at intermediate points in time too). The range of prices is calculated by stepping through a binomial tree after assigning probabilities to both an upward price move and a downward price move at each step. The payoff of the option at expiry for each of the potential underlying prices is then calculated.

- Monte Carlo approaches: Like binomial tree methods, the Monte Carlo approach (Boyle, 1977) involves the calculation of a range of potential future prices and the resultant payoffs at expiry. But whereas the final pay-offs in the binomial tree method are arrived at after stepping through a lattice, the Monte Carlo method calculates the range of potential future prices by simulating paths of a large number of the asset's prices using the appropriate stochastic model. Each path generated by the stochastic model simulates the time evolution of the underlying asset after making assumptions about the expected return and volatility of the asset.

- Finite difference techniques: Finite difference techniques approximate the solution to the partial differential equation that Black & Scholes, above, demonstrated that an option must satisfy when assumptions about asset price dynamics hold. The Black-

Scholes Partial Differential Equation (PDE) models how an asset value changes through time.

Discounting cash flows

The second step in the two-step process for calculating the value of derivatives is the discounting of the derivatives (uncertain) future cash flows. Almost all financial derivatives' pricing in practice comes down to the basic valuation formula described below. The net present value (NPV) of a derivative contract can be expressed as a sum of expected cash flows (C), weighted by the probabilities of that payout (P), discounted to valuation date using discount factors (D).

$$\mathrm{NPV} = \sum_{i=1}^{n} \mathrm{D}_i \mathrm{P}_i \mathrm{C}_i$$

One example of a derivative with uncertain future cash flows is an option. In the case of a call option the uncertain future cash flow = Max [Spot - Strike, 0]. When the future cash flows are looked at in continuous time, the sum in the NPV equation above simply becomes an integral. The basic two-step approach to valuations, however, remains the same.

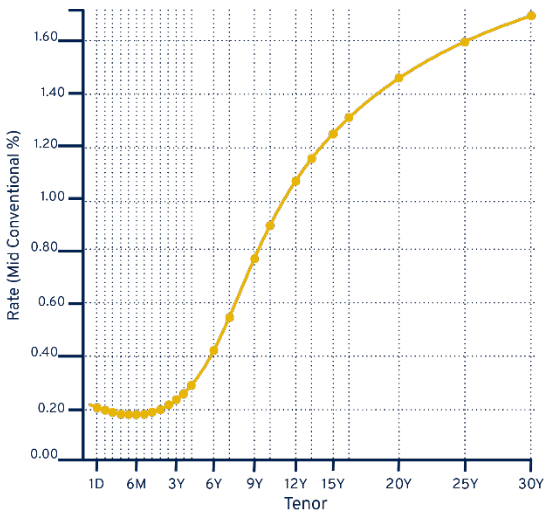
So, which discount factors should be used for the D in the NPV equation above? Prior to the financial crisis of 2008/2009, academic literature assumed broadly that the interest rate curves that were used for step 1 of the derivative valuation process, i.e. the predicting of future cash flows, could also be used for step 2 of the derivative valuation process (the discounting of those predicted future cash flows). The literature described the building of the zero coupon LIBOR curve using LIBOR fixings at the short end, LIBOR futures in the middle part of the curve, and LIBOR interest rate swaps further out the curve. Using a technique called "bootstrapping" and various interpolation techniques for handling gaps in the swap curve, a zero coupon curve was built. The discount factors to be used for derivatives pricing were then derived from that zero coupon curve.

But after the financial crisis market practitioners questioned the use of LIBOR inter-bank deposits that clearly contained inter-bank credit risk as the basis for the indexes underlying the financial instruments that make up a currency's interest rate curve. Was a

LIBOR curve really the right curve to use for the discounting of a derivative's cash flows? As the use of bilateral master derivatives agreements (including appropriate Credit Support Annexes or "CSA") become more prevalent in the aftermath of the crisis, the answer was, clearly, no. With derivatives that are subject to CSAs, each counterparty is expected to post margin to eliminate the credit risk inherent in in-the-money derivatives. But what then was the true cost of financing the derivative's cash flows? And if that true cost is known, then surely it is a set of rates associated with this financing cost that should be used to discount the derivative's cash flows as opposed to LIBOR or any other rates. As various solutions were explored a consensus emerged. The rate of return of the collateral posted as margin under the CSA was assumed to be the rate that best reflected the cost of financing the derivative. And that rate of return on the collateral would, therefore, also become rate at which derivatives' future cash flows should be discounted. Depending on the type of collateral posted, further standards emerged with respect to the the choice of discount curve to use. Broadly, two types of collateral can be posted against a derivative:

- cash
- bonds

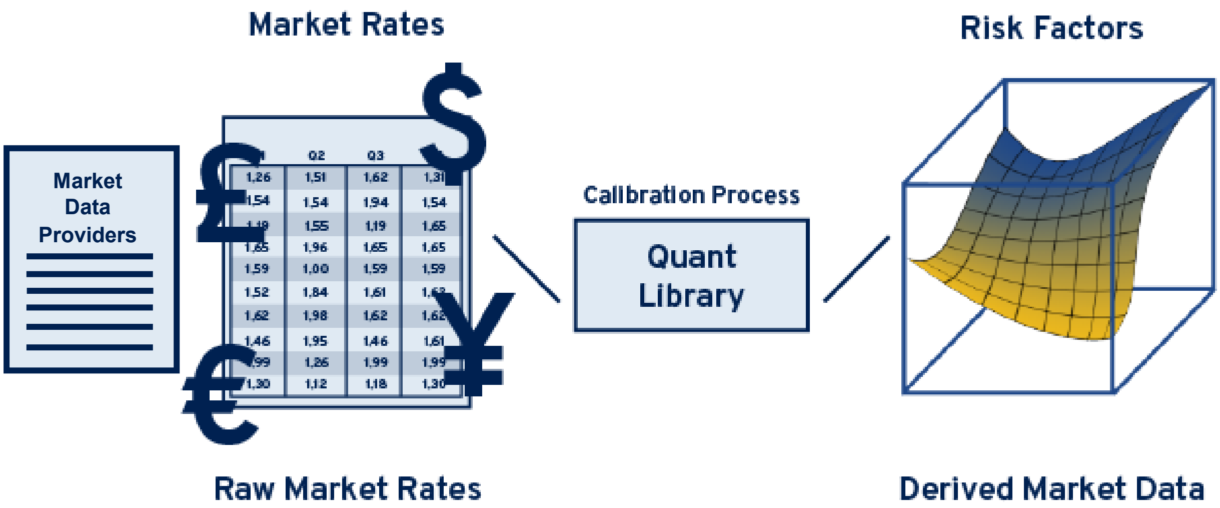
Where cash is posted as collateral, practitioners use a curve that captures a term structure of overnight (i.e. cash) rates. In most currencies there is an active market to price overnight rate expectations years into the future. It is often based on the Overnight Index Swap (OIS) market of the currency in question, e.g. the chart below shows the term structure of JPY OIS.



When this OIS term structure is converted into a zero coupon curve (via the appropriate bootstrapping process) a set of discount factors is generated that are deemed to reflect the true cost of financing the derivative's cash flows. Where bonds are posted as collateral, the rate of return of the bond is used to create the discount factors used. Most market participants agree that the bond's repo rate is a good approximation for its rate of return.

Market Rates, Risk Factors and the process of Calibration

Typically market rates are not used directly by quantitative models. The market rates first need to be transformed into curves of smoothed rates or parameters that the models are programmed to understand. The diagram below illustrates the process:



Market Data: these are the prices or quotes of traded instruments in the market place. Examples are Swap rates, CDS Spreads, Volatility quotes, and futures prices. Market rates, par rates, market quotes are all examples of terms that fall under the broader category of Market Data.

Risk Factors: Valuations and P&L is created by quantitative libraries by passing in risk factors Changes in risk factors result in P&L. However, and this is a critical point, the risk factors that are created using par rates from curves and surfaces need to go through a transformation process before they can be passed in those valuation routines. This is because a bank's proprietary routines are bespoke by nature and require versions of those curves that are transformed into a format that the valuation engines can accept. The transformed market rates can be said to come in two forms:

- Calibrated rates
- Calibrated parameters

From purely a data perspective, both of these data types do the same thing: they are inputs to a routine that produce valuations.[1] In market risk terms, calibrated rates and calibrated parameters are referred to as risk factors.

[1] The other inputs into valuation routines are position data, static or reference data, and conventions data

Control & Regulatory Calculations

Finance and Risk teams have modeling ecosystems that perform control and regulatory calculations. The responsibility matrix below includes a typical list of calculations and the teams across Finance and Risk that own them.

