Long Dated Derivatives

Claudio Albanese

⋆ ⋆ ⋆

Presented at Warwick University, July 11th, 2008

Why this is a Great Time to be a Math Finance PhD Student

- The credit crisis showed that standard models give rise to highly correlated model risk. There's a felt need for new modeling technology.

- Massively parallel multi-core architectures are a major technology shock that is bound to have repercussions on model building.

- Innovations in Math Finance can lead to innovations in the base sciences, such as probability, numerical analysis, economics.

The Courant Condition

$$\boxed{\Delta t \leq \frac{1}{2} \max \mu(x_j) \Delta x_j}$$

This is the stability condition for explicit Euler discretization schemes. In the context of parabolic PDEs of use in financial modeling, the Courant condition typically implies that the time step should not be longer than just a few hours for explicit methods to be applicable.

The Twin Curses of Numerical Analysis

- The curse of the Courant condition
- The curse of dimensionality (path dependent options and baskets)
- Noisy Sensitivities

It turns out that these three difficulties are all intertwined with each other.

By resolving the curse of the Courant condition at the engineering level (as opposed to avoiding it) one can make substantial progress toward resolving the other two.

Computer Technology and the Courant Condition

Traditionally, fast memory has been a scarce resource. This technical limitation motivated the development of

- sparse matrix linear algebra methods
- weakly (i.e. marginally) stable methods which allow for a time step longer than dictated by the Courant condition
- chipsets optimized to double precision arithmetics (about 10 times slower than chips optimized to single precision with the same number of transistors)

Computer Technology and the Courant Condition

Nowadays the technology environment has changed.

- There is enough system memory (most real-life applications I implement require not more than 1GB)
- Very efficient implementations of full matrix-matrix multiplication algorithms, i.e. `dgemm/sgemm`, are available
- Massively parallel multi-core architectures with fast on-chip interconnects (like GPUs, FPGAs, CELL BE) are available, implement `dgemm/sgemm` and are optimized to single precision.

In the new technology environment, there is a strong motivation to devise strongly stable algorithms that respect the Courant condition as opposed to marginally stable ones that avoid it.

Explicit Methods can be fast (in addition to being accurate)!
Sample benchmarks for a semiparametric interest rate model, 672 dimensional lattice, 3 hours time step, maturities from short to 30-into-20 years.

Task	GPU-O	GPU-D	Host-O	Speedup
Initialization	4.79	5.16	3.77	0.79
Calibration to term structure (392 GF)	2.42	2.88	61.88	25.57
585 Bermuda swaptions (580 GF)	3.61	3.97	85.30	23.62
30240 Bermuda swaptions (830 GF)	5.14	5.65	138.68	26.98
12 callable CMS spread range accruals snowballs (1620 GF)	10.02	11.04	134.59	13.43

GPU: Tesla D870, **Host:** 4-core Xeon.

GPU-O: using the GPU `sgemm` with host side optimized code.

GPU-D: using the GPU `sgemm` with host side debug code.

Host-O: using the host `dgemm` only.

Explicit methods can be fast (in addition to being accurate)!

Sample benchmarks for a CDO model, 125 reference names, 600 dimensional lattice, 1 hours time step, maturities up to 10 years.

Task	CPU Time	Memory requirement
Preprocessing	137.57 sec	308 MB
Single Name Calibration	2.69 sec	77 MB
CDO Tranche Pricing	8.63 sec	181 MB

Execution times on a single processor Xeon machine, 2 GHz, with a nVidia Tesla GPU coprocessor.

The Key Concept is Smoothing

The Courant condition ensures a degree of smoothness that is lost in unconditionally stable methods with longer time steps.

If one respects the Courant condition, then probability kernels are smooth, meaning smoothness is numerically observed even if initial (or final) conditions are delta functions or first or second derivatives of delta functions.

Operator Methods for Direct Kernel Manipulation

If the Courant condition is respected, then one can evaluate efficiently probability kernels and the resulting kernels are smooth.

A number of new numerical methods becomes available if one can manipulate kernels directly. These methods help one tackle high dimensional situations in new ways.

Constructive Probability

Numerical Analysis is an eminently constructive field: algorithms need to be communicated to computing machines in the form of software codes. To work at the intersection between Numerical Analysis and Probability Theory one needs to simplify the mathematical and logical framework.

No Axiom of Countable Additivity

No Axiom of Choice

No Principle of the Excluded Middle

Different meaning of existence qualifiers $\exists$: One can claim that an object exists only if one can construct it explicitly

Constructive Probability

Probability is reduced to Numerical Linear Algebra and leverages on multi-core technology as opposed to leveraging on special function valuation

New mathematical landscapes are revealed

Focus on smoothness and kernel manipulations

One learns how to avoid the curse of dimensionality and Montecarlo simulations

Plan for the rest of the talk

✓ Introduction

♢ **A Toy Model**

□ Smoothness and Convergence Estimates

□ Semi-parametric Modeling

□ Path Dependencies (Abelian, non-Abelian, moment methods)

□ Dynamic Conditioning

□ Constructive Probability Theory

□ Dynamic Conditioning Model for CDOs

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

Plan for the rest of the talk

✓ Introduction

✓ A Toy Model

♢ **Smoothness and Convergence Estimates**

□ Semi-parametric Modeling

□ Path Dependencies (Abelian, non-Abelian, moment methods)

□ Dynamic Conditioning

□ Constructive Probability Theory

□ Dynamic Conditioning Model for CDOs

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

The fundamental problem of fundamental solutions

Semi-parametric models are viable from the engineering viewpoint at the condition that one is able to evaluate fundamental solutions (i.e. probability kernels) with a low noise-to-signal ratio. Low noise is needed not only with respect to the uniform norm for the kernel, but also with respect to its derivatives with respect to the arguments and the internal parameter sensitivities.

Analytic solutions are useful precisely because they allow one to express the kernel in terms of special functions. The purpose is to find numerical methods that yield the kernel to similar or higher accuracy with a performance independent of the process specification.

