GoldenSource®

Risk Factors & Data Lineage in Derivative Pricing

XT25
Charlie Browne

Charlie Browne
Head of Market Data, Risk & Quant Solutions

June 2025

Agenda

1 **What is a Risk Factor?**

2 **A Conceptual Framework**: Observable versus Fitted Risk Factors

3 **The Market Data Journey**

- Where does market data come from?
- Where does market data go to?
- What happens to market data along the way?

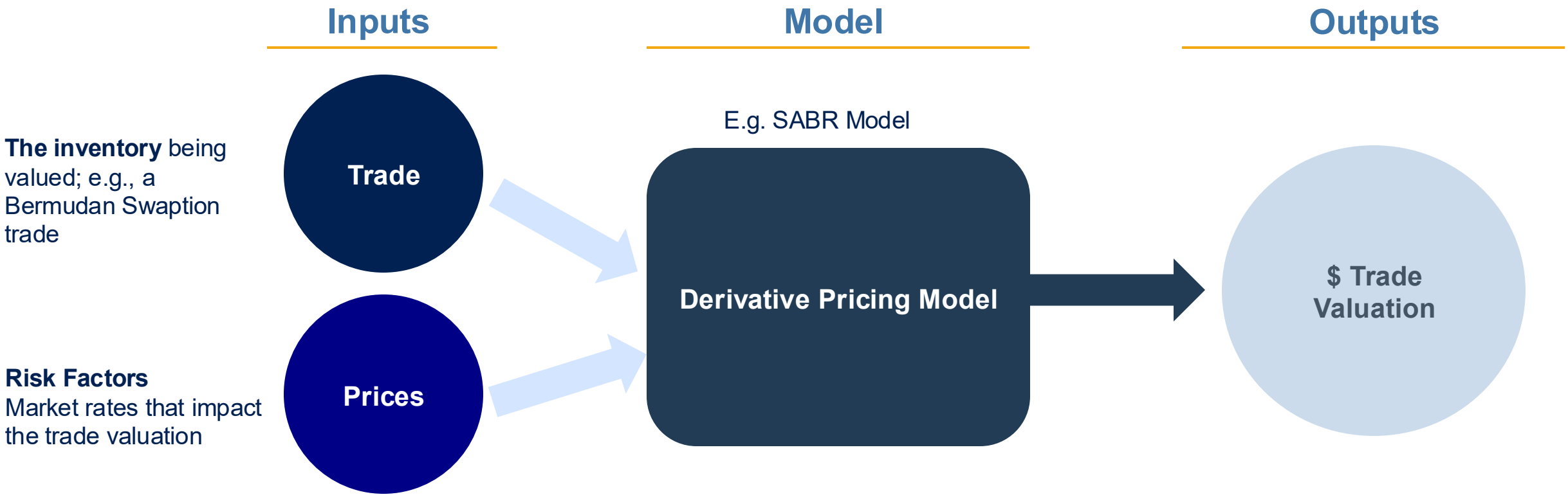
4 **Data Lineage**

- Regulation: BCBS 239 & RDARR
- Visibility and Transparency of the Journey

5 **A SABR Model Example**

What is a Risk Factor?

In the context of derivative pricing models (DPMs), risk factors are the prices (market rates) that are the models' inputs. A central objective of DPMs is to produce the $valuation of the derivative trade.



The diagram shows the inventory, e.g., an IR Swaption trade, and the market prices that determine the valuation of that inventory.

What does Regulation say about Risk Factors?

The main regulation covering market risk factors is FRTB. It has a section called "Specification of market risk factors." But is it really a specification?

FRTB **Fundamental Review of the Trading Book**

"FRTB is biggest change in market risk for two decades (Farag, 2017a)."

FRTB on Risk Factors

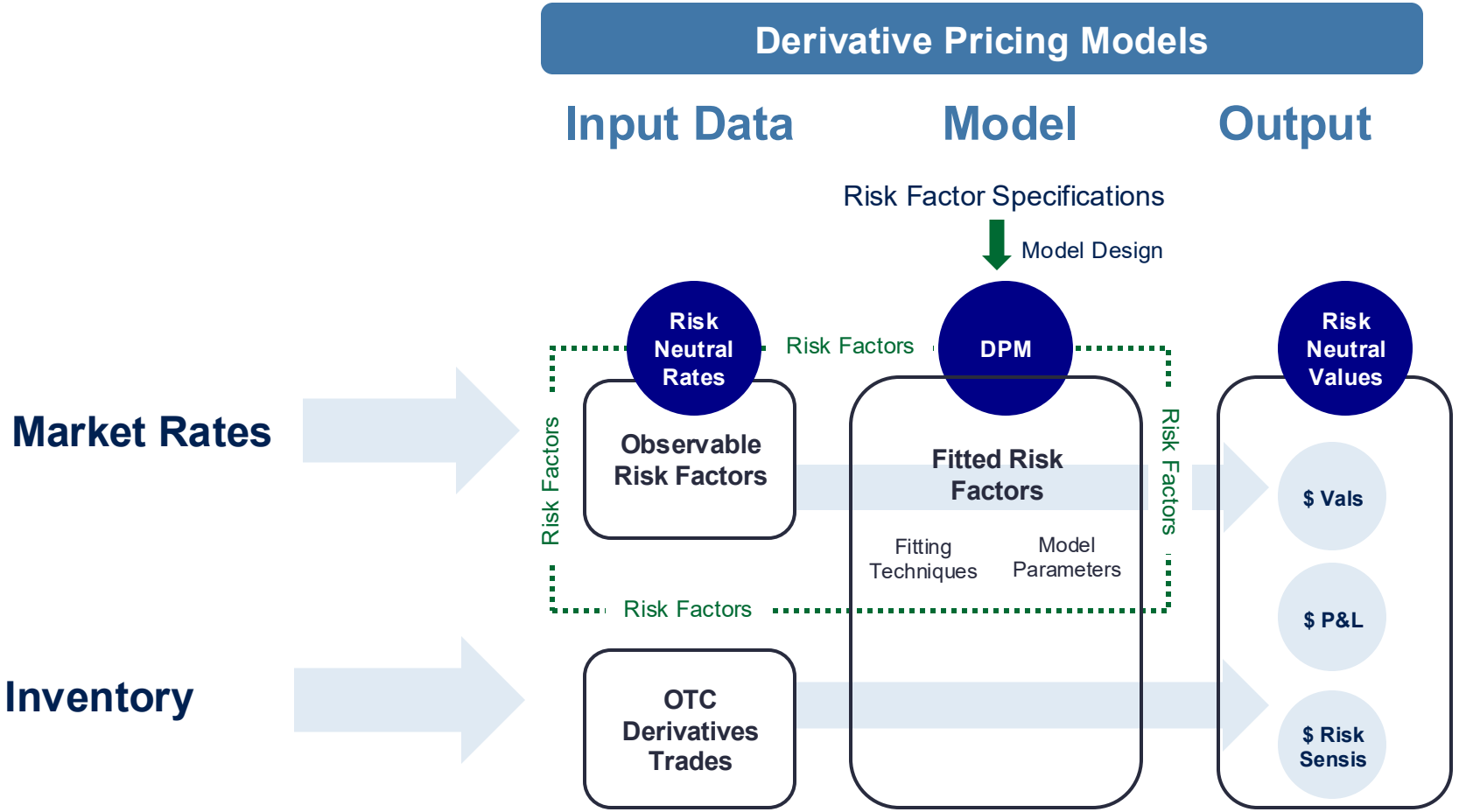
"One of most challenging area of FRTB is Risk factors (Aresi and Olivo, 2017)"

