A Darwinian Theory of Model Risk

Claudio Albanese[1], Stéphane Crépey[2], and Stefano Iabichino[3]

February 19, 2020

Abstract

Performance assessment of derivative pricing models revolves around a comparative model-risk analysis. From among the plethora of econometrically unrealistic models, the ones that survive Darwinian selection tend to generate systematic short term profits while exposing the bank to long term risks.

This article puts forward an ex ante methodology to analyse this pattern for the broad class of structures, whereby a dealer buys long-term convexity from investors and resells hedges to be used for risk management purposes. As a particular case, we consider callable range accruals in the US dollar, a product which has been traded in size in recent years and is currently being unwound. We find 3d animations useful to visualise sources of model risk.

Keywords: model risk, structured products, short rate models.

JEL Classification: D81, G01, G13, G32.

Mathematics Subject Classification: 91B25, 91B30, 91G20, 91G30.

Contents

[1] Global Valuation, London.

[2] University of Evry in Paris-Saclay. The research of Stéphane Crépey has been conducted with the support of the Research Initiative "Modélisation des marchés actions, obligations et dérivés" financed by HSBC France under the aegis of the Europlace Institute of Finance. The views and opinions expressed in this presentation are those of the authors alone and do not necessarily reflect the views or policies of HSBC Investment Bank, its subsidiaries or affiliates.

[3] JP Morgan, London. This paper represents the opinions of the authors and it is not meant to represent the position or opinions of JP Morgan or their members.

1 The Darwinian Principles of Model Risk

This article develops a model risk framework for structured products sharing similar features. These are products where the bank buys volatility and convexity from investors with the purpose of reselling out-of-the-money options trading at a premium in the market, as these are used for risk management purposes. Such structures typically combine an upfront payment and a stream of regular coupon payments over a typically long lifespan. Model risk is then related to features such as optional callability or auto-callability which are inserted for the purpose of yield enhancement and without which these products could typically be statically replicated at no model risk. Examples include callable range accruals hedged by digital swaption streams (our case study focus), equity structured products such as autocallables and cliquets, power-reversal dual currency options, and target redemption forwards.

Quality ranking for models are of course subjective and a function of the utility of the model user.

To a user with interests aligned with the bank and a time horizon as long as the lifetime of his portfolio, a good model is one that enables a strategy whereby the bank makes a systematic profit out of selling out-of-the-money options at a liquidity premium in the market and hedges the risk exposure to the investor. To minimize the impact of model risk, one requires ideally a hedging strategy predicated on a model that is not subject to frequent and material parameter recalibration, as model risk is realized in the form of recalibration risk.

A user with a shorter time horizon however might prefer a model generating a stable stream of short-term profits. This shifted utility may motivate her to prefer a lower-quality model, often with the motivation that simplified models are also easier to implement and interpret.

From among the vast plethora of econometrically unrealistic models, the ones which stand a chance to survive and become broadly used have to satisfy two adverse selection conditions. The first one is related to competitive pricing:

First Darwinian Principle *A lower-quality model surviving the test of time must over-value the structured product at inception.*

An over-valuation at inception implies that the upfront payment paid to investor must be greater than the payment implied by other models. This makes the model competitive in the market, a necessary condition or else it would be disregarded as no trades would be closed.

The initial over-pricing results in systematic time-decay of valuations as the model is re-calibrated through time, a behaviour that makes realized discounted price processes deviate from martingales and thus reveal the existence of arbitrage opportunities on an ex-post basis. Hence, the aggressive initial valuation must turn into a mark-to-market loss for the user as reality unavoidably sets-in.

The first Darwinian principle raises a profitability puzzle: how the systematic alpha-leakage of a model can be sustained over long time horizons? Over-valuation is by itself a necessary condition for the survival of a model but is not sufficient. We need a second condition to guarantee survival of the "fittest".

Second Darwinian Principle *A surviving lower-quality model must over-hedge and harvest profits from excessive hedge positions which offset and surpass the systematic losses engendered by the initial over-valuation, in the short to medium term.*

To explain this principle with a simple example, assume for simplicity's sake that the physical (real world) measure equals the risk neutral measure $\mathbb{Q}$ of a good model and that a trader chooses to use a bad model $\mathbb{Q}^\dagger$ which over-values transactions from the viewpoint of the bank. To generate short term profits, she selects a bad model whose hedging strategy, when seen from the viewpoint of the good model $\mathbb{Q}$, results in selling far out of the money options and harvesting their premium when they expire worthless. This is the stereotypical proprietary trading strategy. The valuation process for the portfolio of excessive hedges follows a martingale process with a systematic stream of initial gains upfront which can last to the medium term followed by large occasional losses which, on average compensate and restore the martingale character of the process. However this is only revealed in a forward looking analysis. If the trader carries out a backward looking analysis, she will be able to claim a long consistent stream of initial wins which more than offset alpha leakage from the the derivative position.

This scheme works even better when the physical measure differs from $\mathbb{Q}$ in the sense that the hedges trade at a liquidity premium as they are used in the market for the purpose of macro-hedging. In this case, the initial profits are particularly attractive. However, in this case losses are also greatly amplified by a gamma trap event: when the trader realizes that assets are over-valued on books, that she has to call transactions and hedges are excessive, she will be prompted to increase issuance to re-use the existing hedges and will realise that any effort to hedge convexity moves the market against herself (cf. Devasabai (2019)).

In Section A we formalise the two Darwinian principles in a mathematical stylized setup.