Pseudo-spectrum

spectral methods are not sufficient because of pseudo-spectrum pathologies. Pseudo-spectrum emerges with particular virulence when the drift is stochastic (as for instance one would have if one wanted to adequately model the interest rate process in the example above).

Smoothness and the Courant Condition

Even assuming infinite precision arithmetics, I also realized that discretization schemes are effective only if the time step is either zero or satisfies the Courant condition.

Unconditionally stable methods are only weakly stable and they don't converge well in the uniform graph norm with respect to the Markov generator. Instead, it is essential for operator methods to have convergence in the uniform graph norm. This can only be ensured by respecting the Courant condition.

Linear Fast Exponetiation

I find that linear fast exponentiation allows one to satisfy the Courant condition and still be the basis of efficient numerical methods.

Fast exponentiation based on more complex Pade approximants are not as efficient empirically and can be proven not to converge in graph norm in some particular cases.

Finite versus double precision arithmetics

Empirically one observes that linear fast exponentiation performs very well even in single precision. The ideal platform to carry out such calculations is given by the emerging GPU hardware (nVidia CUDA, IBM CELL BE, etc.)

Conversely, weakly stable methods require double precision arithmetics. They also require less memory, reflecting the memory-precision trade-off that historically took place.

Convergence Estimates

One can prove convergence estimates in the graph norm for diffusions even in case coefficients are rough (i.e. Hölder continuous), thus reobtaining and refining results proven nonconstructively by Nash, Farbes, Strook and Varadhan in the 50s and 60s. Such estimates extend to join distributions between diffusions and Abelian processes including stochastic integrals.

Full versus sparse matrix numerical linear algebra

I mostly use BLAS level-3 methods based on `sgemm` and `dgemm`, the matrix multiplication routines, with a preference for single precision.

I never take advantage of sparsity patterns, use LU factorizations or SVD decompositions.

Probability theory

Sigma algebras were invented and became a popular topic of study because of the desire of describing processes with continuous state spaces (as opposed to discrete Markov chains). Measure theory methods are unfortunately not constructive.

Semigroup methods for Galerkin schemes are useful to establish the convergence of discretization schemes but not powerful enough to give direct control on kernel convergence in the uniform graph norm and path-dependent processes. My methods instead make use of renormalization group transformations.

Kernel convergence estimates can be considered as the first steps leading to a fully constructive theory of stochatic processes based on Markov chain approximations.

Summary

✓ Introduction

✓ A Toy Model

✓ Smoothness and Convergence Estimates

♢ **Semi-parametric Modeling**

□ Path Dependencies (Abelian, non-Abelian, Moment Methods)

□ Dynamic Conditioning

□ Constructive Probability Theory

□ Dynamic Conditioning Model for CDOs

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

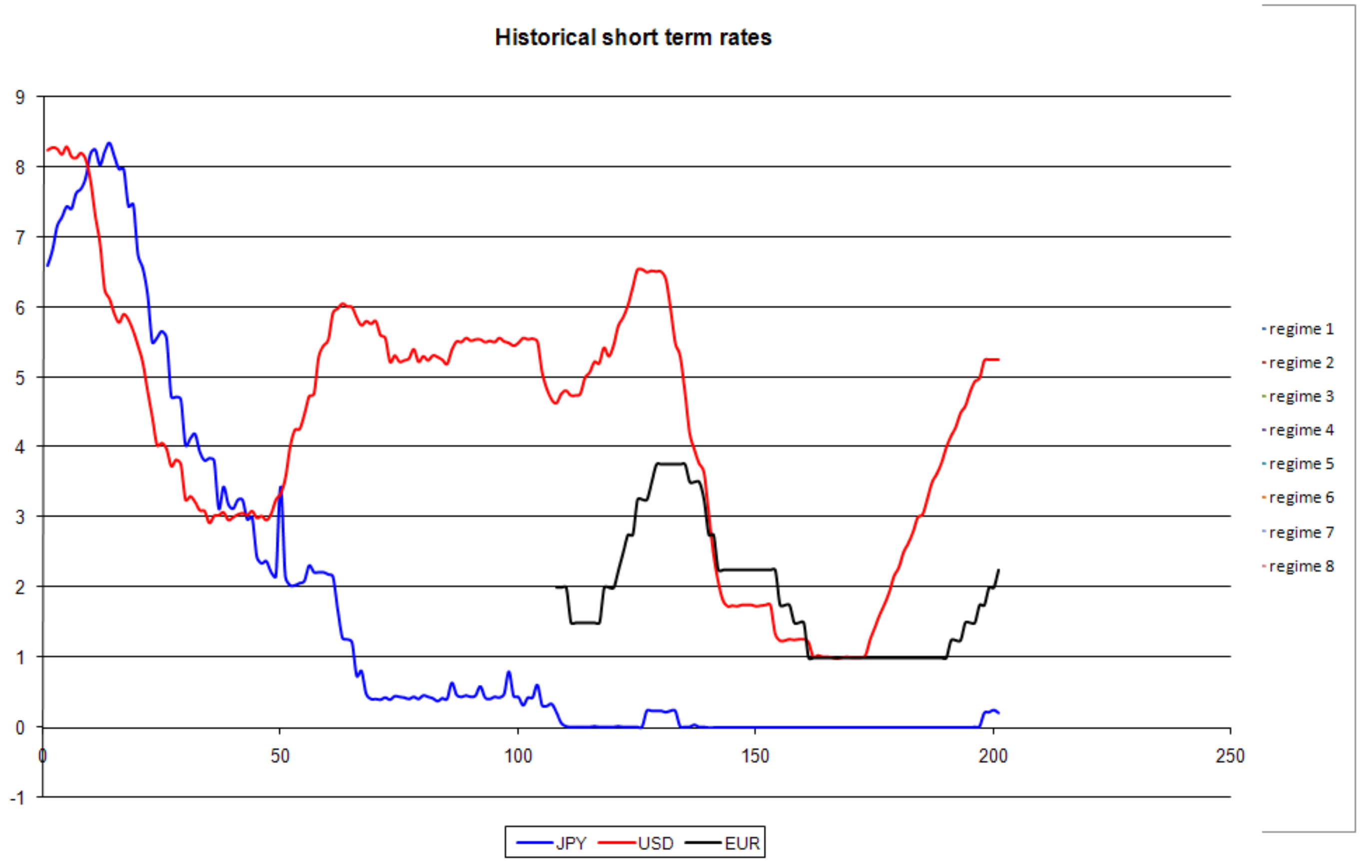
Semi-parametric models are the natural model class

If the main computational engine is given by matrix-matrix multiplcation routines as opposed to being given by special function evaluation and random number generation, the natural model class to consider is given by semi-parametric or even non-parametric models.