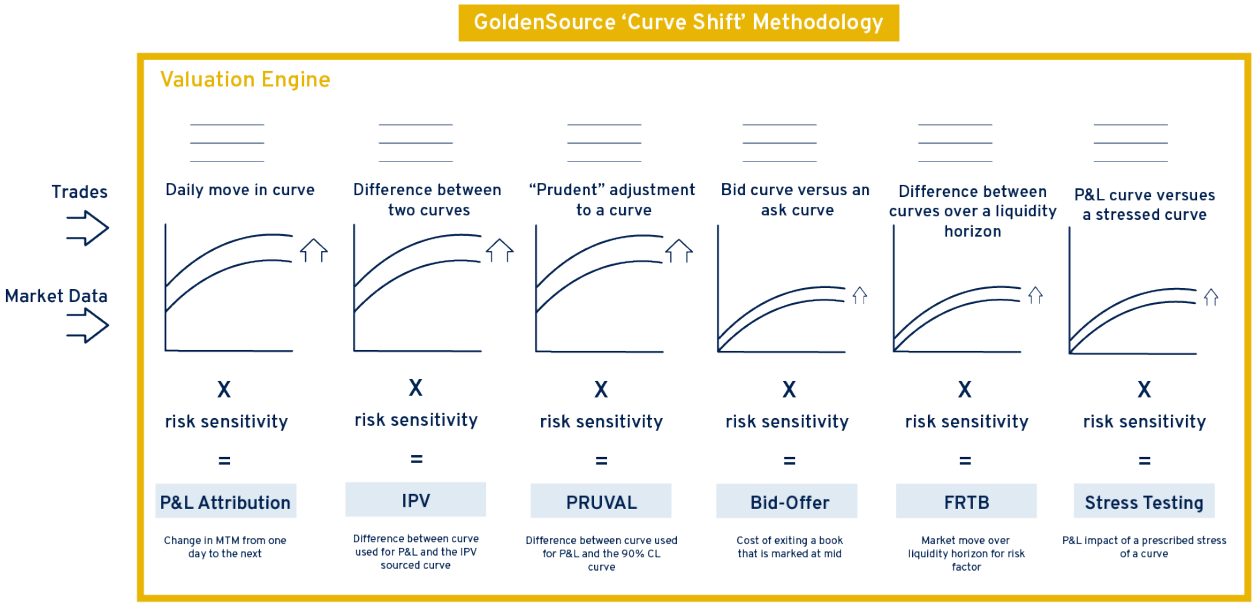
Area	**Team**	**Calculation**	**Valuation Adjustment**	**Engine**
Risk	Market Risk	Expected Shortfall		VaR Engine / Quant Library
Risk	Market Risk	Value At Risk		VaR Engine / Quant Library
Risk	Market Risk	Risk Sensitivity		Quant Library
Risk	Market Risk	P&L Vector		Quant Library or Risk Sensitivity based calculation
Finance	IPV	IPV Variance		Quant Library or Risk Sensitivity based calculation
Finance	IPV	IPV Adjustment	Y	IPV Engine
Finance	Product Control	Bid-Ask Reserve	Y	Reserves Engine or Risk Sensitivity based calculation
Finance	Product Control	Official Valuation		Quant Library
Finance	Product Control	Clean Official P&L		Quant Library
Finance	IPV or Model Validation	Model Reserves	Y	Reserves Engine /Quant Library
Finance	IPV or Product Control	Credit Valuation Adjustment (CVA)	Y	XVA Engine / Monte Carlo Simulation
Finance	IPV or Product Control	Funding Valuation Adjustment (FVA)	Y	XVA Engine / Monte Carlo Simulation
Finance	IPV or Product Control	Capital Valuation Adjustment) KVA	Y	XVA Engine / Monte Carlo Simulation
Finance	IPV or Product Control	Margin Valuation Adjustment (MVA)	Y	XVA Engine / Monte Carlo Simulation
Front Office	Trading	Valuation		Quant Library
Front Office	Trading	Clean Flash P&L		Quant Library

- The team ownership structure used above is indicative and used to aid understanding. The exact team structures and terminology used will vary depending on the bank
- IPV stands for Independent Price Verification

- Valuation Adjustments are typically owned and approved by Finance but are often calculated using Risk and Front Office Systems
- Value at Risk is the calculation used by Market Risk teams to calculate the amount of regulatory capital a firm or desk needs to set aside to cater for market price fluctuations. For example, a 1-day, 95% VaR calculation will calculate the worst P&L day over a 1 year period to a 95% confidence level.
- Expected Shortfall (ES) is deemed, post FRTB, to be a better methodology for calculating market risk capital under the internal model approach than VaR. ES goes a step further than VaR in that it calculates the actual expected loss in the event of a bad P&L day.
- A risk sensitivity is a calculation that approximates the P&L of a position or portfolio if a market data input is shifted by a standardized amount (e.g. 1 basis point).
- Most of the calculations above can be calculated on either a risk-sensitivity basis or a full-revaluation basis.
    - Risk sensitivities are typically produced using the bank's Quant Library by shifting curves or other risk factors by a standardized market data amount (e.g. a basis point). The risk sensitivity is the difference between the valuation produced before the risk factor shift and the valuation produced after the risk factor shift. See the section on Curve Shift Approaches below for an overview of how risk sensitivities are used to generate the P&L & Risk Charges required by Finance and Risk.
    - A full revaluation approach does not require risk sensitivities to be produced. It instead takes market data (or risk factor) inputs, applies them to the trades and positions within the portfolio and revalues the entire portfolio using the bank's quantitative models. Full revaluation is a more accurate approach than the risk sensitivity approach. But it is also computationally more expensive, requiring more powerful IT systems to perform the calculations.
- Model Reserves are valuation adjustments that impute the P&L that needs to be charged to a desk for risk that models produce P&L that is not complete / contains non-linear risks that are difficult to capture within the model.
- XVA are valuation adjustments that describe the various adjustments that are required to account for adjustments that were deemed immaterial prior to the crisis but are now required and commonplace within finance and risk modelling ecosystems.
    - CVA: Credit Valuation Adjustment. This is an adjustment to the value of a derivative to account for the fact that there is credit risk inherent in derivative positions that are in-the-money to the bank.
    - FVA: Funding Valuation Adjustment. This is an adjustment to the value of a derivative to account for the fact that banks hedge their uncollateralized derivative portfolio with collateralized derivatives. The collateral that is posted in the collateralized derivative portfolio needs to be funded. There is a cost associated with this funding.
    - KVA: Capital Valuation Adjustment. This is the lifetime cost of holding regulatory capital for an OTC derivative trade. It is an up-front amount that would generate a specific return on capital via rolling profits over the lifetime of the transaction
    - MVA: Margin Valuation Adjustment. This is the cost of funding initial margin that arises where (as in most cases) the re-hypothecation of the initial margin is not allowed.

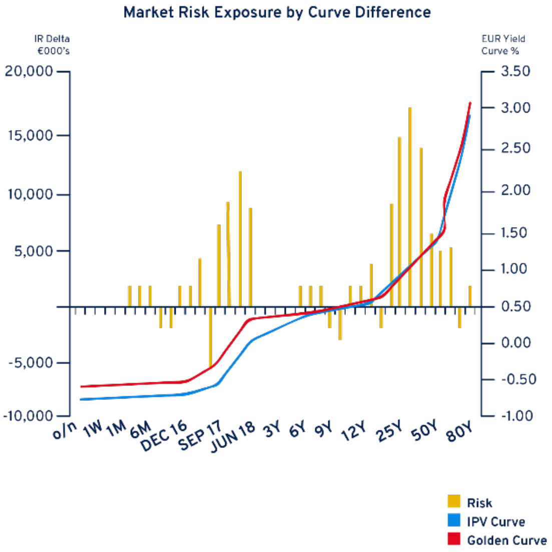
Curve Shift Approaches

The diagram below provides an illustration of how changes in market rates (or risk factors) can be applied to risk sensitivities to create P&L and capital charges within Finance and Risk modeling ecosystems. These are risk-sensitivity based approaches to Finance and Risk calculations.



In each calculation in the diagram above the market data (or risk factor) move is applied to (multiplied by) a risk sensitivity in order to generate a P&L amount or a capital charge.