Is this really a specification?

Regulatory Document: **FRTB**

Conceptual Framework: A Theoretical View

Derivative Pricing Models, like all models, have inputs and outputs. There are broadly two types of inputs into DPMs, 1) the inventory, i.e., the derivative that requires valuation, and 2) the market rates that act as the model's risk factors



Thesis

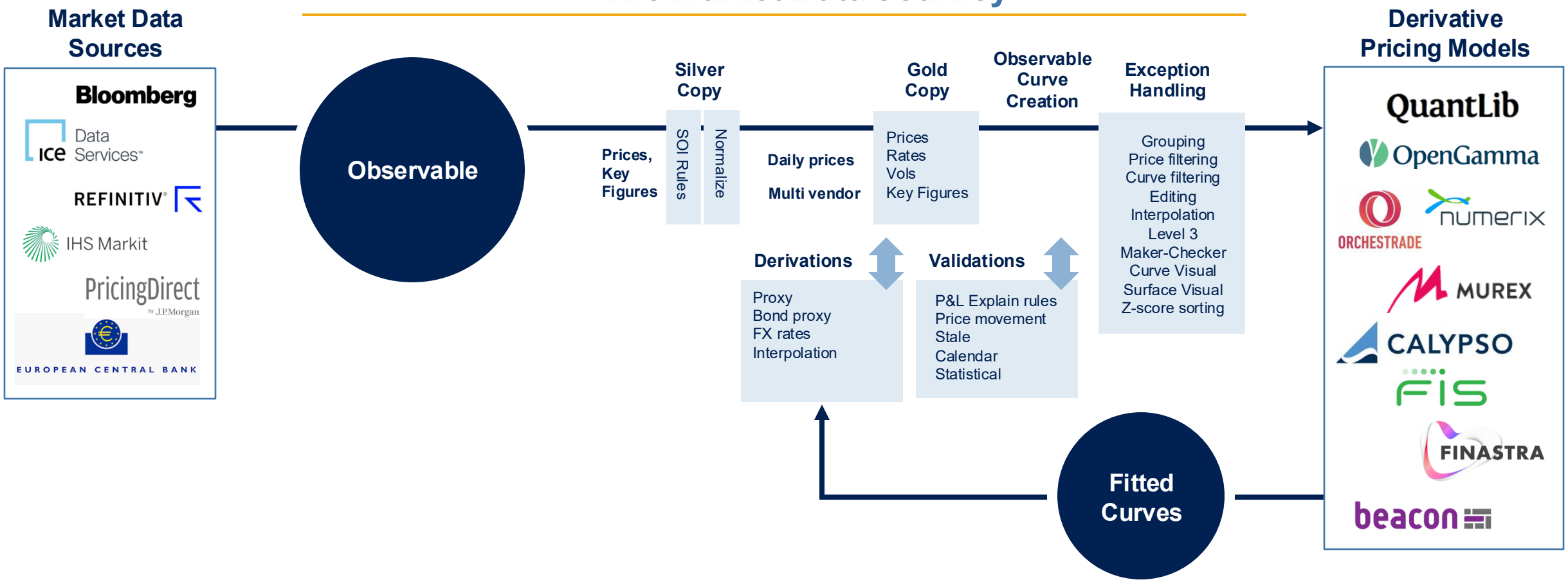
There are only two types of risk factors:

1. Observable
2. Fitted

Conceptual Framework: A Functional View

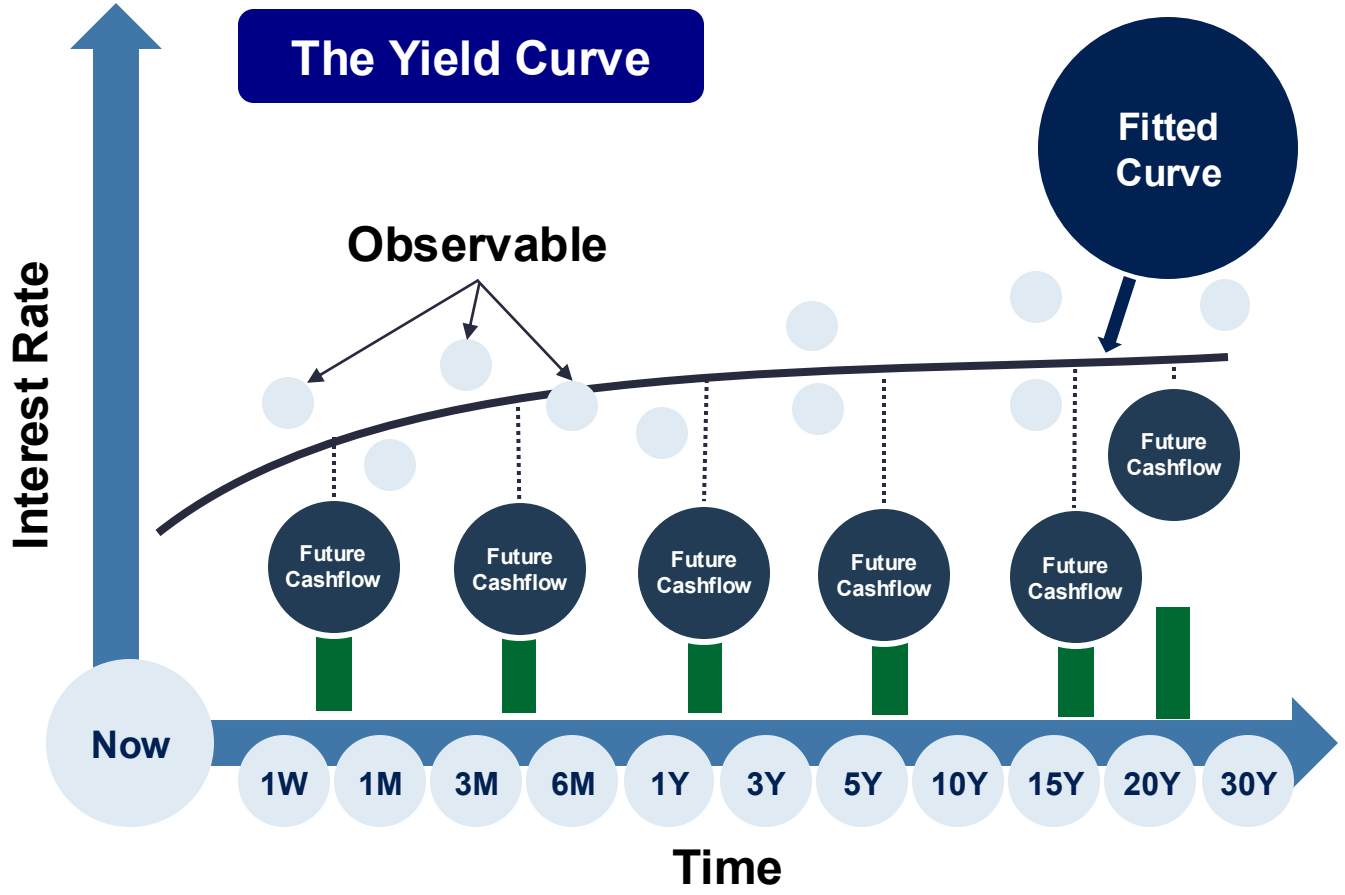
Derivative valuation libraries require validated, derived, and sometimes corrected observable market data.
One of the functions of the libraries is to convert observable curves into fitted curves.

The Market Data Journey



Case Study: The Yield Curve

The yield curve is a risk factor that is used by many derivative pricing models. Its observable market rates are obtained from traded instruments such deposits, futures, FRAs, and swaps. A smooth continuous yield curve is "built" from the observable data.

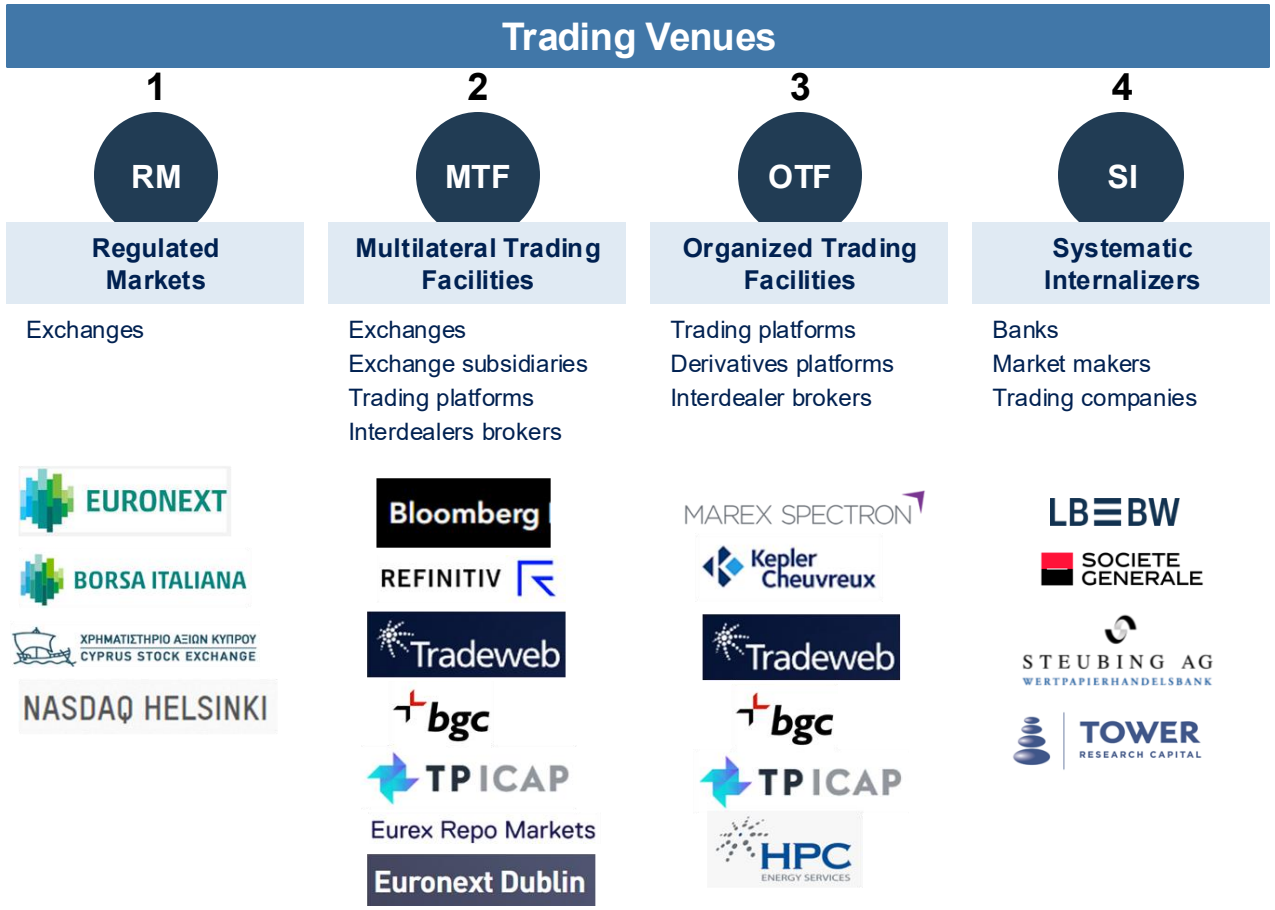


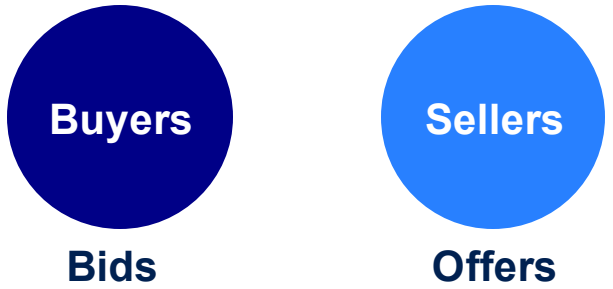
- The observable instruments on a curve are typically highly correlated
- Their market behaviours are similar
- So, while they are different instruments

=> they can be treated as the same risk factor

Observable Risk Factors: Where are they created?