The reason again is smoothness, the low signal-to-noise level for sensitivities, which allows one to calibrate.

Why are semi-parametric models interesting?

The process for short term rates does not resemble one solvable in closed form.



Semi-parametric models for derivatives

Long dated derivatives and baskets are complex and have substantial model risk. Different models that calibrate equally well give possibly very different answers for prices and hedge ratios.

The better models are the ones which, in addition to calibrating, capture econometric evidence and embed economic views.

Engineering Challenges

From the mathematical and engineering viewpoint, the challenge is to devise models that are as flexible as possible but which still can be calibrated.

Montecarlo methods are very flexible from the model specification viewpoint but their applicability is limited by the necessity of calibrating. Calibration requires the availability of low-noise methods for calibration targets, i.e. they ultimately need to be based on

- closed form solutions (possibly with Fourier integrals)
- asymptotic expansions
- lattice methods.

Analytically solvable models

The majority of derivative models currently used is tractable in analytically closed form.

Several models are used. It appears that these models do not fit the standard Lie classification scheme. In the years 2000-2003 I developed a more general classification scheme (with A. Kuznetsov) that encompasses all known solvable models for which the probability kernel is known in closed form. I also developed a second classification scheme (with S. Lawi) for models for which the generating function of the joint distribution between a diffusion process and a stochastic integral can be expressed in closed form.

In this research, I used spectral methods and the theory of special functions.

Semi-parametric models are more general than the ones solvable in closed form and this allows one to achieve a higher degree of agreement with econometric evidence. My research was driven by the following real-world applications:

- A stochastic monetary model for interest rate derivatives
- A structural model for CDOs.
- A stochastic skew foreign exchange model.
- A 3-factor model for long dated equity derivatives and equity baskets.
- A model for electric power derivatives.

Summary

✓ Introduction

✓ A Toy Model

✓ Semi-parametric Modeling

✓ Smoothness and Convergence Estimates

$\diamondsuit$ **Path Dependencies (Abelian, non-Abelian, Moment Methods)**

□ Constructive Probability Theory

□ Dynamic Conditioning

□ Dynamic Conditioning Model for CDOs

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

Abelian versus non-Abelian processes

The notion of Abelian process emerges naturally when one learns how to associate an operator algebra to a stochastic process adapted to the filtration of a driving one. If the operator algebra is commutative the process is Abelian.

Stochastic integrals are Abelian processes as well as the sup of a process. The structured leg of a clicket, a TARN or a range accrual are Abelian processes.

The payoff of snowballs, soft call convertibles and flexicaps are example of non-Abelian processes.

The Cameron-Martin-Feynman-Girsanov-Ito theorem in one formula

Consider the diffusion process of Markovian

$$\mathcal{L} = \mu(x,t)\nabla + \frac{\sigma(x,t)^2}{2}\Delta. \tag{1}$$

Consider also the process given by the integral

$$I_t = \int_0^t a(x_s,s)dx_s + b(x_s,s)ds \tag{2}$$

where $a(x,t)$ and $b(x,t)$ are smooth functions in both arguments.

The Cameron-Martin-Feynman-Girsanov-Ito theorem in one formula

We have that

$$I_t = \phi(x_t, t) - \phi(x_0, 0) + J_t. \tag{3}$$

where $\phi(x,t) = \int_0^x a(y,t)dy$.

$$J_t = \int_0^t \left(b(x_s, s) - \frac{1}{2}\sigma(x_s, s)^2 a'(x_s, s) - \dot{\phi}(x_s, s) \right) ds. \tag{4}$$

and

$$\begin{aligned} &E_0\Big[e^{ipJ_t}\delta(x_t = y)|x_0 = x\Big] \\ &\quad = P \exp\left(\int_0^t \left(\mathcal{L}(s) + ipb(s) - \frac{ip}{2}\sigma^2(s)a'(s) - ip\dot{\phi}(s) \right) \right)(x,y). \end{aligned} \tag{5}$$

Theory of Abelian Processes

The fundamental theorems of stochastic calculus can be seen as an application of the theory of Abelian processes to the particular case of diffusions processes and stochastic integrals.

The theory of Abelian processes however

□ applies to more general Markov processes (with jumps, regime switching, etc.)

□ covers more general path-dependencies (like integrals, the sup, discrete sums)

□ supports more general harmonic analysis (replacing the Fourier characteristic functions with transforms adapted to the problem).

Block-diagonalization methods

Block-diagonalization methods can be used for Abelian processes, both numerically and theoretically.

A theory of stochastic integrals can be built on this basis, thus handling the hypo-elliptic nature of the joint evolution operator.

Numerically, block-diagonalization can be achieved either by means of Fourier transforms or by more elaborate but sometimes more efficient harmonic analysis constructs in different bases.