If a model satisfies both Darwinian principles, then it stands a chance at survival and to become an accepted market standard for modelling a particular derivative product type. From an executive compensation perspective, model risk is particularly insidious because better models typically have lower short term profitability levels and far less long term risk. To a short-sighted decision maker, a lower-quality unrealistic model might be preferable as its usage could be of help to extract short term wealth. Ex-post statistical analysis is quite deceitful for this sort of statistical analysis, unless large losses have already occurred and one has all the data to carry out a post-mortem study. The ex-ante methodology for model risk in this article is based instead on state space analysis of possible outcomes over the lifetime of the trade and is capable to distinguish between good and bad models in the light of the above Darwinian principles. Ex post martingale test for hedge effectiveness are very useful for tracking and as a post mortem forensic tool. The ex ante methods we propose rely on model comparisons and visual rendering of hedging strategies over the trade lifetime. Model risk is intimately intertwined with reverse stress testing. A model can not be considered as valid if it is known to break down on a path leading to a stress scenario. While in Albanese, Crépey, and Iabichino (2020) we focus on the discovery of stress scenarios, in the present article we are concerned with understanding the pattern that leads to blow-up under stress. We conclude that the preference for short-term profitability skews returns in stress conditions giving rise to large losses.

2 Alpha-leakage

What is typically referred to by model risk has numerous facets but its fundamental drivers are remarkably simple. The "Fundamental Theorem of Finance", established by Bruno de Finetti in 1932 and recently published in English as de Finetti (2017), states that in an economy without arbitrage, the universe of all financial contracts can be valued by generating a unique set of scenarios for all the relevant market risk factors and taking expectations of cash flows discounted to present time using a compounded riskless rate.

In such perfect markets, a bank could theoretically value all assets out of a unique, globally calibrated Monte Carlo engine. The Monte Carlo engine would evolve not only elementary risk factors such as short rates, foreign-exchange rates, and stock prices, but also the valuation of derivatives such as swaps and options. Future derivative valuations can be computed with nested simulations or nonparametric regression. The discounted value process stemming from future valuations of options is a martingale. We can simplify things further by valuing all assets in terms of the money-market account starting with one unit of account at valuation date, so that no discounting is required.

Asset prices in such universal Monte Carlo pricing engine do not necessarily reflect the unbiased forecast that one would make on the basis of historical time-series. It is quite possible that, for valuation purposes, a bank would assign a higher or lower likelihood to a particular scenario to account for their degree of risk appetite for that future state of the world.

Black and Scholes (1973) and Merton (1973) fundamental discovery, which lies at the foundation of modern banking, is that: if hedging is perfect, then risk adjustments are irrelevant. The argument is simple: if a delta-hedging strategy which replicates the payoff of a trade under all scenarios can be implemented, then whether or not these scenarios occur in the real world with the same probability used in the simulation itself or with an adjusted probability is immaterial.

Although we might consider that market risk is hedgeable, model risk could still materialize, as we do not quite know how to build the de Finetti pricing engine. Therefore, numerous mathematical assumptions and approximations have to be used. To assess model risk we have to go through equations and assumptions. To illustrate our point we need another mathematical theorem called the "Doob-Meyer semimartingale decomposition" (see Doob (1953) and Meyer (1962)). The Doob-Meyer theorem applies to a large class[1] of stochastic processes V_t, where t is calendar time. The theorem states that V_t can be uniquely represented as the sum of a martingale process M_t plus a finite variation and predictable process A_t, i.e. as

$$V_t = M_t + A_t,$$

where the process A is the difference between two nondecreasing adapted processes and A_t is knowable "right before t", a property intuitively close to autocorrelation. A typical path for the returns of A, unlike a Brownian motion, does not flip sign on a daily basis but either trends upward or downward (going up for a while, reverting for another while, before changing directionality again).

In an asset pricing context, and using the risk-free asset as a numéraire, V_t is the (cum-dividend, i.e. already perceived coupons inclusive) mark-to-market valuation of a derivative asset plus the value of all of its hedges. In perfect markets without model risk, valuations would be martingales and A_t would be zero. If we instead have model risk, then A is possibly non zero. Assuming A is not only predictable, but even absolutely continuous, then A is of the form $\int_0^{\cdot} \alpha_s ds$. The letter α is a reference to the "alpha" remainder in CAPM theory (see e.g. Feibel (2003)). Model risk analysis is about assessing the size of A_t either historically or on a forward-looking basis, and carry out a risk-reward analysis against the martingale term M_t.

A predictable term is quite troublesome from a risk management viewpoint because it is arbitrageable and, for this very reason, it can not be replicated or hedged in the traditional sense, i.e. by means of a portfolio of offsetting trades. If such portfolio existed, we could make consistent riskless profits by acquiring it. By virtue of our second Darwinian principle, we expect that low-quality models in prevalent use have a negative alpha, correlated long-time returns, and a tendency to blow-up at times of stress.

While predictability is bad news for a market maker wishing to hedge and reduce risk, it is a very good news for a skilled counterparty seeking to make consistent profits. The alpha leaked by a market maker is the alpha gained by the hedge fund counterparty which takes option positions and hedges them back with a model different from the

[1] The technical condition states that V is a semi-martingale, i.e. (essentially) non arbitrable if a price process, and that its running supremum absolute value process is locally integrable (see Protter (2004, page 30)), a mild integrability restriction for a pricing model.

model of the market-maker (see Dupire (2019)). Therefore, if an investor trades with a bank and delta-vega hedges a trade to maturity, they can intentionally generate an alpha of the correct sign to manufacture consistent gains over extended time periods. When they are on a winning streak (and the bank is on a losing streak), the strong autocorrelation between the $\delta A_t = \alpha_t \delta t$ (assuming a regular α) makes it very difficult for the dealer to stop-gap the losses.

VaR, Expected Shortfall or Stressed VaR models do not detect alpha leakages because they focus on short-time horizons and on moments of the second and higher orders for return distributions. Instead, model risk derives from the cumulative effect of daily recalibrations and feeds into the first moment of returns, giving rise to a negative alpha term in the Doob-Meyer decomposition. The time horizons over which we observe losses related to model risk are far longer than the typical time horizon for market risk metrics and are more closely related to long-run stress testing analysis.