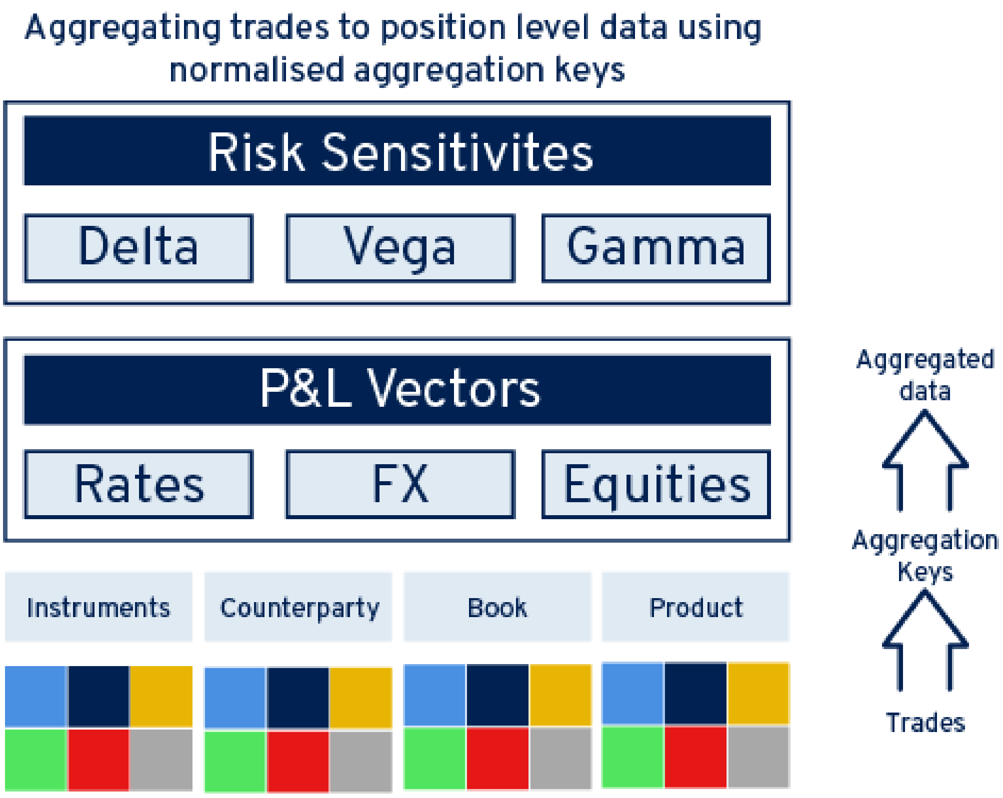
The ability to visualise risk sensitivities adjacent to the market data they are sensitive to is key to the approach. At each point along a curve or surface, the difference between two market data points can be visualized. This difference, whether in basis points, volatility points or credit spread, has a meaning in P&L or Capital Charge terms.



Seeing curve differences aligned with the buckets to which they are sensitive empowers end-users and analysts.

Aggregating Risk Sensitivities

Each of the regulations described in this document, and in fact many calculations in finance and risk modeling ecosystems, require the calculation and aggregation of risk sensitivities. The P&L of every trade contained in the desk's portfolio is sensitive to the daily changes in the model's risk factors. These sensitivities are called risk sensitivities. Risk sensitivities can be used in calculations at many levels throughout the bank. The ability to aggregate them using a commonly agreed approach using standardized, firm-wide identifiers is, therefore, essential. The diagram below provides an overview of the process of aggregating risk sensitivities using standardized identifiers (aggregation keys).



Modeling Ecosystems Data Flows

Data inputs into a typical modeling ecosystems can broadly be broken into two categories

- Market Data
- Position and Transaction Data

An alternative way of viewing these two categories of data is to view them as external versus internal data blocks. Market data can largely be considered external data as it relates to prices of instruments traded outside of the bank and is sourced primarily from external data vendors and exchanges. Position and transaction data, on the other hand, is internal data. It is information about the banks own trades, risk exposures and P&L.

Market Data

Market data can be introduced into a modeling ecosystem from a number of different sources:

- Snapped and End-of-Day vendor feeds
- Ticking prices (snapped or filtered)
- Front offices trading systems

With MIFID II, it is also likely that qualifying price transparency vehicles such as APAs and CTPs will become commonly used sources for market data (pre and post trade).

Clean validated time-series market data is required for most Internal Models Approaches within modeling ecosystems. Tools for choosing the best price (a golden price) for an instrument or a point on a curve each day are critical for success in a modeling ecosystem. The clean, validated daily golden prices (with full lineage to underlying raw market data) become the prices that are each day added to the time series required for calculations. Time series prices are required for many calculations, including:

- P&L Attribution
- Value-At-Risk
- VaR back-tests
- Expected Shortfall
- Stressed Expected Shortfall
- Standard Deviations for Day-on-Day tolerance movements
- Principle Component Analysis
- Correlation Matrices
- Risk Factor de-compositions
- Default Risk Charges

Where a bank decides that it wants to trade new products (and those products require new risk factors or instrument market data), it is important that the modeling ecosystem has functionality that will allow the introduction of the associated time series market data. Because the products being traded are new, this time series market data will not be available. Typical time-series functionality required when new instruments are traded include:

- Time Series Validation (zero checks, null checks, missing data)
- Time Series filling or proxying rules (interpolation, use of basis spreads, cross instrument proxying)

Banks in many cases will have thousands of risk factors and instruments for which they require time series data. FRTB requires that Expected Shortfall calculations use time series data that goes all the way back to 2007. The ability to store and access such large volumes of data is critical for success in FRTB in particular.

Instruments and market data

Three basic categories of instruments can be defined with respect to market data:

- Type 1: Securities or cash instruments
- Type 2: Quote instruments
- Type 3: OTC Derivatives

Instrument	Market Data Instruments	Tradeable	Contain Notional Values	Contain MTM Values
Type 1: Securities or cash instruments Examples: Equities, Bonds, FX Spot	Yes	Yes	Yes	Yes
Type 2: Quote instruments Examples: IR Swap rate, FX forward rate, FX Fixing, Gold Fixing	Yes	No	No	No
Type 3: OTC Derivatives Examples: IR Swap, FX Forward, FX Option	No	No	Yes	Yes

Some inferences …

- Market data instruments cover type 1 and type 2 above
- Type 3 need to be linked to market data instruments
- Fixing instruments are not traded and are therefore Quote instruments (type 2)
- Swap rates are quote instruments (type 2) whose identifiers are created by data vendors so they can publish the average (e.g. fix-for-floating rates) that they see/collect
- Type 1 instruments above act as both traded instruments and quote instruments

Position and Transaction Data

Position and Transaction (P&T) data can be introduced to modeling ecosystems from the bank's front office, risk or finance systems. P&T data is internal, proprietary data that often contains sensitive information about the bank's positions or its customers. The following categories of internal P&T data can be classified:

- Trade level Risk Sensitivities
- Position level Risk Sensitivities
- Trade level P&L
- Position level P&L

Risk and Finance Calculations and related Data Requirements

The table below describes the type of data that each Risk and Finance calculation requires

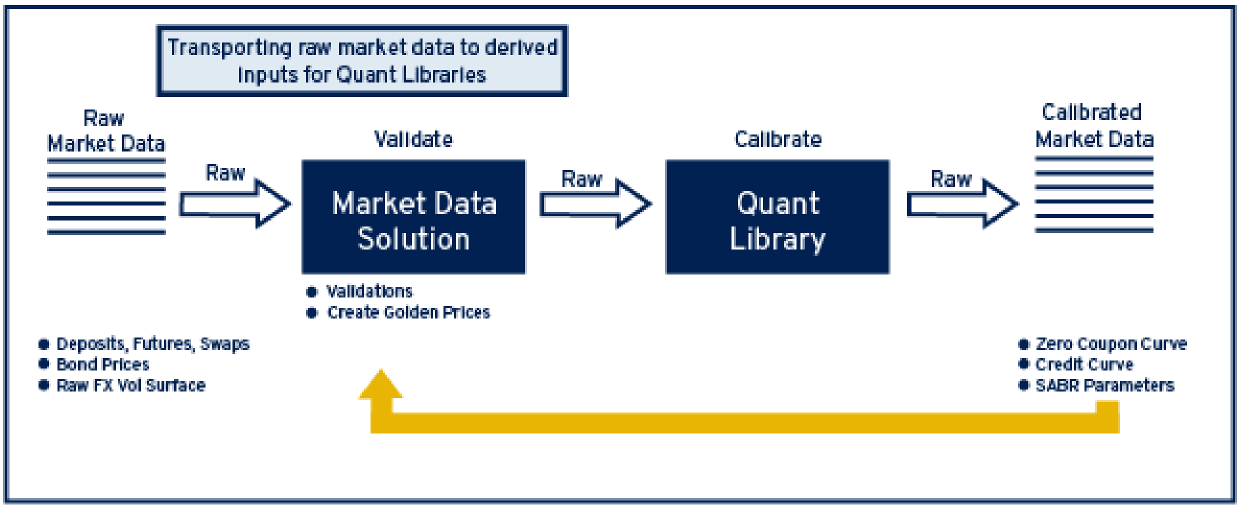
Calculation	Point in Time P&T Data	Time Series Market Data	Time Series P&T Data
Expected Shortfall	Yes. Risk Sensitivity as at valuation date. E.g. position-level IR Vega on the date for which the ES calculation is being run	Yes. Time series market data is required to calculate the N-day market moves that are applied to the point-in-time position. E.g. 10-day move in EUR Swaption Vol is derived from time series of daily market data moves	Yes. Full revaluation P&L (based on N-day market moves) is required where a full revaluation approach is chosen over a risk sensitivity approach
VaR back-test	Yes. Time series risk sensitivities are required to support investigations where VaR back tests fail	Yes. Time series market data is required irrespective of whether VaR is calculated using a risk sensitivity approach or a full revaluation approach	Yes. Time-series P&L by Risk Factor. The P&L can be calculated on either a risk sensitivity basis or a full revaluation basis Time series risk sensitivities by Risk Factor will be required if the P&L used in the VaR back-