Other Useful Operator Methods for Path Dependencies

□ Moment methods (based on the Dyson-Kac formula)

□ Block-factorization methods for non-Abelian payoffs (useful when block-diagonalizations are not available)

Summary

✓ Introduction

✓ A Toy Model

✓ Semi-parametric Modeling

✓ Smoothness and Convergence Estimates

✓ Path Dependencies (Abelian, non-Abelian, Moment Methods)

$\diamondsuit$ **Dynamic Conditioning**

☐ Constructive Probability Theory

☐ Dynamic Conditioning Model for CDOs

☐ Stochastic Monetary Policy Model for Interest Rate Derivatives

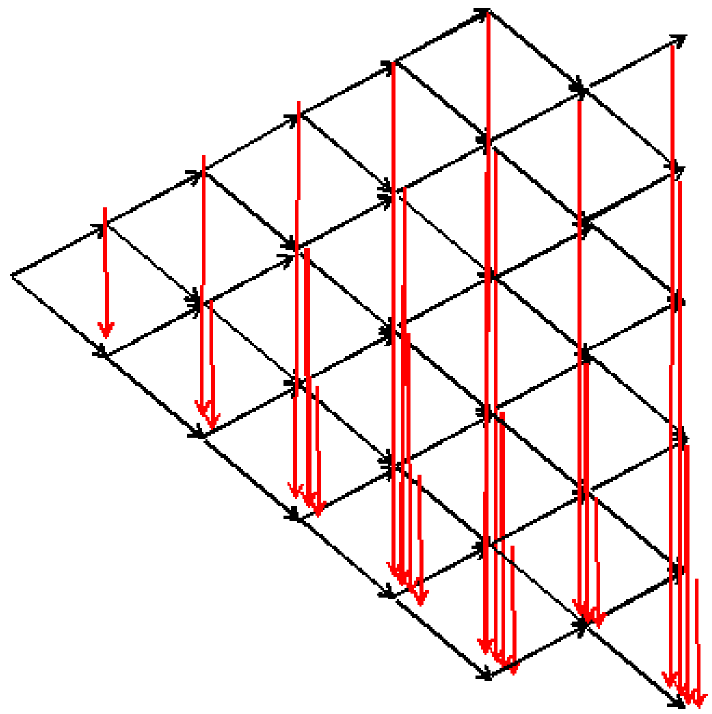
Dynamic conditioning

I have been working for over 4 years now at refining a method of dynamic conditioning. The method is intended to beat the so called "curse-of-dimensionality" and enables one to build lattice models in situations where one has many risk factors (such as for instance for CDOs).

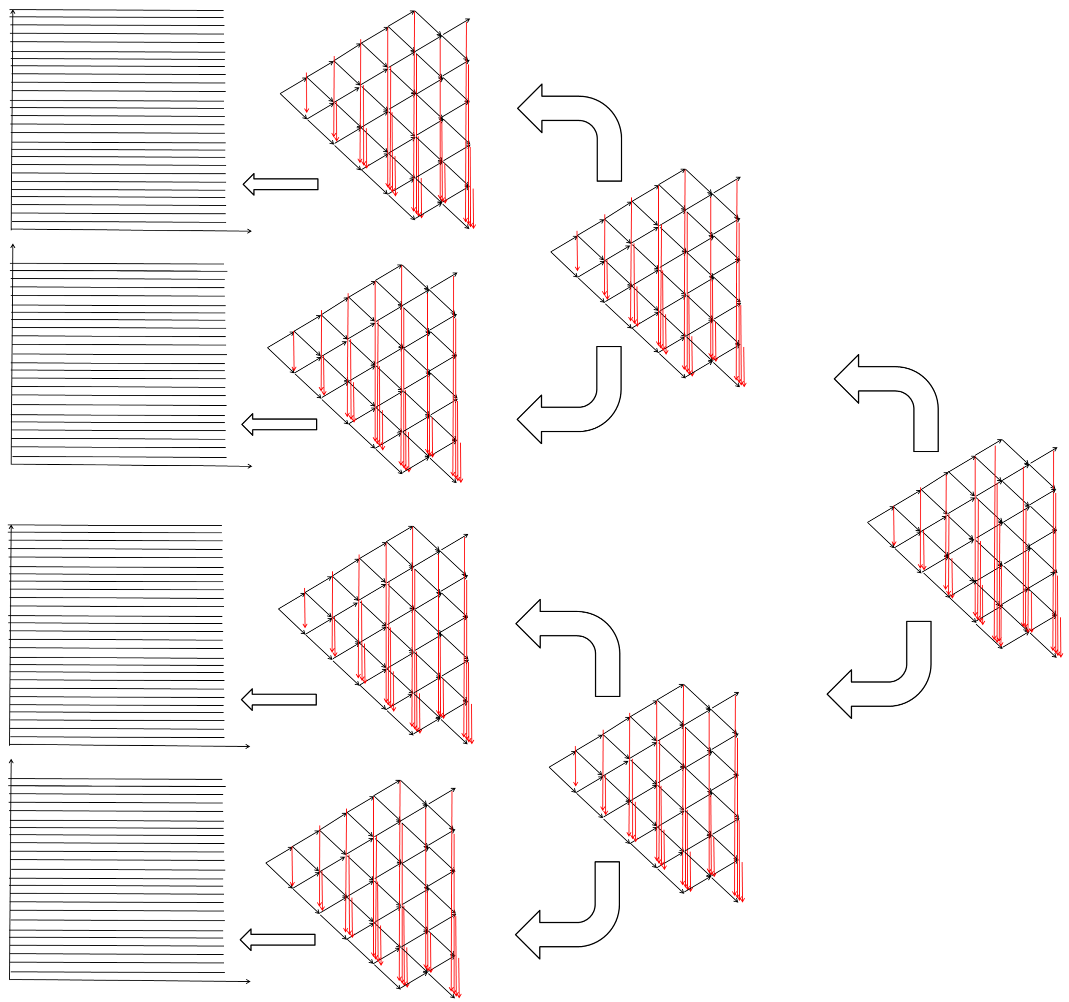
I just completed coding a third version of dynamic conditioning method applied to CDOs.

The key observation is that by using direct kernel manipulations one can correlate by conditioning.

Conditioning tree expressing a binomial process with jumps to the bottom vertex.



Scheme for multifactor conditioning involving industry sector factors and a global economics factor.