3 Sources of Model Risk

The factors contributing to model risk are generally speaking deviations from martingality which lead to systematic, autocorrelated losses or gains. Root causes and remedies are many and varied and they can be organized in several categories.

3.1 The Model Risk of Strategies

Hedging strategies typically consist of the combination of a robust strategy, which is model independent (see for instance Carr and Madan (1998), Dupire (1993), Hagan (2002)) plus a model driven strategy which is subject to model risk. Examples of model risk sources are optionality clauses embedded in derivative contracts such as early termination as in American style options, where different models typically lead to different conclusions regarding optimal termination decisions. Typically, different models would lead to different conclusions regarding optimal termination decisions and risk sensitivities.

To properly assess the model risk related to hedging strategies, we focus on the total price process for the trade itself plus its static and model driven hedges. We also assume that optional exercise decisions are executed consistently with the corresponding model.

Assuming daily rehedging, on any given day we would value an instrument using a certain model specification calibrated to current market conditions. The following day, market conditions evolve and, in parallel, model parameters change.

As calendar time skips a day, (cum-dividend) prices follow a martingale evolution under the pricing model. The moment the model is recalibrated, valuations are affected and portfolio martingality might break. Following the Doob-Meyer theorem, the change in portfolio value can be decomposed into a proper martingale term plus a predictable return. In formulas (see also Figure 1), we obtain

$$\delta V_t = \delta N_t + \delta R_t + \delta A_t,$$

where $R + A$ is the Doob-Meyer decomposition of the valuation change due to model recalibration. The martingale return δN_t, related to the intrinsic risk factors that are

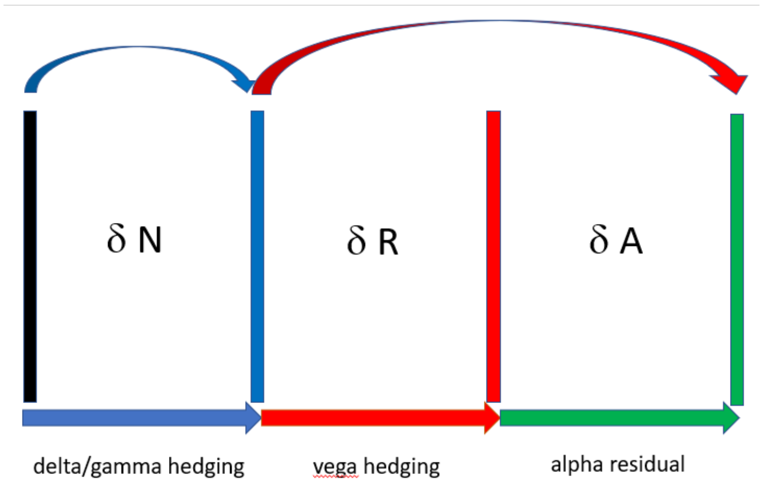

Figure 1: A recalibration step involves delta hedging with respect to intrinsic risk factors, and recalibration risk that can be handled by vega hedging with respect to model parameters (extrinsic risk factors) modulo an alpha residual.

explicitly modeled, is mitigated by delta and gamma hedging. The martingale return δR_t due to recalibration is handled by techniques such as vega-hedging, i.e. by offsetting sensitivities to model parameters (see Hagan (2002)). The predictable remainder δA_t instead captures model risk and can not be mitigated by dynamic hedging strategies, it can only be prevented by using models that embed the correct market views.

As a rule, rapidly varying hedge ratios for either delta or vega hedging suffer of transaction costs. Budget constraints on the cost of hedging limit the ability to reduce market risk and motivate an optimization exercise between the cost of risk capital and the cost of hedging. Although the fast variability of hedge ratios per se does not directly contribute to the predictable alpha term in the Doob-Meyer decomposition, it affects the cost of ownership of a model and it has to be part of a broader model quality assessment. However, unstable hedge ratios are also an indication of material alpha in the hedging process.

For the purpose of assessing alpha, we ignore transaction costs on hedging strategies and assume that, on each rebalancing date, the position of the selected hedging instruments is reset to the values indicated by the model itself. The process δA_t can be estimated historically as the hedging remainder and, since the diffusion term is explicitly offset by delta-vega hedging, the resulting estimate should exhibit predictability, autocorrelation, arbitrageability, and accumulate over long time periods.

3.2 Correlation of Alphas

A general question about capital reserves for model risk is whether they have to be apportioned at the portfolio level or at the individual instrument level in proportion to price uncertainties. In case one uses a globally calibrated model, the answer can only be that capital reserves are determined at the portfolio level, since the entire hedging set is regarded as a single instrument. If one instead uses locally calibrated models with parameter choices specific to each instruments, the answer depends on the extent by which alphas are mutually correlated. As locally calibrated models are estimated on a transaction-by-transaction basis, we are confronted with the correlation between alpha rates due to re-calibration. When taken in aggregate at the portfolio level, the mutual inconsistencies of locally calibrated models also break martingality and give rise to alphas.

Reduced-form models are particularly problematic because it is intrinsically impossible to promote them to the role of global models and use them at the portfolio level. An example of a reduced-form model is the SABR model (see Hagan, Kumar, Lesniewski, and Woodward (2002)), which models only one particular swap rate for a forward starting swap, and is thus not applicable to callable swaps or in full generality across fixed income derivatives. Copula models (see Li (2000)), which target a specific tranche of a CDO, are also narrow models, since they retain no information on the individual credit spread dynamics and they can not thus be used to model credit exposures holistically.

Since alphas are by nature strongly autocorrelated, one in general expects that they are also correlated between trades. The portfolio-level forward looking analysis with alternative models we propose reveals that α rates are state dependent and thus mutually correlated.