Calculation	Point in Time P&T Data	Time Series Market Data	Time Series P&T Data
			test is derived using risk sensitivities
P&L Attribution	Yes. Time series risk sensitivities are required to support investigations where P&L Attribution tests fail	Yes. Time series market data is required for P&L Attribution to support investigations where P&L Attribution tests fail	Yes. Time series "Risk Theoretical P&L" and time series "Hypothetical P&L" are required
Risk Charge	Yes. Risk Sensitivity as at valuation date. E.g. risk weights are applied to position-level IR Vega	No. Not Required	No. Not Required
Standard Deviations	No	Yes. Standard Deviation is a measure of the variability (and therefore risk) of market data movements around a mean movement over a period of time	Yes. Required to generate P&L vectors for VaR and Expected Shortfall calculations
Principal Component Analysis	No	Yes. E.g. calculate the covariance matrix from a time series of day-on-day movements in the rates. Use the covariance matrix to calculate the first principal component which captures variations in the curve due to shifts in the level of the curve	No
IPV Variance	Yes. (IPV Rate less Front Office Rate)*Notional or Risk Sensitivity	No	No
Bid-Ask Reserve	Yes. Bid-Ask spreads * Risk Sensitivity	No	No
Official Valuation	Yes. Trade and position data is required to generate valuations	Yes. Historical market data is required to generate P&L vectors for VaR and ES	Yes. Trade and Position data is required to generate P&L vectors for VaR and ES calculations
Clean Official P&L	Yes	Yes. See Official Valuation above. Clean P&L excludes valuation adjustments	Yes. See Official Valuation above. Clean P&L excludes valuation adjustments

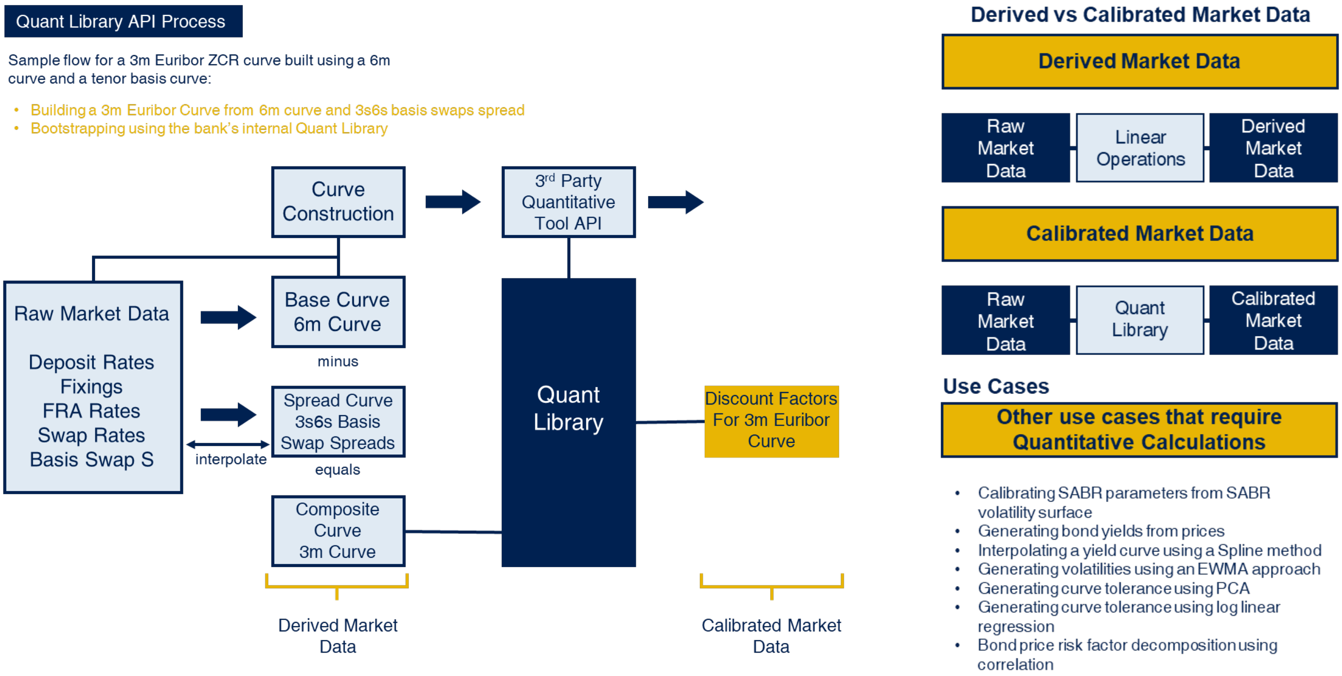
Calculation	Point in Time P&T Data	Time Series Market Data	Time Series P&T Data
Model Reserves	Yes. Market	Possible	No
Credit Valuation Adjustment	Yes. See CVA above in Control and Regulatory Calculations	Possible. Depending on how exposures are generated	Possible depending on how exposures are generated
Funding Valuation Adjustment	Yes. See FVA above in Control and Regulatory Calculations	Possible. Depending on how exposures are generated	Possible depending on how exposures are generated
KVA	Yes. See KVA above in Control and Regulatory Calculations	Possible. Depending on how exposures are generated	Possible depending on how exposures are generated
MVA	Yes. See MVA above in Control and Regulatory Calculations	Possible. Depending on how exposures are generated	Possible depending on how exposures are generated
Valuation	Yes	No	No
Clean Flash P&L	Same as Clean Official P&L except generated by the front office	Same as Clean Official P&L except generated by the front office	Same as Clean Official P&L except generated by the front office

Calibrating Market Data

Models require inputs that they can understand. See Market Data versus Risk Factors above. These inputs come in the form of smoothed / transformed market data inputs. The diagram below illustrates the process and provides some examples:



In the more specific example below, a 6-month Euribor curve is created from a 3-month curve and basis swap spreads using linear derivation functionality. The resulting 6-month curve is passed to a quant library to produce the discount factors (using a non-linear process).



Recording Models and How They Have Been Executed

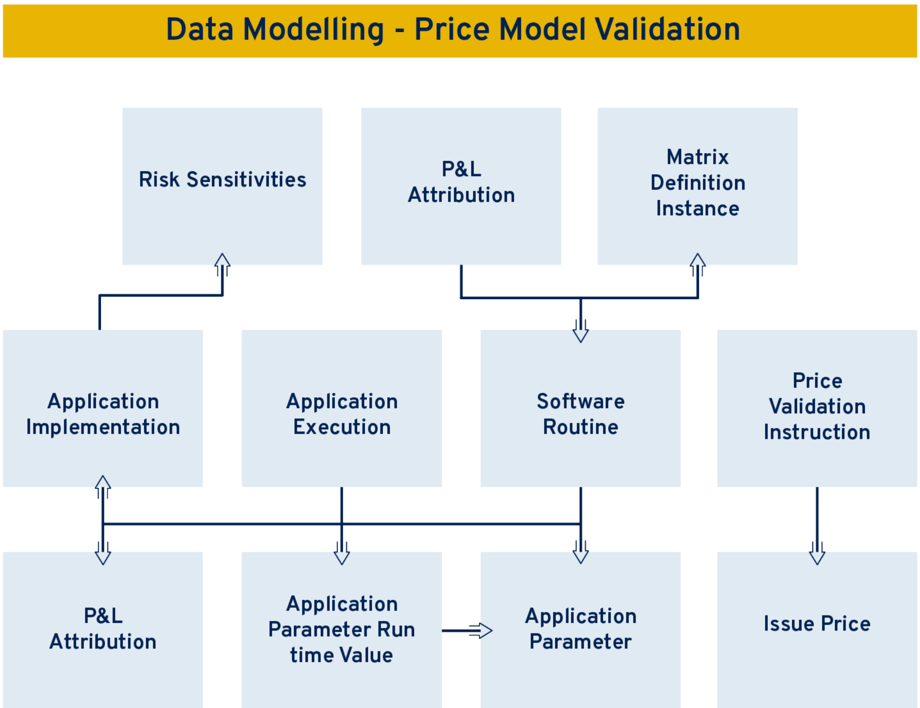
A compliant modeling ecosystem will be able to record definitions, inputs, outputs and approval statuses of "Pricing Models". As described in the introduction, a common and well known example of a model is "Black-Scholes" – which is the industry standard model for valuing European call and put options. Other examples are SABR, Black's Model, simulation-type models (e.g. Monte Carlo) and Base Correlation. Models are used to calculate the prices/values of positions in securities or derivatives. Models can be thought of as the implementations of the mathematics required to calculate fair value valuations. A simple example of a pricing model is V=P*Q. Where V= the valuation of the security, P is the price of the security and Q = the amount of the security the bank is holding. When calculating the V for derivatives, e.g. call options, however, you need to take account of the fact that the option will only have a value if the P (P in this case would be the Price of the underlying Security) is greater than the strike. So, you need to know the probability of P being greater than the strike at maturity – so the model becomes more complicated, requiring interest curves and probabilities as inputs. Standard Deviations, which are used to measure the riskiness of the underlying prices, are derived using market volatilities of those prices.