Summary

✓ Introduction

✓ A Toy Model

✓ Semi-parametric Modeling

✓ Smoothness and Convergence Estimates

✓ Path Dependencies (Abelian, non-Abelian, Moment Methods)

✓ Dynamic Conditioning

♢ Constructive Probability Theory

□ **Dynamic Conditioning Model for CDOs**

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

Mathematics, Numbers and Logic

Mathematics is about numbers and logic. In this paper we follow Brower's Intuitionistic Logic, a subset of classical Mathematical Logic according to which before considering any object we need to explain how to construct it. Integers up to ten can be counted on the fingers of our hands and many hands represent large integers. Fractions and rational numbers can also be conceptualized. Real numbers are a bit more of an awkward concept as a real number is described by a Cauchy sequence of rational numbers (f_n) such that for all $\epsilon > 0$ there is a $N > 0$ for which $|f_n - f_m| < \epsilon$ for all $n, m > N$.

Probability Theory and Colored Marbles

Probability theory takes its premises from the concept of random draws of equiprobable integers out of a finite set. Rational probabilities can be defined by taking a basket of equiprobable marbles, coloring them in various ways and associating the result of a draw to a color as opposed to an individual marble. So far, probability and combinatorics have the same conceptual root. The two fields depart when the probabilist wants to consider draws with a general positive real valued probability. An operative definition of such draws involves considering not a single basket of marbles, but a sequence of baskets of increasing size. It also involves computing an asymptotic.

Classical Probability Theory

The classic works on Probability Theory by Bernoulli, Laplace, de Moivre and Poisson are all based on computing limits and asymptotics. These derivations are understandable constructively within the confines of Intuitionistic Logic. A general description of a classic probabilistic result involves considering not one but a sequence of finite sets describing possible events, such as finite sequences of coin tosses. Next, one looks at an observable such as the mean, the variance or the distribution law for the number of outcomes of a given type. Finally, one takes a limit as the number of coin tosses goes to infinity and establishes results such as the law of large numbers, the central limit theorem or arrives at the Poisson law. Although continuum distributions as a rule emerge in the limit, these results are naturally formulated and established on finite event spaces as convergence estimates in the limit in which the number of possible events diverges to infinite.

Continuous Functions and Riemann Integrals

Continuous functions can be understood constructively because all one needs is to define them on the set of rational numbers. Integrals over continuous functions can be defined in the Riemann sense. Limits of continuous functions can also be conceptualized constructively similarly to what one does when obtaining real numbers out of rational numbers, by introducing the notion of Cauchy sequence with respect to a certain distance function. A natural distance function is the integral of the absolute value of the difference $||f_n - f_m||_1 = \int |f_n(x) - f_m(x)| dx$.

Lebesgue Integrals as a Convention on Limit Ordering

At the end of the nineteenth century, work on the theory of integration motivated efforts aimed at understanding the effect of exchanging the operation of limit over a sequence of functions with the operation of integration. Lebesgue proposed that limits would commute if one (i) agrees to regard all measurable functions as limits of continuous functions and (ii) whenever one encounters a double limit one of which is a space limit for a Riemann integral, one agrees to reshuffle the order of the limits by convention in such a way that the continuous space limit is carried out first. Regarded this way, Lebesgue theory is more a convention than an invention.

Borelians

Lebesgue's work can also be seen as an elaboration of the concept of countable additivity introduced by Borel. The family of Borelians and more general sigma-algebras of measurable sets give a geometric interpretation of sorts to this convention on the order of limit taking. The conceptual problem with this construction is that the family of Borelians can only be defined out of open and closed intervals of the real line by taking unions and intersections only if one accepts the concept of transfinite induction. This concept cannot be explained to a computer and exceeds the limits of Intuitionistic Logic.

Borel on Borelians

Being a real world science based upon Statistics and Physics, convention on limit ordering and non-constructive set-theoretical concepts seemed not justified as a matter of principle. Both Borel and Kolmogorov proposed two types of proofs for the Laws of Large Numbers, shorter ones based on measure theory and longer, fully constructive and more detailed ones with explicit limit taking. Borel considered constructive arguments as the most appropriate ones to Probability Theory and in line with the classic work of Bernoulli and others.

Kolmogorov

Kolmogorov's expository paper setting out an axiomatic framework for Probability Theory [**?**] represented a turning point. There, for the first time, the step was taken to introduce countable additivity as a *useful expedient* in Probability Theory on grounds that it allowed one to simplify derivations. Although not justifiable as a first principle according to Kolmogorov, countable additivity still seemed harmless and was thus introduce on a ”*why not*” basis. In the same years, Hilbert posted quite derogatory comments on Brower's Intuitionistic Logic, sealing the debate on constructive analysis until recent times.

Countable Additivity

The history of Probability followed a path all to common in human history. Traces of these early debates vanished from the Mathematical literature and was relegated to Philosophical circles. Measure theoretic Probability was funded upon the postulate of countable additivity so that limits and integrals could be freely interchanged. Once cast in stone in the textbook literature, countable additivity graduated from being a harmless expedient to becoming a defining property of the Theory of Probability. Being an eminently non constructive concept, countable additivity became an article of faith as most objects referenced cannot be constructed and communicated in full detail. Even worse, countable additivity was and still is often confused as a necessary parameter for mathematical rigor. What started as a convention and expedient became a human law and a limitation on mathematical speech, a meta-mathematical social convention applicable to academic human earthlings only and not communicable in full detail.

Constructive Analysis According to Brydges

Bridges work on constructive analysis was an important milestone. By adhering to Brower's Intuitionistic Logic and writing in the informal style of Classical Analysis, Brydges demonstrated that important chapters of Analysis could be understood constructively. Regarding the theory of Lebesgue integration, this work is instructive as it shows the merit of calling spade a spade and interpreting Lebesgue theory for what it truly is from a constructive standpoint, a convention on the order of limit taking.