In this context, regulators have reportedly focused their attention in recent times on model interconnectedness, i.e. cross links between model calibrations which can possibly lead to correlated model risk losses (see for instance Devasabai (2017)). The related article on reverse stress testing by Albanese, Crépey, and Iabichino (2020) shows an example of model risk interconnectedness. This article and the current one revolve around the model risk embedded in economically unrealistic modelling assumptions, which for example are identified in the symmetric short rate process assumption of the Hull-White model. However, while the present article dwells on the model risk embedded in derivatives valuation, Albanese, Crépey, and Iabichino (2020) also include the model risk which can affect XVA metrics. In the case where one looks at the model risk impact of the Hull-White model on the XVA metrics generated by a portfolio of exotic derivatives, the two forms of model risk intertwine and interact non-linearly.

3.3 Modeling Realism and Risk Adjustments

A further contribution to the alpha term is related to the discrepancy between calibrated scenario probabilities and the risk adjusted probabilities one would estimate with an unbiased statistical method based on time series.

Replication strategies are typically devised in the pricing (risk-neutral) measure, not the real world measure. This works well in stylized situations such as a Black-Scholes world, where perfect dynamic replication is possible, so that it does not matter how the probabilities between the two measures differ. Whenever there is a major discrepancy between the risk-neutral and the risk-adjusted measures, the difference between the two measures will greatly amplify hedging errors.

The relevant mathematical result in this context is Girsanov (1960)'s theorem, according to which a risk-adjusted (univariate) diffusion process has the same volatility as the corresponding risk-neutral process but differs by the drift. In other words, if a portfolio valuation follows a diffusion process and we risk manage overnight the risk using the value-at-risk, then, whether we look at it from a risk neutral or a risk adjusted viewpoint, the market risk is nearly the same. However, differences emerge over medium to long time periods due to the Girsanov drift, which is a predictable correction of bounded variation similar to alpha remainders.

4 Example of Model Risk Analysis: Callable Range Accruals

A range accrual is a derivative product very popular among structured-note investors. The investor in a range accrual bets that the reference underlier, usually interest rates, will stay within a predefined range.

4.1 The Payoff

Consider a range accrual issued at time 0 with maturity T. A typical payoff at time $t \in [0, T]$ is $\Phi_t dt$ to the bank, where

$$\Phi_t = K_0(I_s^- + I_s^+) - K_1.$$

Here $I_s^- = \mathbb{1}_{\{I_s < R_0\}}$ and $I_s^+ = \mathbb{1}_{\{I_s > R_1\}}$, where $[R_0, R_1]$ is a range and I_s is an index process, typically of the form

$$I_s = \eta_1 \mathrm{SR}_s^{(1)} - \eta_2 \mathrm{SR}_s^{(2)},$$

where $\mathrm{SR}_s^{(1)}$, $\mathrm{SR}_s^{(2)}$ are two swap rates and $0 \le \eta_1, \eta_2 \le 1$ are constants. The swap rates are functions of the time s discounting curve and Libor rates. Hence $I_s^{\pm}$ are payoffs of European style digital swaptions in the index maturing at time s.

Some range accruals are callable by the issuer as, in order to entice the investor with a larger coupon, the bank reserves the right to terminate the transaction prior to maturity. From the investor viewpoint, termination could be acceptable since it will occur only in a state of the world where the bank has incurred in large losses, hence the investor has received excess revenues.

The time-0 valuation of a callable range accrual under the risk neutral pricing measure, with corresponding expectation $\mathbb{E}$, is given by (we skip the risk-neutral discount factors for notational simplicity)

$$\Pi_0 = \mathbb{E}\big[\int_0^T \mathbb{1}_{\{s \le \tau\}} \Phi_s ds \big],$$

where τ is the stopping time for the early exercise held by the bank.

Non callable range accruals are relatively simple to analyze as the issuer can statically hedge them by writing a set of digital swaptions struck at the corridor boundaries (see Figure 2). In this case, we have that $\Phi_s = K_0(I_s^- + I_s^+) - K_1$. Hence the valuation

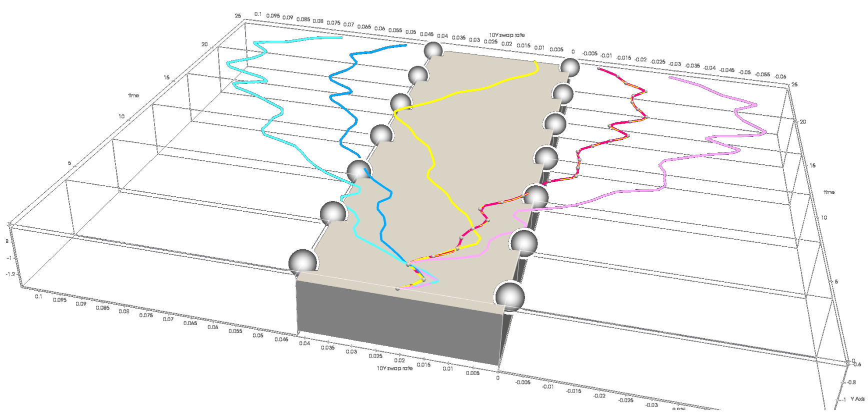

Figure 2: Range accrual corridor (the grey box), static hedges (on the sides of the corridor), and sample paths.

at the inception of the non-callable variant of the range accrual can be decomposed as follows

$$P_0 = \int_0^T \mathbb{E}(\Phi_s) ds = \int_0^T \big(K_0(D^+(s) + D^-(s)) - K_1\big) ds,$$

where $D^{\pm}(s) = \mathbb{E}(I_s^{\pm})$. The time 0 valuation of a callable range accrual instead is given by

$$\Pi_0 = \int_0^T \mathbb{E}\big[\mathbb{1}_{\{s<\tau\}} \big(K_0(I_s^+ + I_s^-) - K_1\big)\big] ds = \int_0^T \big(K_0(\Delta^+(s) + \Delta^-(s)) - K_1 \Xi(s)\big) ds.$$