The data modelling for pricing model validation needs to be able to capture

- The definition of pricing libraries
- The definition of parameters used by pricing libraries
- The definition of the market data inputs used by pricing libraries

- The recording of run-time values associated with pricing libraries
    - Market data run-time values
    - Parameter run-time values
- The recording of the results of pricing library calculations
- Run-time values for parameter calibration processes
- The recording of parameter calibration results
- The association of pricing routines with risk sensitivities generated by those routines
- The association of pricing routines with P&L vectors generated by those routines
- Approval statuses for pricing routines (model validation requirement)
    - Approvers
    - Date approved
- The products that are approved to be valued by the pricing library
    - Asset classes or instrument level approval
    - Model validation requirement

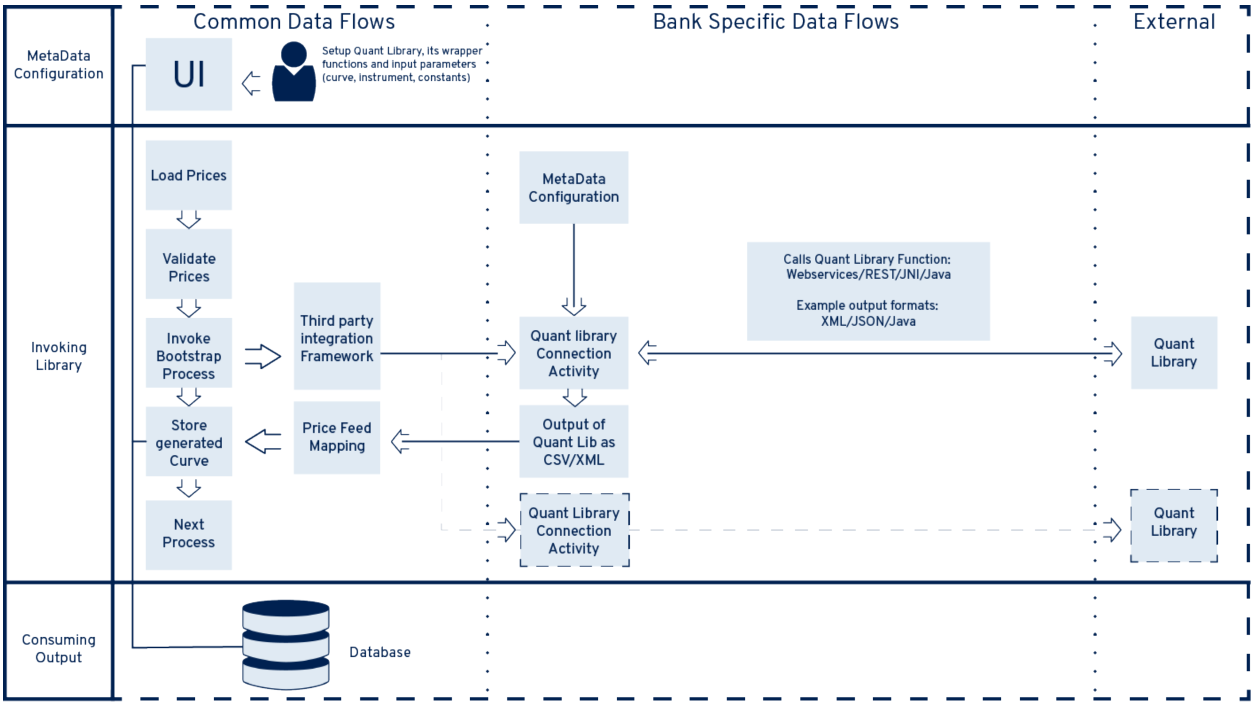
The diagram below gives an overview of the type of data that would be required for modelling pricing model validation



The role of model validation is increasing in importance as market risk, counterparty risk, pricing and XVA models are subject now to an onerous set of regulations that require more monitoring and control. This increased scrutiny means Model Validation teams have to validate more models more quickly and more accurately than ever before. Having a model governance framework in place that supports the various data lineage and audit requirements that come with model validation is essential for successful model validation.

The Integration of a Quantitative Library with the Firm's Data Management Systems

As described above, a well-defined model validation framework will allow users to easily maintain the formal list of libraries, routines and their associated inputs. The implementation of these models happens when they are physically integrated within the banks end-to-end environment, wrapper functions, input feeds, REST/SOAP APIs, data formats (XML, JSON, CSV), data platforms (SQL, NoSQL), message queues, hardware, connectivity and data volumes. These are all part of the discussion for how the modelling ecosystem is physically implemented. The diagram below shows an example end-to-end flow for a bootstrapping process



A Summary of Recent Model Risk Regulation

As described in the introduction, the four pieces of model risk regulation this paper will be focusing on are:

- **FRTB:** Basel Committee on Banking Supervision, Minimum capital requirement for market risk, January 2016, also known as "Fundamental Review of the Trading Book"
- **PRA Model Risk Management:** The Bank of England, Prudential Regulation Authority, Supervisory Statement | SS3/18, Model risk management principles for stress testing, April 2018
- **TRIM:** The ECB's Guide for the Targeted Review of Internal Models (TRIM)
- **Fed Model Risk Management:** Federal Reserve / OCC SR 11-7 "Supervisory Guidance on Model Risk Management"

Fundamental Review of the Trading Book (FRTB)

Banks need to have sufficient capital set aside to protect against adverse movements in market prices. This is often referred to as market risk capital. FRTB[2] defines the requirements for this. The regulation covers the following areas:

- The Trading Book and Banking Book Boundary
- The treatment of credit risk in market risk calculations
- Calibration to stressed market conditions
- Movement of VaR to Expected Shortfall
- The comprehensive incorporation of the risks of market illiquidity
- The treatment of hedging and diversification
- The relationship of internal models with standardized approaches
- A revised Internal Models Approach
- A revised Standardized Approach

While all of these areas are important for banks that are subject to FRTB regulation, the primary question that most banks are interested in answering is:

- Should they use the internal models approach (IMA) or should they use the standardized approach (SA)?

It is less expensive from a capital charge perspective for a bank (or trading desk) to use the IMA, i.e. each trading desk will be required to set aside less market risk capital under the IMA. However, from an infrastructure, systems and operational cost perspective, the calculation of capital under IMA is more expensive. Banks, therefore, will choose to go with IMA as long as the infrastructure costs of doing so are not prohibitive.

For a trading desk to be allowed to choose IMA over SA, two key tests need to be passed:

[2] See GoldenSource white paper for "Data Considerations for FRTB success" for further details of data lineage requirements for FRTB

- a VaR back-test. This test was also required under previous market risk regulatory regimes (e.g. Basle 2.5). However under FRTB it will be more rigorously enforced.
- a P&L Attribution test

If a trading desk fails to pass either of these quantitative tests, it will need to revert to the Standardised Approach.

The FRTB VaR Backtest tests the 1-day 99% value-at-risk (VaR) measure against actual P&L (APL) and hypothetical P&L (HPL) over the prior 12 months. The Standard defines hypothetical P&L as the gains or losses that would have arisen from holding position quantities constant and simply applying market data moves to them over the period in question. An exception or an outlier occurs when either the actual loss (APL) loss or the hypothetical (HPL) loss of the overall trading book (cross desk) is greater than the daily VaR measure. VaR backtesting breaches for actual losses are counted separately from breaches for hypothetical losses; the overall number of exceptions is the greater of these two amounts. The regulation specifies backtesting zones, which will be allocated based on the number of breaches using a pre-defined approach, i.e.

Green: Back-testing breaches within tolerance => No impact on Capital Multiplier
Amber: Back-testing breaches outside of tolerance => Capital Multiplier increase
Red: Back-testing breaches => regulator potentially disallow IMA usage

As you can see, these zones are used as indicators to flag potential overall failures in the testing.

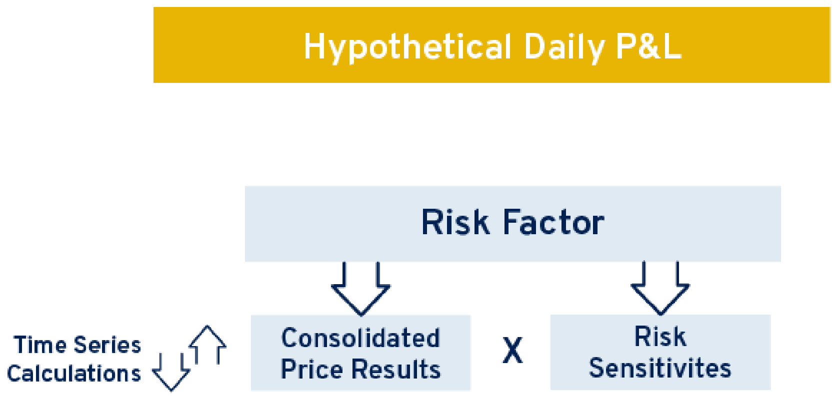
If any given trading desk experiences in the most recent 12-month period, either

- more than 12 exceptions at the 99th percentile or
- 30 exceptions at the 97.5th percentile in the most recent 12-month period,

then the standardized approach will apply. This is a hard rule.