The Courant Condition

In this paper, we propose to go a step further with respect to Brydges and keep an open mind about ordering of multiple limits. This is more than an academic exercise and is actually motivated by engineering considerations. Our main objective is to discuss the impact of the Courant condition when defining constructively stochastic processes. The Courant condition is an inequality implying stability of explicit differentiation schemes. In the case of stochastic process, one has to evaluate a space limit along with a continuous time limit. To respect the Courant condition, the two limits should be taken together following a diagonal direction, whereby the time step is proportional to the space step.

Smoothness

The result of the limit would not change if one were to take the space limit first and then the time limit as when applying the Trotter product formula. However, the rate of convergence is affected by the Courant condition, as this ensures convergence in the uniform graph norm, a norm involving not only the sup of the difference but also the sup of the difference of a combination of first and second space derivatives. Smoothness is of absolutely central practical importance for a multitude of numerical applications. By leveraging on smoothness, one can avoid and bypass the curse of dimensionality in many instances by means of direct kernel manipulations. Perhaps not surprisingly, we conclude that the benefit of avoiding the axiom of countable additivity and restricting to Intuitionistic Logic, deepens our understanding and reveals new and useful mathematical landscapes.

Stochastic calculus	**Operator methods**
Arbitrage free pricing	Arbitrage free pricing
Measure theory based on sigma algebras	Measure theory based on topological spaces and operator algebras
Dynamics described by SDEs	Dynamics described by Markov generators
Implicit differentiation schemes	Fast exponentiation
Double precision	Single precision
Diffusions and sparse matrices	Jump processes and full matrices
Analytic solvability	Reducibility to manipulations of matrices small enough to fit in memory
CPUs and CPU clusters	CPU/GPU pairs and GPU clusters
Market models	"Economic" models without drift restrictions
Measure changes	Operator manipulations possibly without probabilistic interpretation
Stochastic integrals Montecarlo methods	Abelian processes Dynamic conditioning

Summary

✓ Introduction

✓ A Toy Model

✓ Semi-parametric Modeling

✓ Smoothness and Convergence Estimates

✓ Path Dependencies (Abelian, non-Abelian, Moment Methods)

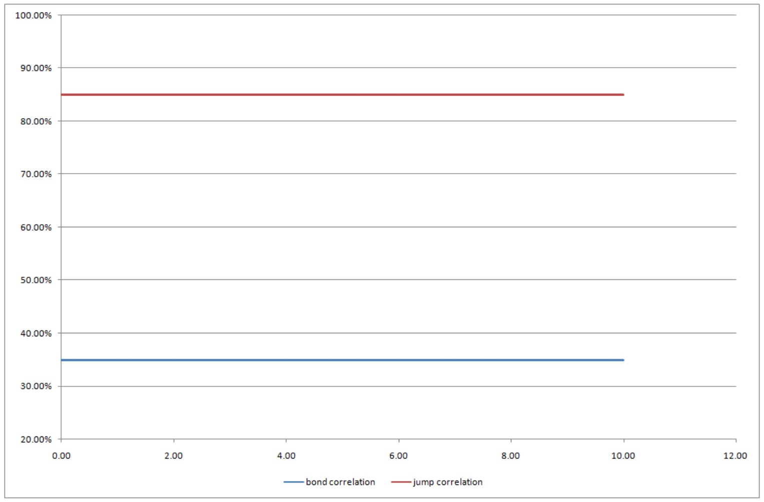
✓ Dynamic Conditioning

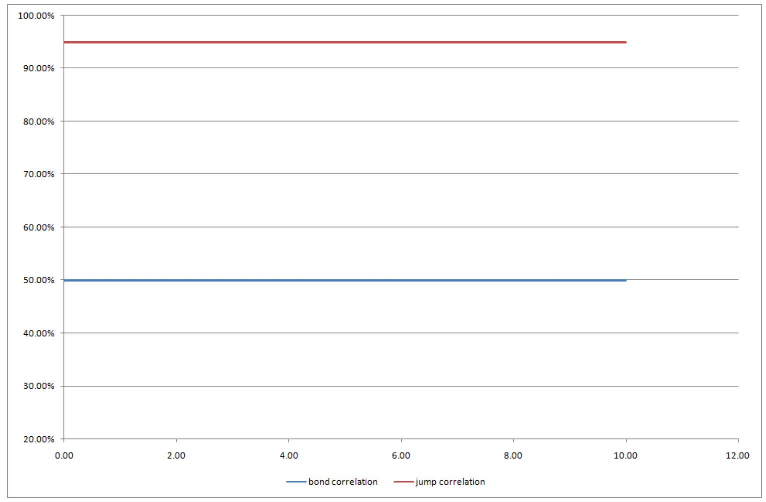
✓ Constructive Probability Theory

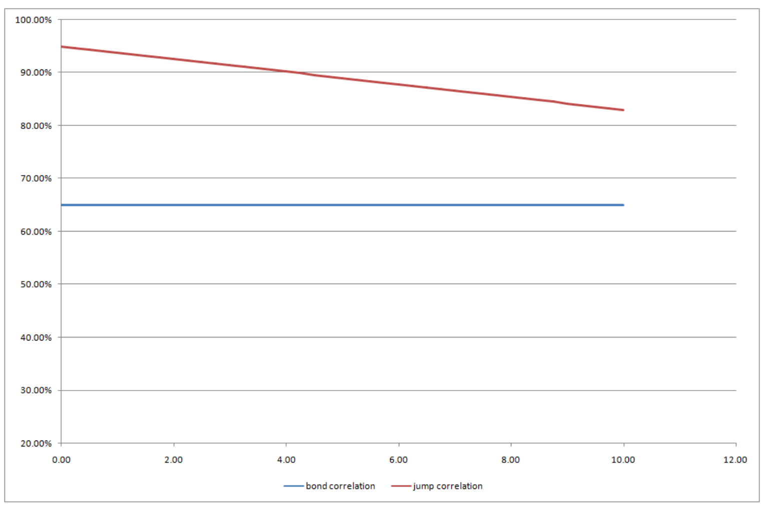
♢ **Dynamic Conditioning Model for CDOs**