Here $\Delta^{\pm}(s) = \mathbb{E}[\mathbb{1}_{\{s<\tau\}} I_s^{\pm}]$ are time-0 valuations of "Canadian style" digital swaptions (cf. Carr (1998)), which expire worthless in case the callable range accrual is exercised, and $\Xi(s) = \mathbb{E}\big[I_s^-\big]$. Assuming that digital swaptions are fairly priced, a quasi-robust hedging strategy for the callable range accrual in terms of European digital swaptions is given by the following static hedging ratios (constant over time t) in the digital swaptions with infinitesimal ds nominals (and the ensuing funding position so as to make the strategy self-financing):

$$a^{\pm}(s) = \frac{\Delta^{\pm}(s)}{D^{\pm}(s)}.$$

We say quasi-robust, because the time-t hedging ratios that would correspond to the same analysis as above, but starting from a time t in the future, are quite stable through time t. In practice, people use these hedging ratios plus a classical delta hedging strategy for the residual hedging error already "smoothed out" by the digital swaptions quasi-static hedge.

4.2 Champion and Challenger

For the model risk analysis, we compare two models, a champion and a challenger.

The champion is the Hull-White (HW) 1-factor model specified as

$$dr_t = \kappa(t)(\theta(t) - r_t)dt + \sigma(t)dW_t.$$

In the HW model, rates are not bounded from below and the classic Hull and White (1990) solver depends on this assumption.

The challenger is a model with two fundamental differences: rates are bounded from below and there are two factors instead of one (so that the steepness of the yield curve is stochastic). An example of such model is the stochastic drift (SD) short rate model of Albanese and Trovato (2008), where the short rate process is defined by an equation of the form

$$r_t = \phi(t) + \lambda(t)\rho_t. \tag{1}$$

Here

$$\phi(t) = \min(0, f(t) - 20\text{bp}), \tag{2}$$

where $f(t)$ is the infinitesimal forward rate, $\lambda(t)$ is a drift adjustment factor, and

$$d\rho_t = \kappa(\theta_t - \rho_t)dt + \sigma(t)\rho_t^{\beta}dW^{(1)}, d\theta_t = k(a - \theta_t)dt + \nu dW^{(2)}, \tag{3}$$

with $d\langle W^{(1)}, W^{(2)}\rangle_t = \gamma dt$. The resulting short rate model is very similar to a shifted 2-factor Libor market model.

To make the qualitative comparison significant, we select specific instances, sharing the properties of the HW and the SD model, of a common framework of discrete hidden Markov models (see Albanese and Trovato (2008)). Figure 3 depicts the probability densities in the resulting models, fitted to the same market data, as a function of time, from which the major differences between the two models can be observed.

We then run both models on a callable range accrual in the USD of maturity 25 years and with corridor $[0, 4\%]$. 3d animations are found the best way to visualise sources of model risk (see Albanese (2019) and the snapshots below). The conclusions of the ensuing comparison between the two models are summarized in Table 1. The

Valuations (cf. the first Darwinian principle)	The pricing function of the HW model dominates the pricing function for the SD model for times to expiry longer than 8 years because of a negative rates effect: when rates descend to deep negative values, a possible scenario in the HW model (but not in the SD model), range accruals become very valuable to the issuer. The mere possibility of deep negative rates raises the valuation of the HW model also in the positive range; but the alpha leakage is pushing down the HW valuation to the SD valuation as time goes on (see Figure 4).
Early exercise	Because of the negative rates effect, the optimal exercise boundary for the HW model arrives much later than in the case of the SD model (see Figure 4). We also observe that the optimal exercise boundary in the SD model is greatly affected by the steepness of the interest rate curve (see Figure 5).
Hedge ratios (cf. the second Darwinian principle)	Hedge ratios in digital swaptions struck at the corridor boundaries for the HW model at maturities exceeding 8 years are far larger than the hedge ratios for the SD model (see Figure 7). The reason is related to the delay with which optimal exercise decisions are signaled by the HW model.
Sensitivities	While the HW model shows that for rates within the corridor the gamma is mostly positive and thus dynamic hedging is relatively safe, the SD model indicates that the crossgamma matrix has a negative eigennvalue for most values of the state variable (see Figure 8). Once represented in terms of the 10y and the 10y-2y swap rate, the angle for the lowest eigenvalue varies from 30°to 70°as one goes across the corridor boundary (see Figure 9).

Table 1: Champion / challenger comparison.

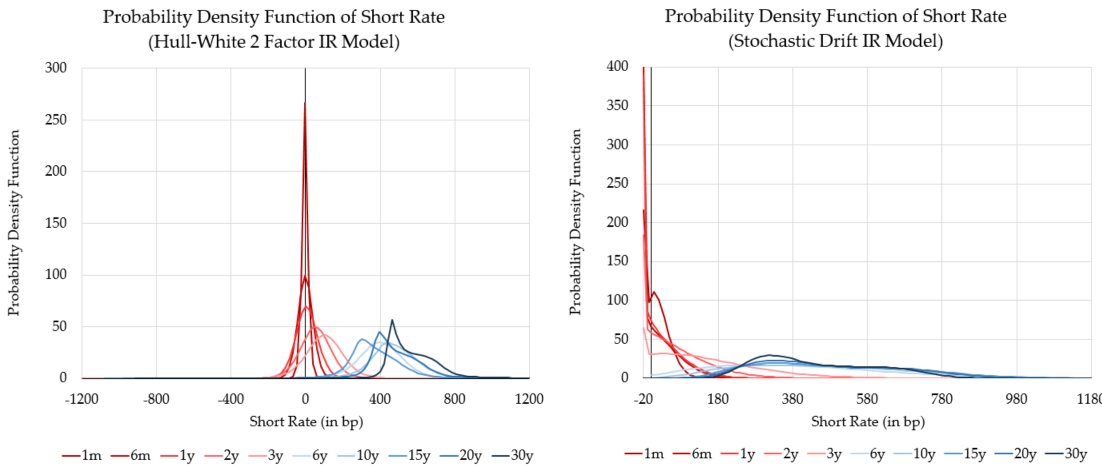

Figure 3: Probability densities of the USD short rate computed using the HW and SD models.

figures that follow show results in the form of three panels corresponding to the 290, 200 and 100 months to expiry, respectively.