Market data is created on trading venues. MIFID II defines four types of trading venues. For lineage and transparency purposes, it is critical that the bids and offers created on these venues can be analyzed using their associated reference data attributes.





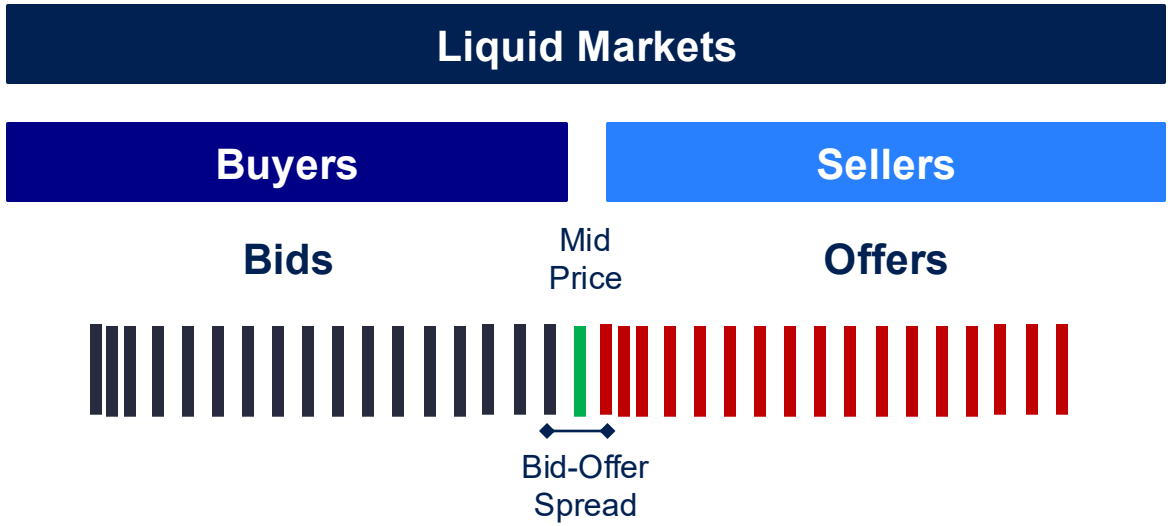
Reference Data Transparency

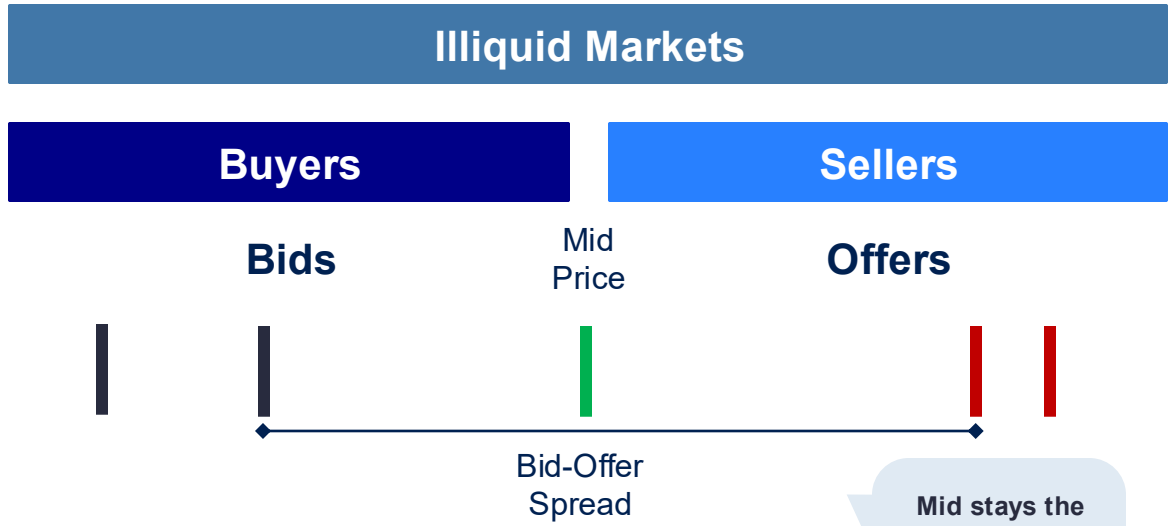
Instrument Identifier

Issuer	Asset Class	Market Conventions
Coupon	Maturity	Terms & Conditions

Observable Risk Factors: How liquid are they?

The volume of bids and offers on a trading venue for an instrument determines the liquidity of an instrument. Visibility into the underlying volumes and liquidity is required for regulatory compliance.