The P&L vectors that are used in the Expected Shortfall calculation in the IMA approach are calculated by deriving price movements from the historical daily time-series of an instrument. These price movements are measured over rolling historical windows called Liquidity Horizons. If a risk-sensitivity approach is used in the Expected Shortfall calculation then, from a data perspective, the approach is similar to the Curve Shift methodology described above, the approach involves:

- mapping market data points to risk sensitivity points
- "shifting curves" by daily moves in market data and by amounts determined by the liquidity horizon
- applying the daily "curve shifts" to risk sensitivity positions



The time-series curve shifts applied to the starting risk positions will result in a distribution of hypothetical daily P&L values by risk factor. 1-day 99% confidence intervals (CIs) can be determined using the distribution. An outlier occurs whenever a historical daily hypothetical P&L is greater than the VaR at the 99% CI. In a year with 260 trading days, at the 99% CI the bank would expect to have no more than 3 back-testing breaks (rounded up from 2.6). If there are greater than 3, the risk management model will be deemed suspect and will require investigation and potential remediation.

It should be noted that this is a test of the VaR model and not a test of a VaR limit, i.e. it is a test of the VaR model that will be used to calculate a VaR risk measure. In fact, the approach tests both the VaR model and the front office pricing models that are used to generate the P&L vectors that are used in the VaR calculation. This is because:

- Where risk sensitivities are used in the calculation they will have been created using the bank's front office pricing models
- There is no 'noise' in the calculations from position amendments, fees, commissions etc., and, finally,
- The VaR CI is calculated using clean historical P&L values create by front office pricing models.

If the VaR model is accurate the number of back-testing breaches should be very close to the number that the model predicts

As mentioned above, a compliant FRTB ecosystem will need the ability to both:

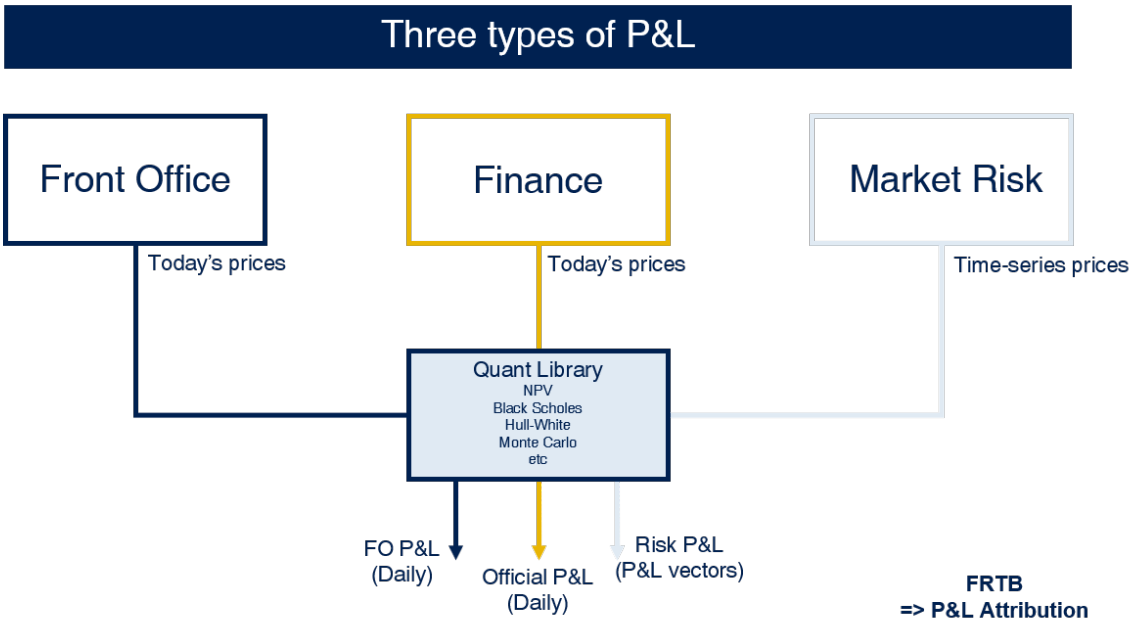
- Store all the time-series data required for the FRTB-IMA VaR calculations and
- Integrate with a model-validated calculation engine to support a parameterized approach to these calculations.

So, should a trading desk use risk sensitivities for its VaR back-test or should it use a full revaluation approach? FRTB is not prescriptive on this question. P&Ls used in the VaR back-test are calculated across all risk factors that the trading desk is exposed to, irrespective of whether a full revaluation approach or a risk-sensitivity approach is used. If a desk uses a risk sensitivity approach then, for example, a 10-day P&L will be generated for each risk sensitivity that the desk is exposed to and the resulting P&Ls will be aggregated into a desk level P&L. If, on the other hand, the desk uses full a revaluation approach for VaR then the full revaluation model will take as input the market data for each risk factor in the model and use this market data to generate fair values for the desk's trades/positions. The difference between this value and the value generated from the market data from 10 days ago will be the full revaluation P&L. A full revaluation approach is always more accurate than a risk sensitivity approach but the full revaluation approach is also computationally more expensive.

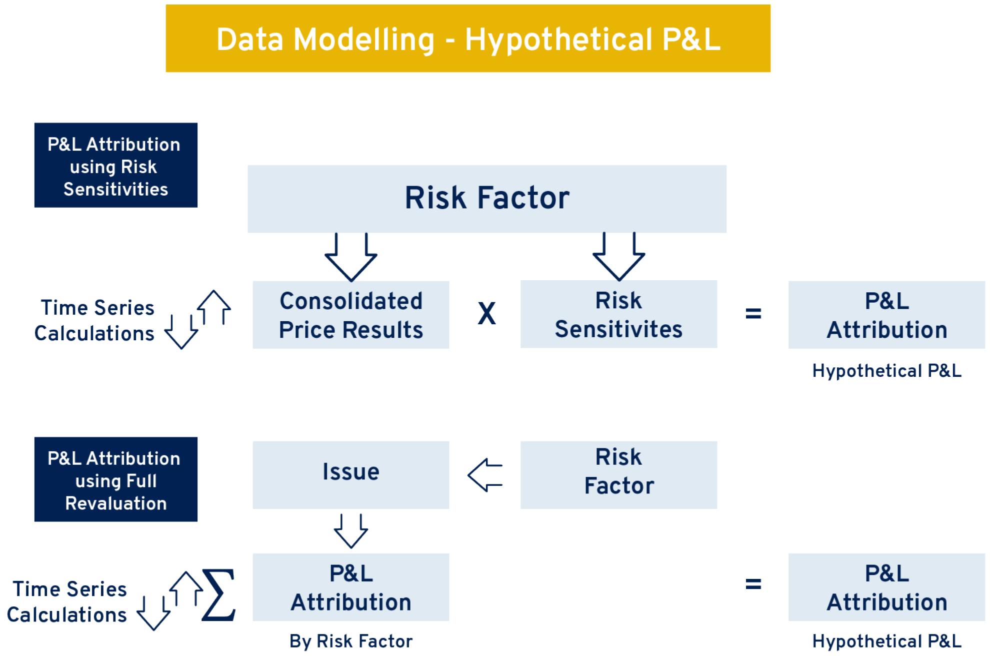
Dashboards and key risk indicators (KRIs) tailored to senior management requirements will be required to highlight any back-test breaches.

The P&L Attribution Test: FRTB is Driving an Alignment of Models and Data Between Risk and Finance

While VaR back-testing has always been a requirement, the need to pass a *P&L Attribution* test is a new requirement introduced under FRTB. Its overriding objective is the alignment between risk-calculated P&L and official P&L. The diagram below illustrates the concept. Due to the need for P&L to be approved and calculated daily by Product Control (a Finance function), Front Office and Finance P&L tend to be more aligned than Risk and Front-Office P&L. FRTB is, however, changing this.