In particular, Figure 6 is based on the Hull-White model and depicts a fairly enticing story from the viewpoint of the bank. The gamma is positive across most of the state-space and the option termination occurs quite late in the lifetime, allowing the bank to seize the benefit of carry of the hedges which are traded at a material liquidity premium. From the viewpoint of the SD model, however, the excessive hedges represent a financial risk (see Figure 10). The X-axis in this plot represents the 10y swap rate, while the Y-axis represents the 10y-2y swap rate spread. the Z-axis is the value of the portfolio of excessive hedges. Over the most profitable period for the trade, the risk profile due to excessive hedges exhibits a saddle point shape centered in the middle of the corridor and a pronounced steep, which signals potential losses in case of falling rates. This is a typical situation for cross-gamma risk, whereby one eigenvalue of the cross-gamma matrix is positive and the other is negative (see Figure 8). The direction of the negative eigenvalue is denoted with white filaments in the other 3d plots (see Figure 9).

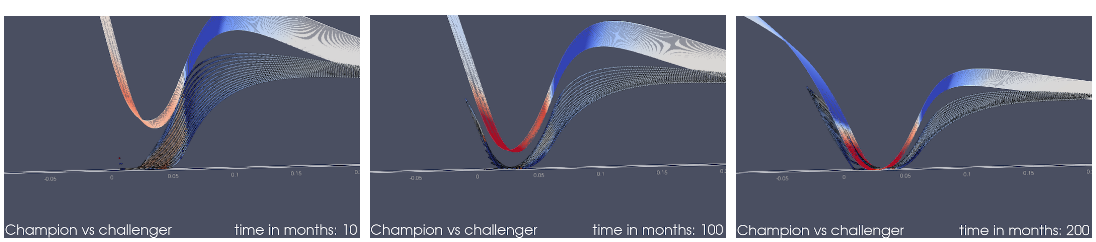

Figure 4: Comparison between the valuations of the challenger and the champion, color-coded for the gamma (red for positive gamma, white for intermediate, blue for negative).

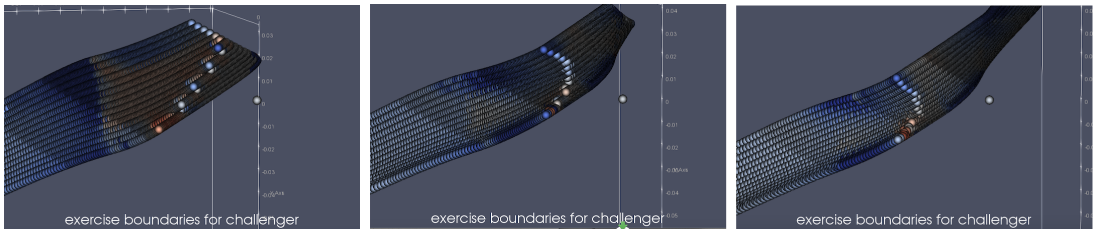

Figure 5: Exercise boundaries of the challenger.

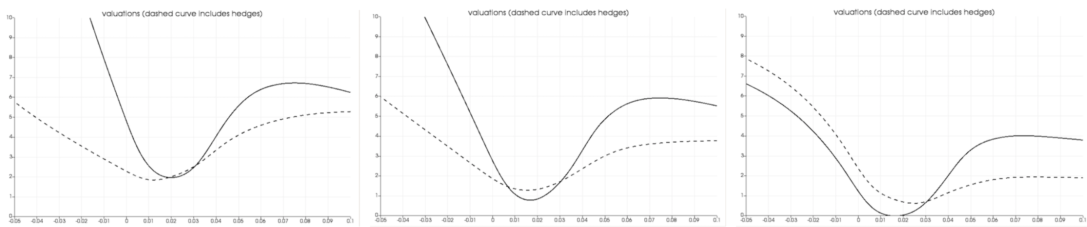

Figure 6: Valuations of the champion.

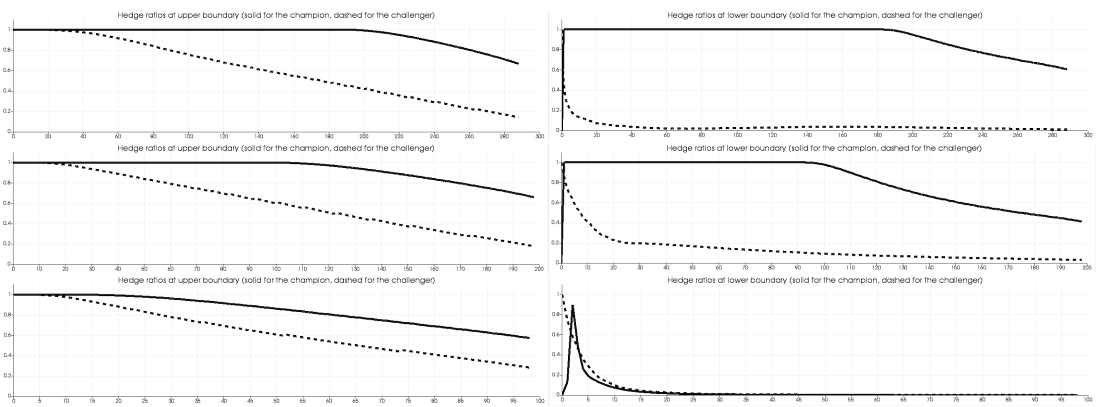

Figure 7: Comparison of the hedge ratios for the champion and the challenger.