□ Stochastic Monetary Policy Model for Interest Rate Derivatives

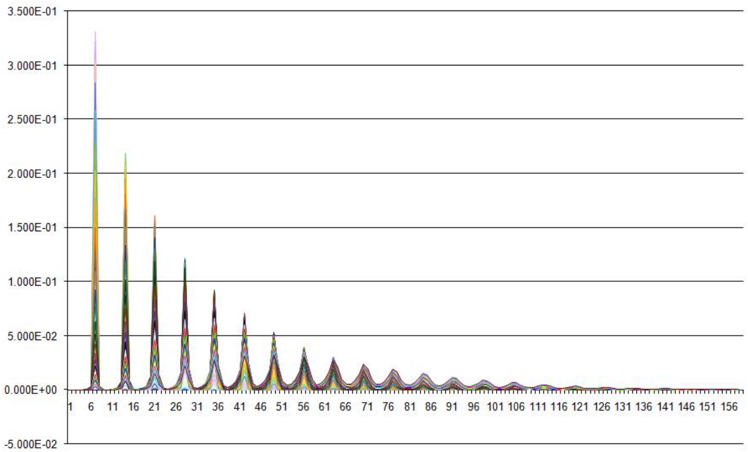
Term structure of bond and jump correlations, Apr-06, Mar-07, Oct-07

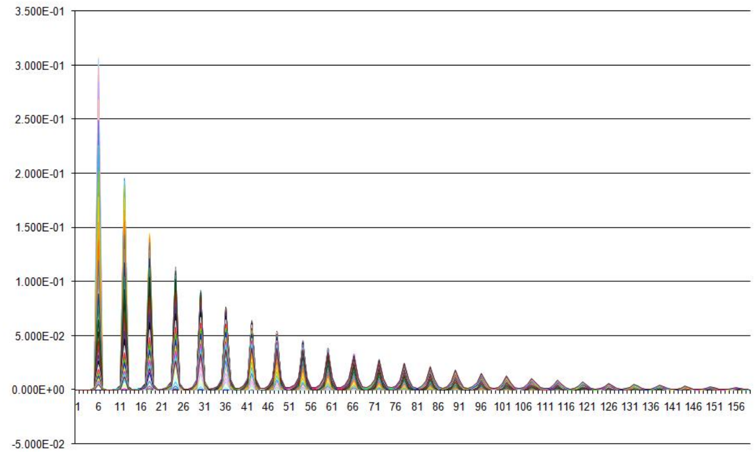


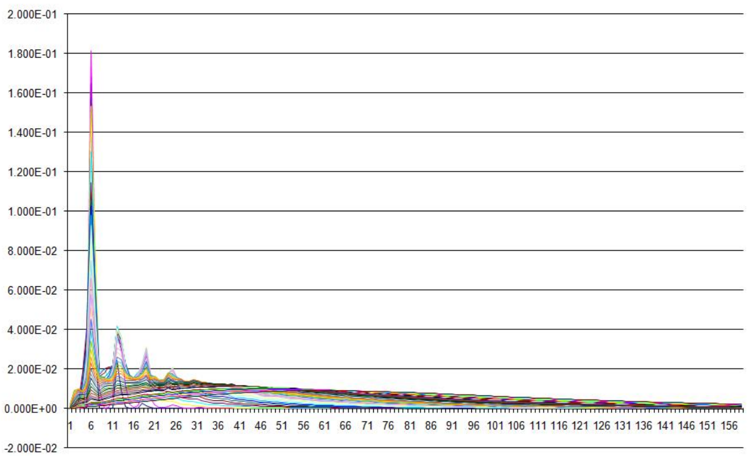




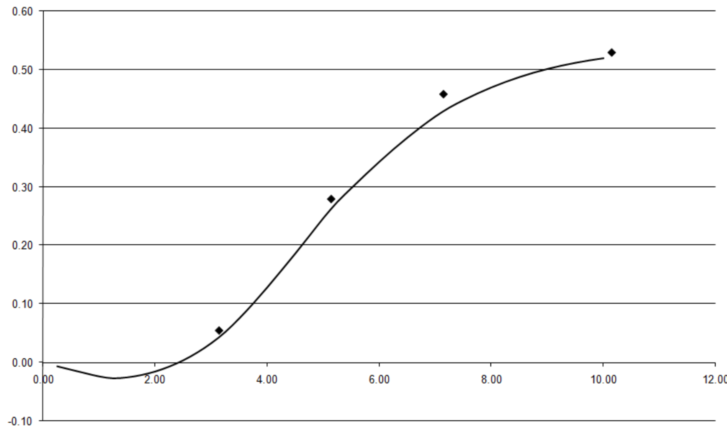
Term structure of loss distributions, Apr-06, Mar-07, Oct-07