Mid stays the same but market becomes more expensive

Liquid
Short Dated Instruments
Mature Markets
Exchange Traded
G10 Countries

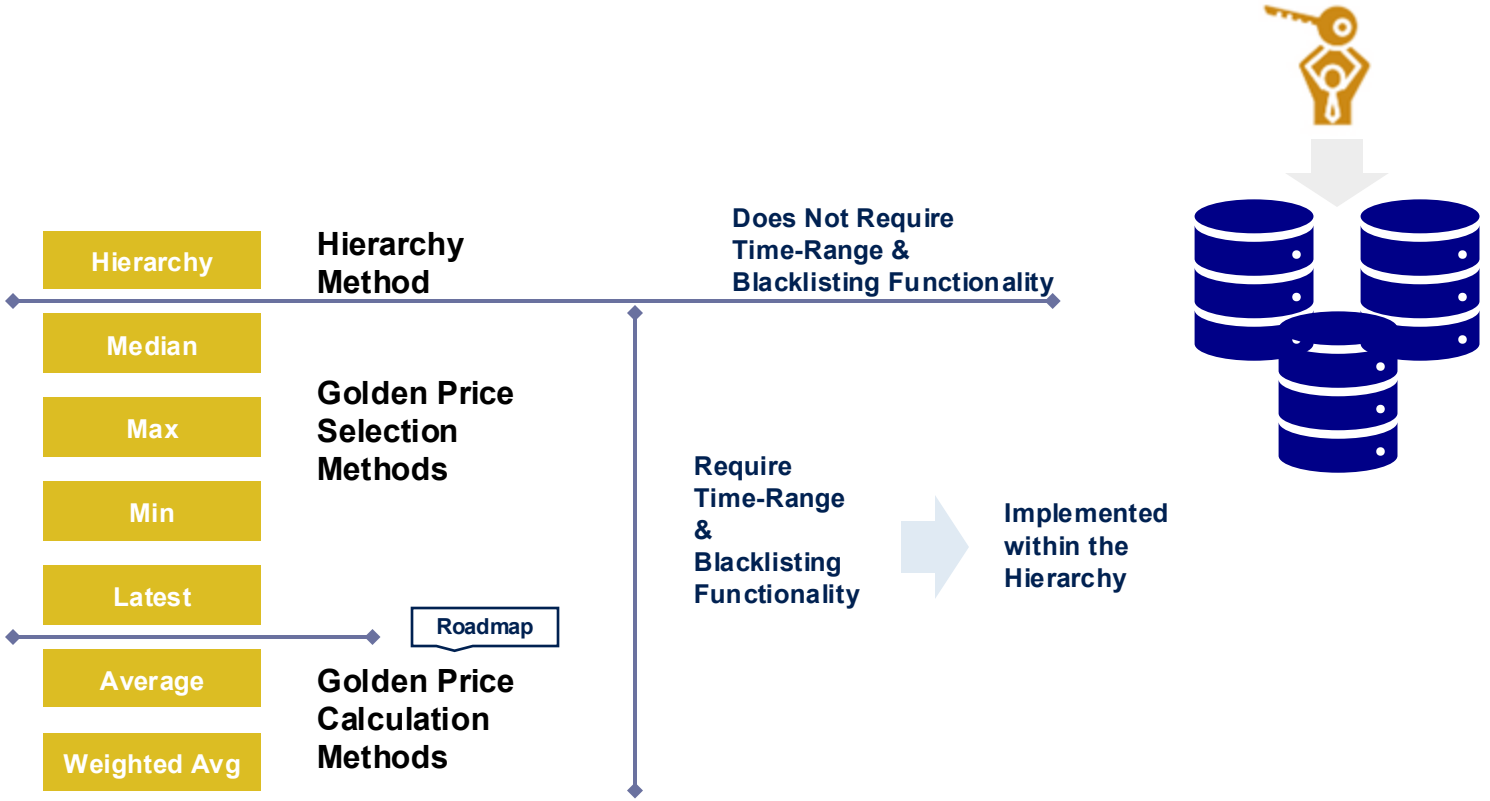
Examples
- US Treasuries,
- Gilts,
- Listed Equities,
- Interest Rate Futures,
- G10 FX

Less Liquid
Long Dated Instruments
New Markets
OTC
Non G10 Countries

Examples
- ABS
- Inflation
- Corporate bonds
- Commodities
- G20 FX

Observable Risk Factors: Golden Price Approaches

End of day or intra-day snapshots of market data require a gold copy concept when multiple sources of market data are available for each instrument. Several golden price approaches are possible. Transparency into the golden price method is required.



- Golden Price Selection Methods select existing one-from the existing set of candidate prices **to be golden**
- Golden Price Calculation Methods use the set of candidate prices as inputs to a **golden price calculation**

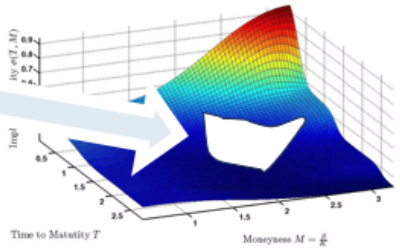
Observable Risk Factors: Derivations & Proxies

Proxy instruments are required where data is missing due to lack of liquidity. Spread curves, adjacent proxies and fixed income proxies are examples of derived and proxied data.

1 Adjacent Proxies

Rate Curves

Volatility Surfaces



Missing Prices

Time Series

% in Asset A	% in Asset B	Standard Deviation	ed
-50%	150%	4.3%	
-40%	140%	3.8%	
0%	100%	2.7%	0.2%
10%	90%	2.7%	0.3%
20%	80%	2.8%	0.5%
30%	70%	3.1%	0.6%
40%	60%	3.5%	0.7%
56%	44%	4.2%	1.0%
66%	34%	4.7%	1.1%
76%	24%	5.3%	1.3%

2 Fixed Income Proxies

Maturity	Geography	CCY	Rating	Proxy Bonds	Z-Spread	ASW Spread
2 Years	Holland	EUR	AA+	Bond 1	Z-Spread 1	ASW Spread 1
				Bond 2	Z-Spread 2	ASW Spread 2
				Bond 3	Z-Spread 3	ASW Spread 3
				Bond 4	Z-Spread 4	ASW Spread 4
				Bond 5	Z-Spread 5	ASW Spread 5
				Average:	ASW Spread	Z Spread