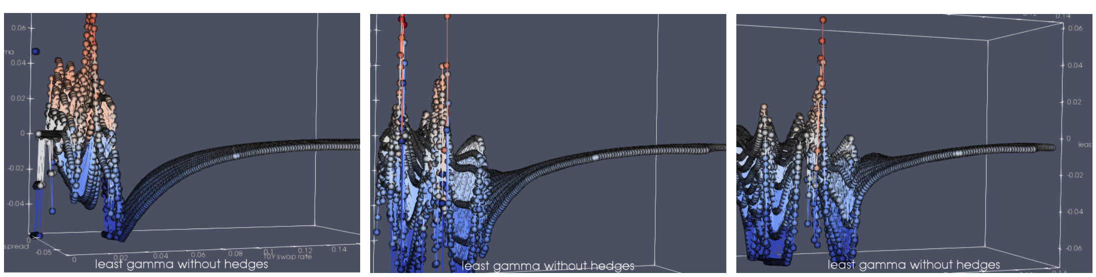

Figure 8: Least crossgamma eigenvalue, color-coded as blue if negative, red if positive.

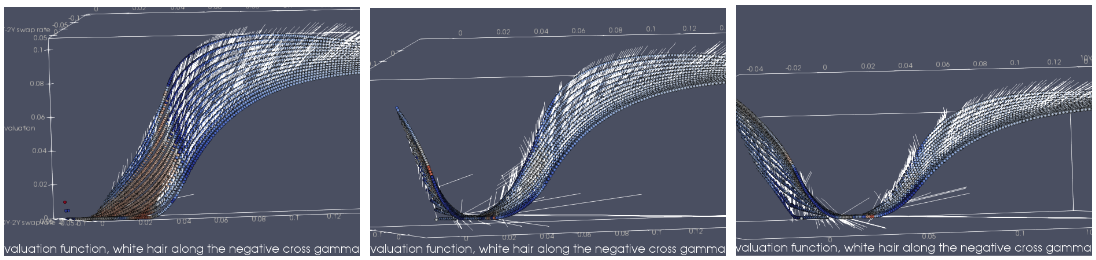

Figure 9: Valuations of the challenger. The white hair point in the direction of the negative crossgamma eigenvector.

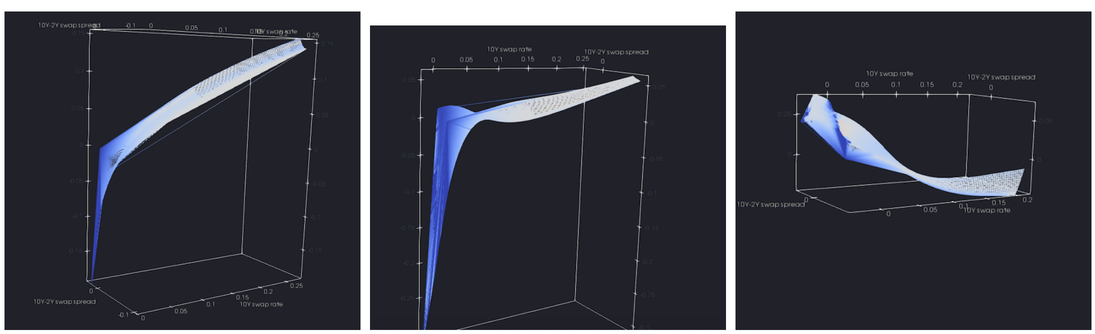

Figure 10: Exposure to the over-hedging by the HW model from the viewpoint of the SD model.

5 Conclusion

Pricing models are built with the objective of hedging and risk managing portfolios of financial derivatives. The risk in a derivative portfolio can be partitioned between market risk, which is focused on second and higher moments of the return distribution, and model risk, focused on the first moment. While delta/vega hedging strategies target the higher moments, typically overnight, model risk results in a bounded variation residual which manifests itself over longer time periods. There is a bias in the sign of the residual from among the sample of low-quality models which stay in use for an extended period of time. Darwinian selection favors models that generate systematic profit in the short and medium term compensated by possible large losses in the long term. These losses are undetectable by market risk models such as value-at-risk (VaR), expected shortfall and stressed VaR.

The state space analysis on which this article is based lends itself to visualization by means of intuitive 3d animations.

A Darwinian Model Risk in Equations

In this mathematical appendix, we illustrate the Darwinian model risk principles of Section 1 on the case of coupon-bearing American options. In doing so, we will assume that a bank issues, at time 0, such an option, which however will be priced within a wrong model class. Furthermore, we also assume that the option will be hedged by the bank, using the same model, with a set of European options to which the wrong model is continuously recalibrated. For notational simplicity, we use the risk-free asset growing at the OIS rate as a numéraire. Assuming Itô processes (without jumps) everywhere for simplicity, we denote by

- X_t, the true market risk factor;
- ρ_t, the parameters of the wrong model continuously recalibrated to the market prices of the European hedging options;
- $\Pi_t = \Pi(t, X_t, \rho_t)$, the price process of the coupon-bearing American option to be hedged;
- $P_t = P(t, X_t, \rho_t)$, the market price process (in vector form) of the European hedging options;
- ζ (of the same dimension as P_t and assumed predictable and nonnegative componentwise), the hedging ratios in the European options, computed in the wrong, continuously recalibrated model;
- Γ and C, the cumulative coupon streams of the American and of the European options.