The diagram below illustrates some of the data modelling concepts that are important for P&L attribution



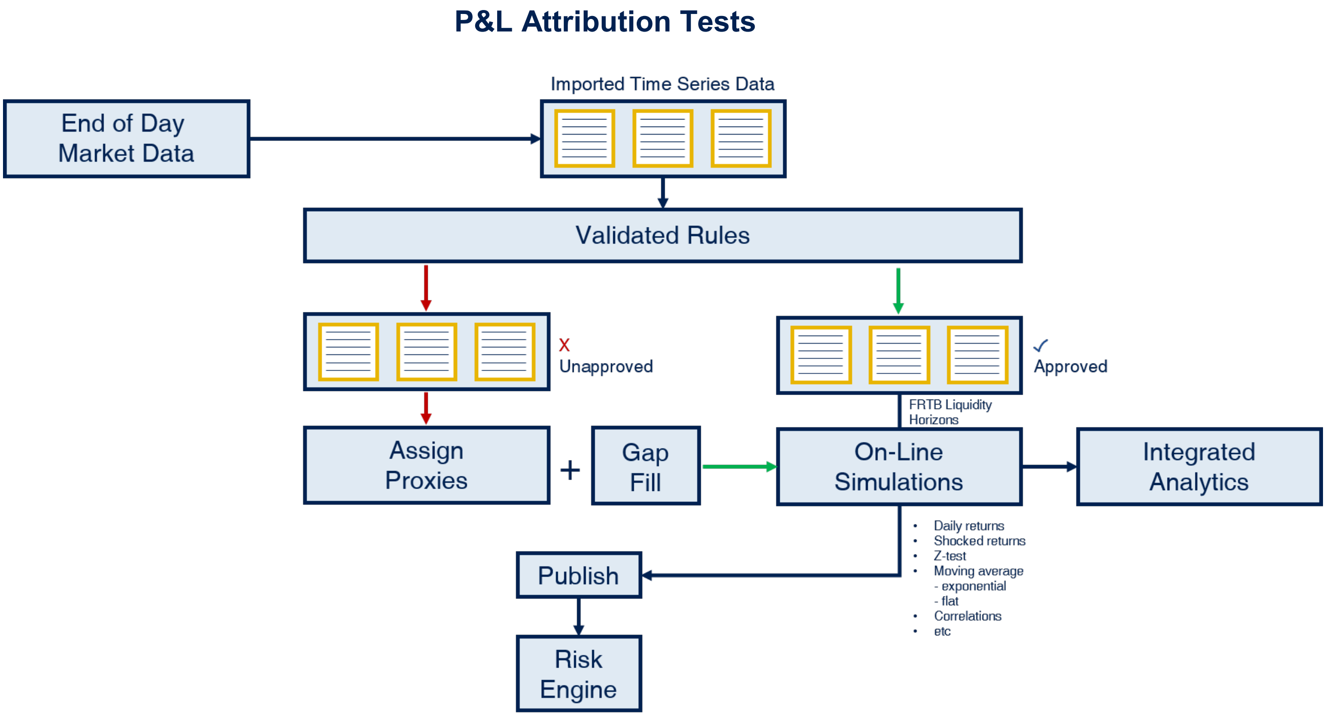
P&L Attribution in FRTB revolves around the concepts of Risk-Theoretical P&L and Hypothetical P&L. Like the VaR back-test, it is a test that is applied at trading-book level. Risk-Theoretical P&L is P&L calculated in Market risk models. Hypothetical P&L is the Front Office or Official P&L over the same period assuming that positions are held constant and market moves are applied to them. From a regulatory perspective, the aim of the P&L Attribution test is to test that risk models used for capital calculations are closely aligned to front office models (i.e. models used for official valuations which are in turn used in Product Control for official daily P&L purposes). There are three broad drivers of divergence between the results of Risk and Finance models:

1. The definition of the risk factors that determine the market data inputs
2. The market data itself
3. The valuation methodology, e.g. Full revaluation vs risk factor sensitivity based approximation, reduced risk factor models, etc.

As mentioned above, with regard to the valuation methodology, FRTB does not prescribe that risk models use a full revaluation approach but one of the arguments being made by industry practitioners at the moment is that full revaluation will essentially be enforced by the Standard in that if the risk models do not use full revaluation then the P&L Attribution test will likely fail.

The PLA requirements are based on two test metrics: (1) the Spearman correlation metric to assess the correlation between RTPL and HPL; and (2) the Kolmogorov-Smirnov (KS) test metric to assess similarity of the distributions of RTPL and HPL. To calculate each test metric for a trading desk, the bank must use the time series of the most recent 250 trading days of observations of RTPL and HPL. For risk calculations to be auditable, valid time-series data is required alongside robust integration with risk engines and analytics tools. Providing the required lineage back to source data is also key.

The diagram below illustrates the concept.



PRA, Model Risk Management Principles for Stress Testing

The Prudential Regulatory Authority Supervisory Statement describes the model risk management practices that the PRA expects firms to adopt when using stress test models. Among these practices are

- model validation policies
- processes for the identification of risks associated with models
- control processes for the risks inherent in use of stress test models
- daily management of model risks

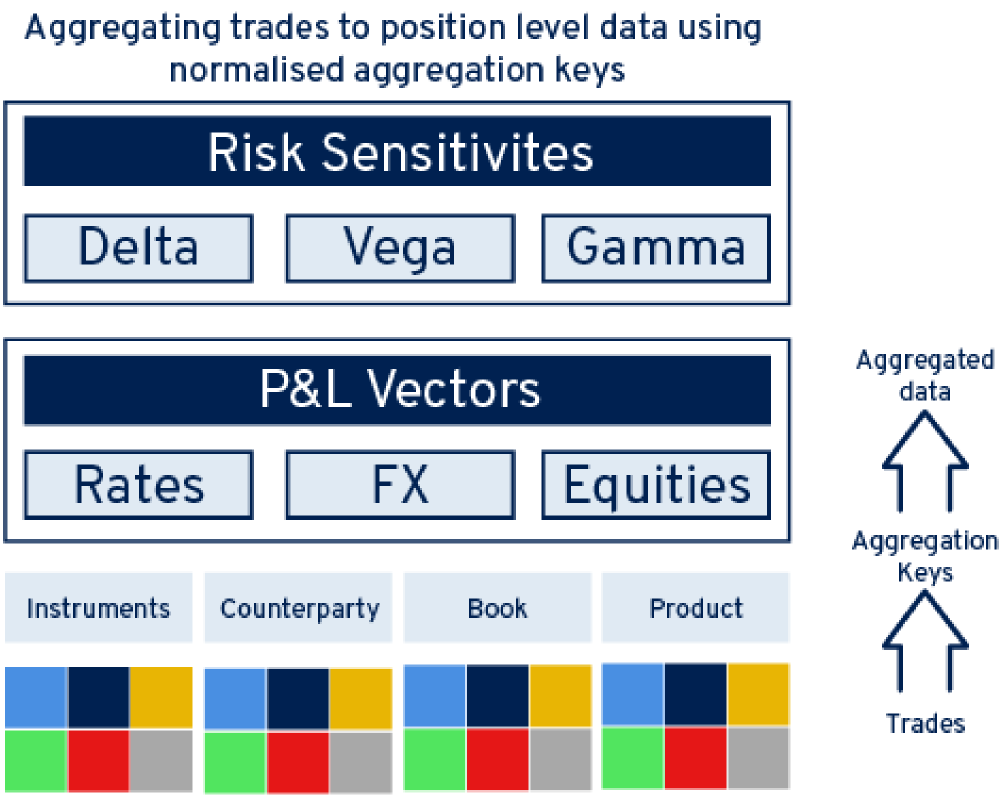
The Supervisory Statement lays out principles it expects a bank to adhere to. These principles cover the definition and maintenance of a model inventory, the governance framework required to manage model risk, the process for the development of new models, and the expectation of a robust Model Validation function.

The Supervisory Statement focuses on materiality. How much risk is inherent in the portfolio? What impact the movement of rates associated with risk factors will have on the bank's P&L or capital.

The PRA expects firms to adopt a risk-based approach to determine the materiality of models focusing on two aspects of size:

- number of positions in the portfolio
- exposure of portfolio (in P&L or capital terms) to changes in market rates

From a data perspective, aggregate position size can only be determined if data standards for aggregation keys have been adhered to, i.e. the netting of long and short risk positions (risk sensitivities) can only be done if the detailed positions and trades that are being aggregated share the same aggregation keys. It is worth showing the diagram below again here to emphasize the importance of this concept when it comes to model risk management.



As described, exposure size when it comes to derivatives is measured using risk sensitivities. The impact of a change in a curve over a period of time (a curve shift) is calculated by multiplying the curve shift by its corresponding risk sensitivity. The impact of a model parameter can be determined in the same way. The PRA wants all firms to focus their validation and independent review activities commensurate with the overall use, complexity and materiality of models (in terms of size, as described above) across the model life cycle to ensure those models that pose most significant risks - financial, capital or other - are adequately managed.

The Supervisory Statement describes a set of detailed principles that cover the ways in which it expects models to be classified, as well as details of the governance framework for model development, maintenance and governance.

The Supervisory Statement states that there can be four different types of models:

- **Statistical/Economic calculations.** E.g. impairment models, income models
- **Parameter transformation models**. E.g. probability of default models, scenario expansion models

- **Qualitative adjustments**. E.g. model reserves, model limitation adjustments that rely on subject matter expertise and judgement
- **Model output transformations**. E.g. the conversion of probability of default or exposure at default models into expected loss

A bank's model inventory should classify its models in line with these definitions. In addition, the development status of each model needs to be recorded, e.g.

- Under development
- Live
- Recently retired

European Central Bank (ECB), Targeted Review of Internal Models (TRIM)

TRIM is a (very large) set of rules and regulations issued by the ECB's banking supervision unit describing the central bank's expectations for banks within its jurisdiction with regard to internal models. The primary objective of TRIM is to bring consistency to approaches used by banks that use internal models to calculate risk weights.

TRIM focuses on the reduction of variability in risk-weighted assets (RWA) caused by poor or inappropriate modelling due to the flexibility in the approaches allowed by current regulation. And it does this in the context of the EBA's existing Capital Requirements Regulation (CRR). It attempts to develop a set of best-practice approaches to credit risk, market risk, counterparty credit risk and general issues related to model governance. It also aims to align as closely as possible with other regulations that prescribe approaches to internal models, e.g. FRTB and the EBA Guidelines on probability of default (PD) and loss given default (LGD).