Where

Z Spread

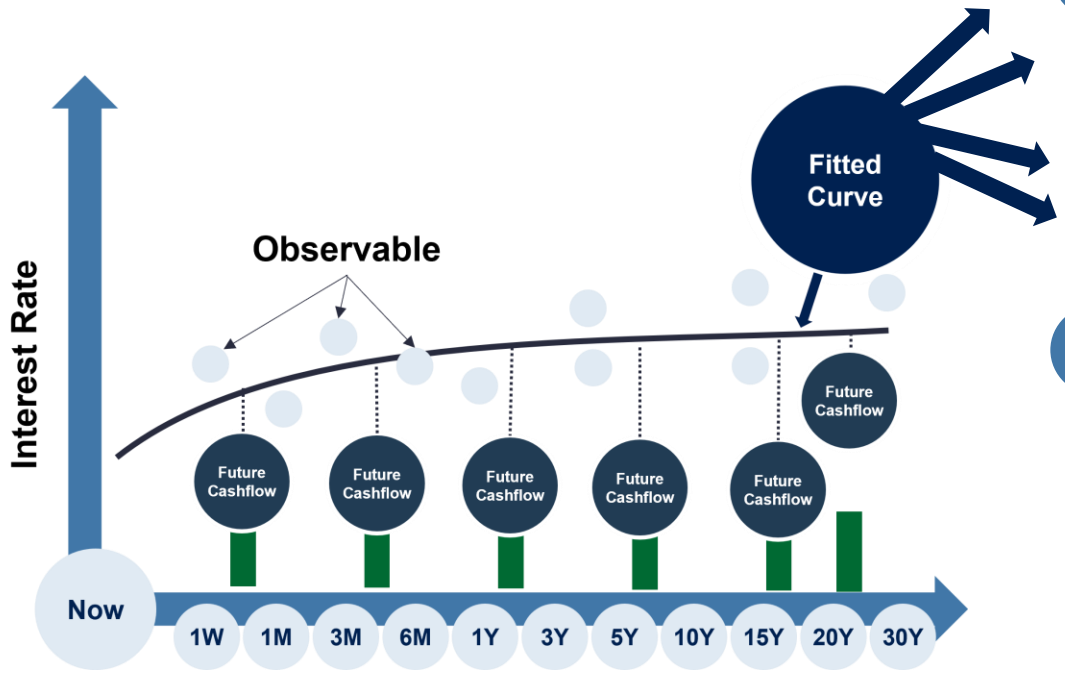
$$PV = \frac{PMT}{(1+z_1+Z)^1} + \frac{PMT}{(1+z_2+Z)^2} + \frac{PMT}{(1+z_2+Z)^3}$$

ASW Spread

$$100\text{-}P + C \sum_{i=1}^{N(FIXED)} z(t_i) - \sum_{i=1}^{N(FLOAT)} \Delta I \, (RFR_i + ASW) \, z(t_i) = 0$$

Fitted Risk Factors: How are they fitted?

The volume of bids and offers on a trading venue for an instrument determines the liquidity of an instrument. While typically only say the bid, mid & offer are required by valuation and risk systems, visibility into the underlying volumes and liquidity is required.



Fitting Approaches

1 Polynomial Splines

e.g., Cubic Spline

$$\partial(t) = 1 + at + bt^2 + ct^3$$

- OLS on discount function
- Piecewise polynomial functions => identify “knots”

2 Parametric

e.g., Nelson Siegel

$$f(m) = \beta_0 + \beta_1 exp(-\frac{m}{\tau}) + \beta_2 \frac{m}{\tau} exp(-\frac{m}{\tau})$$

- Specify functional form
- Level, Slope and Curvature
- Parameters estimated with OLS

3 Bootstrapping

Various Methodologies

- Iterative starting at short
- Derive DFs from market quotes
- Interpolate
- Successively re-calibrate to market

4 Multi-Curve Root-Solving

- Numerical approach to curve construction
- Solve for DFs that reprice market quotes
- Solving for multiple concurrent curves

Observable & Fitted Risk Factors: Validation Methods

Different types of validation methods are possible depending on the requirement. Audit features are required which enables controllers to trace golden prices back to source data, highlighting the validations that were run

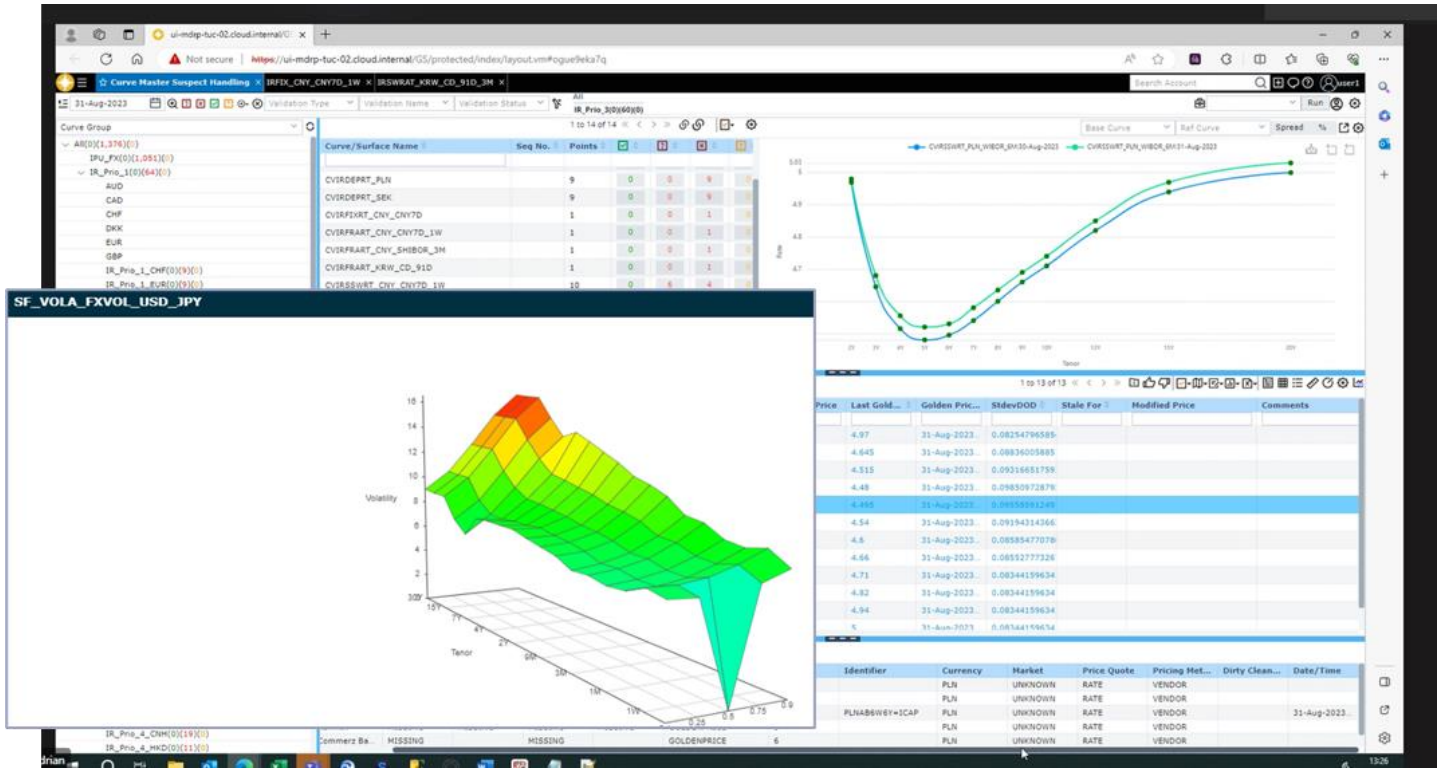
.