As long as the American option is not exercised, we have between the option coupon dates, by the Itô formula,

$$
\begin{aligned}
d\Pi_t = d\Pi(t, X_t, \rho_t) = \partial_x \Pi_t dX_t + \big(\partial_t \Pi_t dt + \frac{1}{2}\partial^2_{x^2}\Pi_t d\langle X\rangle_t\big) \\
+ \partial_\rho \Pi_t d\rho_t + \frac{1}{2}\partial^2_{\rho^2}\Pi_t d\langle \rho\rangle_t + \frac{1}{2}\partial^2_{x,\rho}\Pi_t d\langle \rho, X\rangle_t,
\end{aligned}
\tag{4}
$$

complemented by a no-jump condition for $(\Pi + \Gamma)$ at coupon dates. Because of recalibration, $(\Pi + \Gamma)$ may have a nonzero drift coefficient α.

The first Darwinian adverse selection principle for a wrong model to survive is alpha *leakage* in the sense of $\alpha \leq 0$, a highly implicit condition in view of (4), which can only be attained by "natural selection between wrong models".

Analogous considerations apply to P (componentwise, if P a vector), with coupons C, with related risk-neutral martingale $(P + C)$ (assuming no arbitrage on the market of the hedging assets). The all-inclusive wealth process of the bank accounting for the American option and its European option hedges is given by $V_0 = 0$ and, for $t > 0$,

$$
dV_t = d\Pi_t + d\Gamma_t - \zeta_t\big(dP_t + dC_t\big), \tag{5}
$$

whence a drift coefficient $\alpha \leq 0$ of V.

The second Darwinian principle states that the hedging ratios ζ computed in the calibrated but wrong model are too high with respect to the reality of the market risk face by the bank, so that it effectively puts the bank in a net convexity seller position. In particular, because the option hedged is American, using a wrong model class for the bank implies that the option should be held longer in the model than in reality, which pushes ζ up. In the short to middle term, the bank harvests profit but, in the long run, it suffers from the alpha leakage of its position, which accumulates systematically, and from the realization of its net (hidden) negative gamma risk.

In case of a liquidity premium (positive at time 0) on the hedging assets so that $(P + C)$ deviates from martingality, the value process V of the strategy benefits of increased returns on the short to medium period. However, typically a liquidity premium is associated with a market-wide imbalance between offer and demand, which typically signals the occurrence of a macro-economic stress event. In this situation, if the portfolio must be unwound under stressed conditions, one can expect nonlinear market inefficiencies to arise such as gamma traps, wider transaction costs and amplified losses to a largely unpredictable extent (cf. Section 1).

In this paper we consider the long-run model risk of a bank that is effectively selling volatility to the market. In Karoui, Jeanblanc-picqué, and Shreve (1998), the emphasis is how a hedge fund can guard against model risk by using a conservative volatility hedging model. Their analysis is based on the consideration of the terms in parentheses in (4). Assuming in their case delta hedging and a given (and fixed) local volatility function $\rho_t = \rho$ as wrong model parameter, the hedging error of the fund (akin to the wealth process of the bank) reduces in their setup to this term. Using the pricing equation for Π in their setup, they conclude that convex payoffs can be super-hedged by a hedge fund using an overestimated volatility in their model.

The above-mentioned paper can be considered as an ancestor of an important active branch of model risk related mathematical finance literature in the direction of robust pricing and hedging, with a more recent ramification in the direction of martingale optimal transport: see e.g. Henry-Labordere (2017) for references. While robust finance provides a systematic exploration of all the mathematically possible scenarios (possibly subject to no-arbitrage or calibration constraints), it hardly embeds a notion of realism of the ensuing extreme scenarios, or permits to introduce a (Bayesian) notion of subjective views.

References

Albanese, C. (2019). https://www.youtube.com/watch?v=CVVbcEfaEB.

Albanese, C., S. Crépey, and S. Iabichino (2020). Reverse stress testing. Working paper.

Albanese, C. and M. Trovato (2008). A Stochastic Monetary Policy Interest Rate Model. *Computational Methods in Financial Engineering*, 343–392.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–654.

Carr, P. (1998). Randomization and the American put. *The Review of Financial Studies 11*(3), 597–626.

Carr, P. and D. Madan (1998). Towards a theory of volatility trading. In *Volatility: new estimation techniques for pricing derivatives*, pp. 417–427. Risk Books.

de Finetti, B. (2017). *Theory of Probability: A critical introductory treatment*. Wiley.

Devasabai, K. (2017). Model risk managers grapple with interconnectedness. *risk.net*.

Devasabai, K. (2019). Remembering the range accrual bloodbath. *risk.net*.

Doob, J. (1953). *Stochastic Processes*. Wiley.

Dupire, B. (1993). Model art. *Risk Magazine 6*, 118–124.

Dupire, B. (2019). The perils of parameterization. Slides of the Quant Summit USA Risk.net conference, New York, July 20, 2019.

Feibel, B. (2003). *Investment Performance Measurement*. Wiley.

Girsanov, I. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and its Applications 5*, 285–301.

Hagan, P. (2002). Turning good prices into great prices. *Wilmott Magazine 2*, 56–59.

Hagan, P., D. Kumar, A. Lesniewski, and D. Woodward (2002). Managing Smile Risk. *Wilmott Magazine September*, 84–108.

Henry-Labordere, P. (2017). *Model-free hedging: A martingale optimal transport viewpoint*. Chapman and Hall/CRC.

Hull, J. and A. White (1990). Pricing interest rate derivative securities. *Review of Financial Studies 4*.

Karoui, N. E., M. Jeanblanc-picqué, and S. E. Shreve (1998). Robustness of the Black and Scholes formula. *Mathematical Finance*, 93–126.

Li, D. X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income 9*.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science. The RAND Corporation 4*, 141–183.

Meyer, P. (1962). A decomposition theorem for supermartingales. *Illinois Journal of Mathematics 6*, 193–205.

Protter, P. (2004). *Stochastic Integration and Differential Equations* (2nd ed.). Springer.