Broadly, TRIM describes its expectations with respect to compliance with CRR for the following areas:

- General Topics
- Credit Risk
- Market Risk
- Counterparty Credit Risk

Internal Ratings Based (IRB) Approach Under TRIM

TRIM describes its guidelines for the harmonization of rules and regulations with respect to the Basel Internal Ratings Based (IRB) approach as it is implemented in the EBA CRR. The term Advanced IRB or A-IRB stands for advanced internal ratings-based approach. It refers to a set of credit risk measurement techniques proposed under Basel II.

Under this approach banks may develop their own empirical model to quantify required capital for credit risk. Banks can use this approach only subject to approval by NCAs (National Competent Authorities). Under IRB banks use their own quantitative models to estimate PD (probability of default), EAD (exposure at default), LGD (loss given default) and other parameters required for calculating the RWA (risk-weighted assets). Total required capital is then calculated as a fixed percentage of the estimated RWA.

Credit Risk Under TRIM

TRIM describes its guidelines for the harmonization of rules and regulations specifically with respect to the credit risk calculations that apply for IRB. The following topics are discussed:

- the type of data that can be used for estimating PDs, LGDs, EADs
- calculation of PDs
- calculation of LGDs

Market Risk Under TRIM

TRIM outlines a set of principles that apply to the calculation of market risk capital using internal models. The topics take into account the current requirements of the EBA CRR (Capital Requirements Regulation) as well as the principles of the expected revision to the market risk framework within the CRR arising from FRTB.

As mentioned, the principles prescribed by TRIM for internal models review cover:

- delimitation of the trading book and banking book
- treatment of structural FX positions created in the banking book
- the inclusion of Own Debt exposures in IMA calculations
- the identification of positions in defaulted debt
- collective investment undertakings such as ETFs and equity investments in hedge funds
- regulatory back-testing of VaR models, including:
  - the eligibility of hedge trades also used in CVA market risk calculations
  - the historical period used to perform back-testing
  - the definition of business days
  - the need for a policy describing how they calculate actual vs hypothetical P&L
  - the need for a policy describing the inclusion of valuation adjustments in P&L time-series
  - the rules & policies that should be adhered to when calculating actual P&L
  - the rules & policies that should be adhered to when calculating hypothetical P&L
  - the rules & policies that should be adhered to when calculating VaR back-testing overshoots
  - admissible reasons for withdrawing overshoot notifications
  - expected provision of analysis for VaR back-testing overshoots
    - Portfolio analysis
    - Market move analysis

- analysis of the internal model including
    - Appropriateness of the risk factors used
    - The modelling of risk factors
    - The suitability of the process for calculating VaR, hypothetical P&L and actual P&L
- methodologies for VaR and SVaR
- use of proxies, beta approximation & regression
- treatment of risks not in the model

Counterparty Credit Risk Under TRIM

TRIM describes a set of principles defined for the Internal Model Method (IMM) as part of the EBA Capital Requirements Regulation (CRR). The scope of the principles includes netting agreements, Credit Support Annexes, transactions-in-scope, trade coverage, simulated exposure scenarios, cash flow matching, pre-calibration for expected PV time-profiles, margin period of risk, Expected Positive and Expected Negative exposures.

Federal Reserve Guidance on Model Risk Management

Having been issued in 2011, this is the most mature of the four regulations described in this document. It is a joint Federal Reserve and Office of the Comptroller of the Currency OCC document. It is referred to as SR 11-7: Guidance on Model Risk Management

- Like both TRIM and the PRA guidance it covers: Model development, implementation and use
- Model validation
    - Conceptual soundness
    - Monitoring
    - Outcome analysis
- Governance, policies and controls

A Data Lineage Example

In the sub-section “Instruments and Market data” above, three categories of instrument identifiers were introduced.

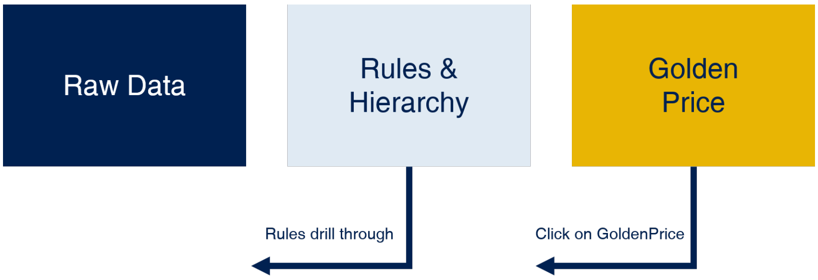
- Type 1: Securities or cash instruments
- Type 2: Quote instruments
- Type 3: OTC Derivatives

This section describes some instrument lineage requirements with respect to a Type 3 instrument identifier example. Type 3 instrument identifiers are identifiers for OTC derivatives.

Some important features of OTC instruments are:

- OTC instruments have a single, typically in-house generated identifier
- OTCs do not have a price, they instead have “mark-to-market” (MTM) or PVs
- OTC instrument identifiers can have data relationships to the Type 1 and Type 2 instruments that either form part of the OTC terms and conditions or are used to value the OTC

Lineage from the OTC instrument identifier to related Type 1 and Type 2 instruments as well as source data should be maintained at all times.



Instead of a single price, an OTC typically has multiple risk factors (each of which has a price). And even though the OTC has multiple risk factors, it has only one MTM value (and an MTM value is not a price). The MTM value (or PV) is instead a $ amount equal to the present value of the future cash flows of the OTC. To calculate the PV of the OTC, a model is required. The model takes each risk factor as an input and revalues the MTM each day. The following is an example.

<u>Type 3 instrument: OTC Equity Option</u>

OTC Instrument id=internalid1234

Terms: right to buy 100,000 shares of ISIN 12345678910 in 6 months at price $21.00

MTM today =$898,500

Risk Factors (per the mathematical definition of the model, e.g. Black Scholes):

- Risk Factor 1: the underlying equity
- Risk Factor 2: the interest rate curve
- Risk Factor 3: Volatility of the equity underlying

Risk Factor inputs to the model today:

- Risk Factor 1: the underlying equity
    - Price of instrument ISIN 12345678910* today (e.g. $20.56), i.e. the OTC option is slightly out-of-the-money today

- Risk Factor 2: the Interest rate curve of the currency of the equity (e.g. USD Libor)
    - Prices today of the 20 instruments** that make up the IR curve that is used to obtain the interest rates that are used to derive the discount factors (via a bootstrapping process usually) that are multiplied by the expected future cash flows of the option in order to calculate the net present value (PV) of the option (i.e. the MTM of internal id1234)
- Risk Factor 3: The volatilities of the equity underlying
    - Prices (volatilities) today of the 400 instruments*** that form part of the volatility surface of underlying =ISIN 12345678910. These volatilities are the market's estimate of the variance of the price of risk factor 1 (the price of ISIN 12345678910), i.e. the market's estimate of how likely the price of ISIN 12345678910 is going to finish in-the-money (and pay out the call option holder)

*Type 1 instrument. Has a price but can also be traded because it is a security.

** Type 2 instrument. These are swap rates for the c.20 maturity buckets (tenors) as provided by vendors and brokers. These swap rates cannot be traded. They are not securities and they are not OTCs. They are rates that are backed out of OTCs (IR swaps) that the vendor and/or broker observe in the market place

*** Type 2 instruments. These are volatilities for the c.400 strike/expiry combinations of the various OTC options in the underlying as observed by vendors and brokers. These volatilities cannot be traded. They are not securities and they are not OTCs. They are rates (volatilities) that are backed out of OTCs (equity options with underlying ISIN 12345678910) that the vendor and/or broker observe in the market place

Conclusions

Daily P&L, balance sheet and capital charges are directly impacted by the quality of the models that are used to generate fair values and capital charges. In the aftermath of the financial crisis of 2008, regulators have been increasingly focused on ensuring that the risks associated with the models that are used to produce these values are quantified and adequately risk managed.

This paper appraises four pieces of regulation that directly address the management of risk within models. The concept of internal models is important in much of the regulation that addresses model risk management. These models are deemed internal because they either partially of fully use existing in-house models to calculate regulatory capital. Regulators allow banks to calculate regulatory capital using these internal models only where certain criteria are met.

Data lineage is critical to effective model risk management. The ability to trace the inputs to models back to source and with a full audit trail of the various transformations that have been applied (derivations, calibrations, golden price rules, etc.), is seen as critical by auditors and regulators alike. Some of the key aspects of data lineage for model risk management are

- Instrument lineage
- Data normalization
- Aggregation keys
- The transformation of raw market data inputs into calibrated inputs that models can understand
- The linkage of market data points on curves and surfaces to risk sensitivity buckets
- The model approval process and ability to have APIs from data systems to internal quantitative libraries

With the advent of recent regulation on model risk management, black-box approaches to the production of fair values and capital charges are no longer acceptable. A data centric approach to model risk management is essential.