Parameterizable instrument level validations

- DoD movement checks (percentage, absolute, relative, standard deviation)
- Multi-vendor checks
- Zero checks
- Missing checks
- IPV Checks
- Calendar checks
- Time-range checks
- Stale checks
- Bid-ask spread checks
- Black-list check list
- Price age
- Time-series statistical checks
- Price level

Curve and surface level validations

- All points valid
- All points from same contributor
- Curve consistency check

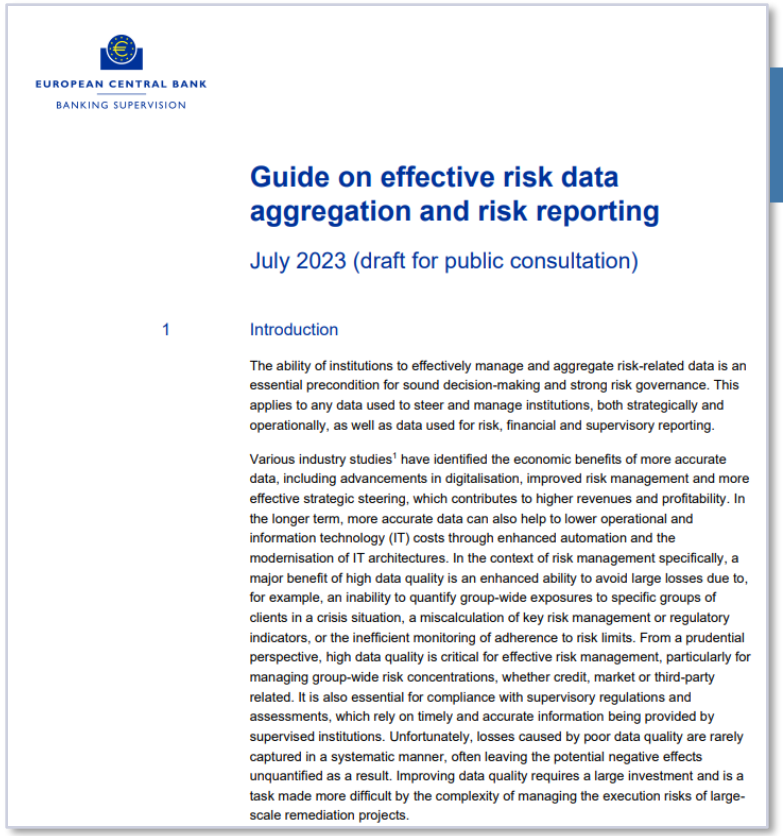


Data Management Regulation

Regulators increasingly focused on data management outcomes.
Data lineage and transparency are now central to audits.

European Central Bank Guidance

BCBS 239, RRDAR, FRTB



EUROPEAN CENTRAL BANK
BANKING SUPERVISION

Guide on effective risk data aggregation and risk reporting

July 2023 (draft for public consultation)

1 Introduction

The ability of institutions to effectively manage and aggregate risk-related data is an essential precondition for sound decision-making and strong risk governance. This applies to any data used to steer and manage institutions, both strategically and operationally, as well as data used for risk, financial and supervisory reporting.

Various industry studies[1] have identified the economic benefits of more accurate data, including advancements in digitalisation, improved risk management and more effective strategic steering, which contributes to higher revenues and profitability. In the longer term, more accurate data can also help to lower operational and information technology (IT) costs through enhanced automation and the modernisation of IT architectures. In the context of risk management specifically, a major benefit of high data quality is an enhanced ability to avoid large losses due to, for example, an inability to quantify group-wide exposures to specific groups of clients in a crisis situation, a miscalculation of key risk management or regulatory indicators, or the inefficient monitoring of adherence to risk limits. From a prudential perspective, high data quality is critical for effective risk management, particularly for managing group-wide risk concentrations, whether credit, market or third-party related. It is also essential for compliance with supervisory regulations and assessments, which rely on timely and accurate information being provided by supervised institutions. Unfortunately, losses caused by poor data quality are rarely captured in a systematic manner, often leaving the potential negative effects unquantified as a result. Improving data quality requires a large investment and is a task made more difficult by the complexity of managing the execution risks of large-scale remediation projects.

- Various studies indicate the benefits of accurate financial, risk data
  - Contributes to higher revenues, more profitability
  - Lowers operational and IT costs
- Enhances ability to avoid large losses
- It is Essential for banks to manage risk and finance related data

ECB Stakeholder, Meeting, 2023

RRDAR: “Data lineage” + “Data governance” + “Golden source of data” + “data validation”

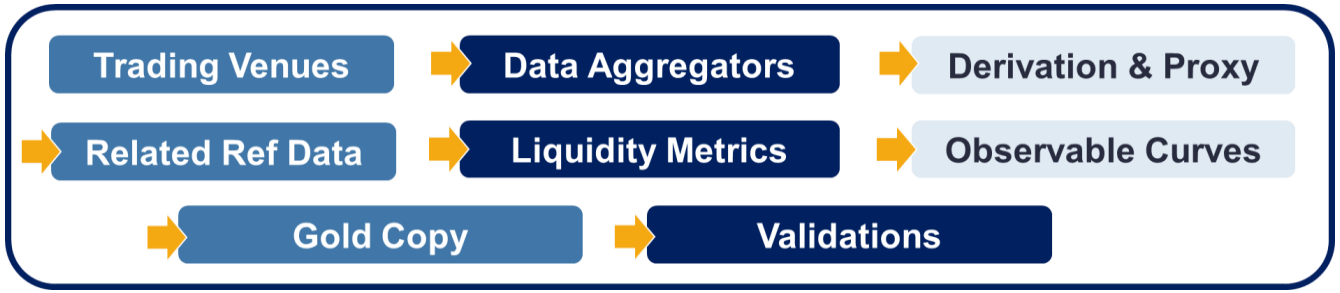
Data Lineage for Risk Factors: So, what is it?

Data Lineage is a reconstruction of the market data journey. The steps required to create the fitted risk factors that are used in the derivative valuations process need to be auditable and re-traceable. The data lineage process usually occurs in EDM systems.

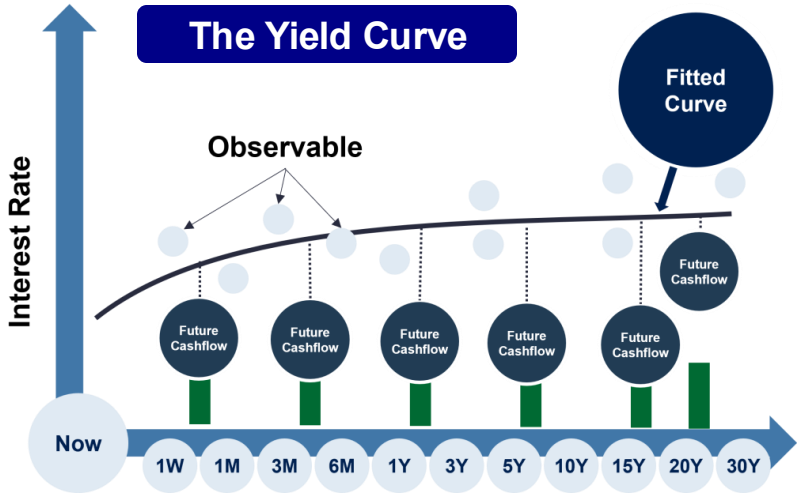
Observable

Model

Fitted

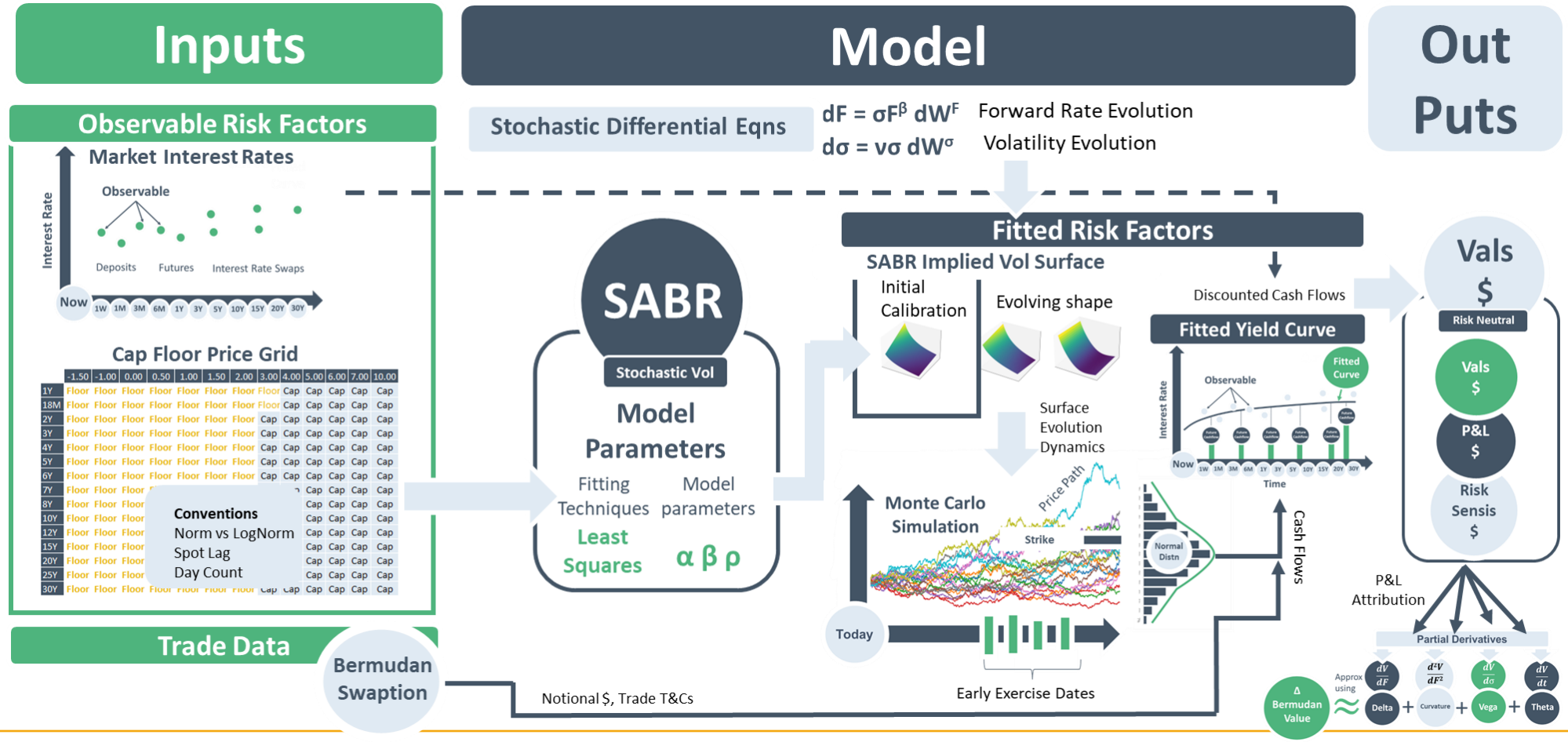


FIS QuantLib OpenGamma ORCHESTRADE numerix MUREX CALYPSO



A More Quantitative View of the Journey: The SABR Model

To price a Bermudan swaption, the volatility of the yield curve is also required. Stochastic volatility models like the SABR model below. Market volatilities are obtained from traded option prices and a model like SABR is used to show how those volatilities change over time



Questions

Thank You

For more details on the GoldenSource Sell-Side Solutions, please reach out to:

cbrowne@thegoldensource.com

Australia

Level 31, 120 Collins Street Melbourne, Victoria Australia

Italy

Piazza Vetra 17
Milano
20123

India

Raiaskaran TechPark, Office-2, Level-6, Towe II, Andheri Kurla Road, Sakinaka, Andheri (East), Mumbai – 400072, India

US

115 Broadway, 13th Floor New York, NY 10006 United States

UK

14 Bonhill Street 4th floor,
EC2A 4BX
United Kingdom

GoldenSource®