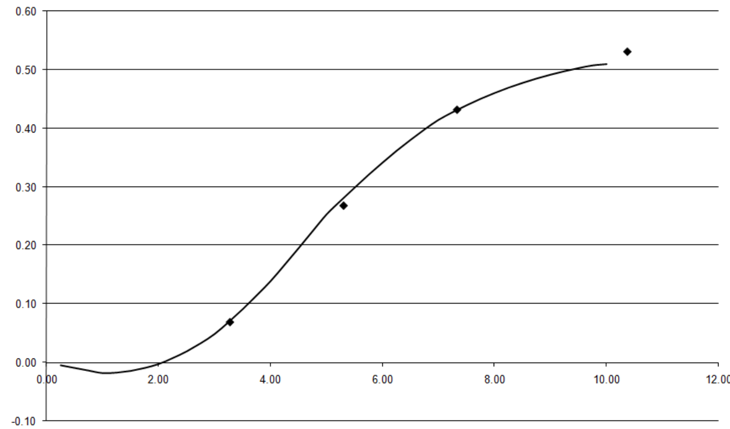


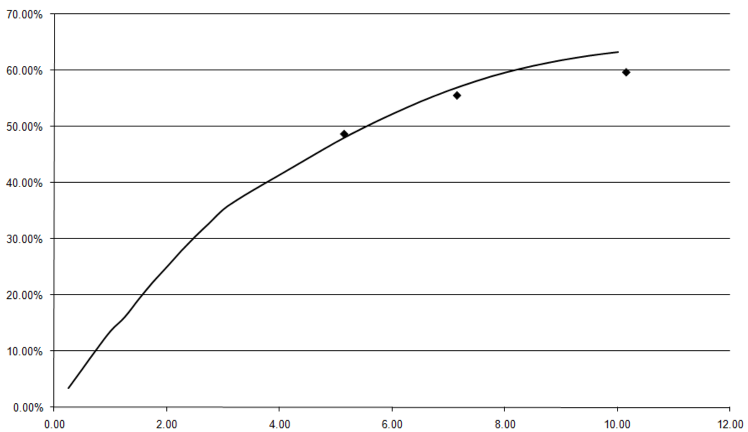




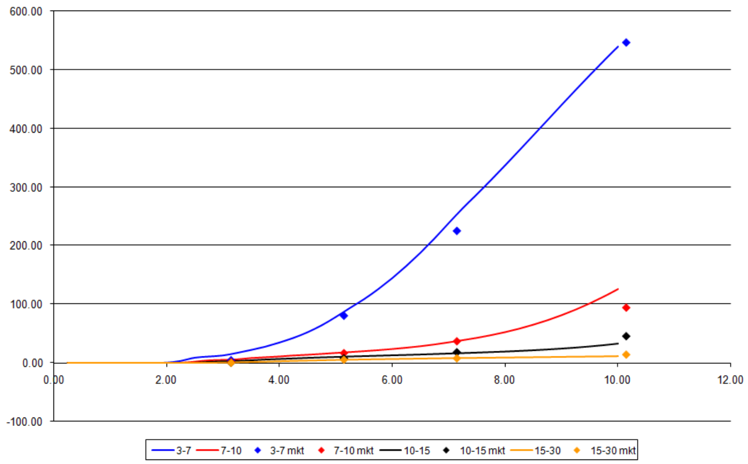
0-3 equity tranche upfront fee, Apr-06, Mar-07, Oct-07

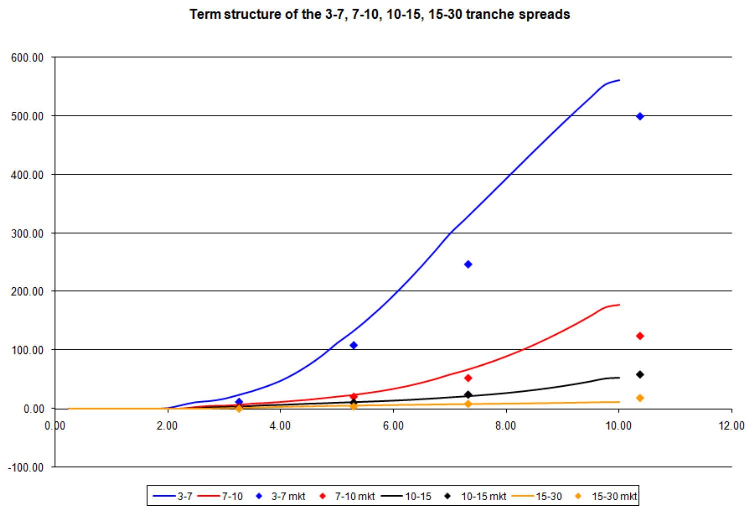


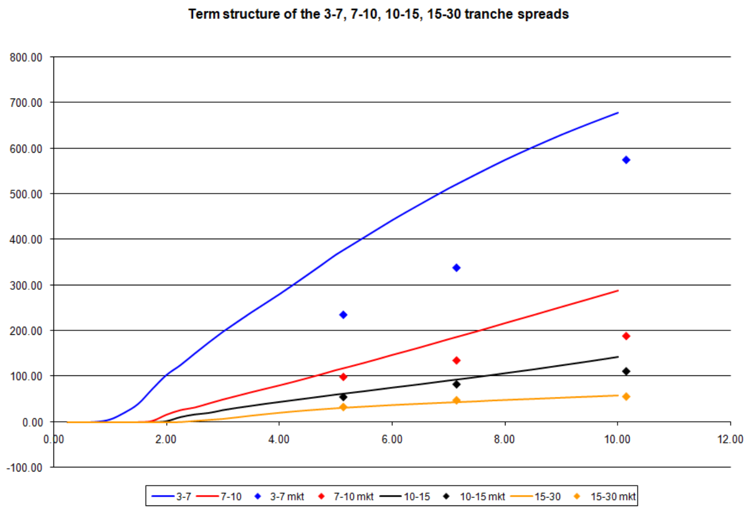




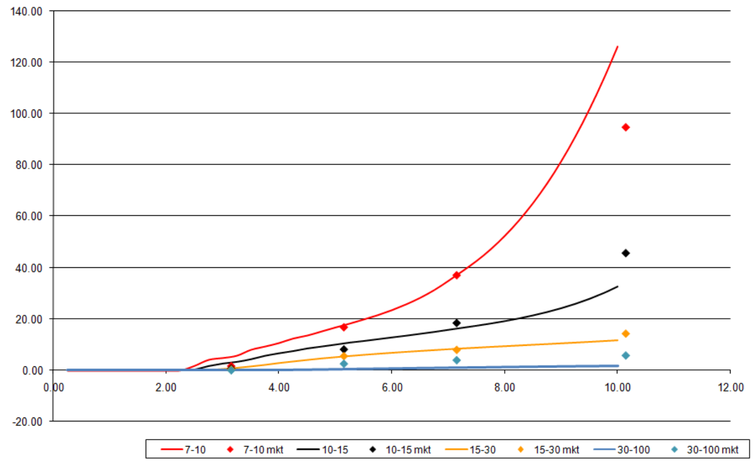
Mezzanine tranche spreads, Apr-06, Mar-07, Oct-07

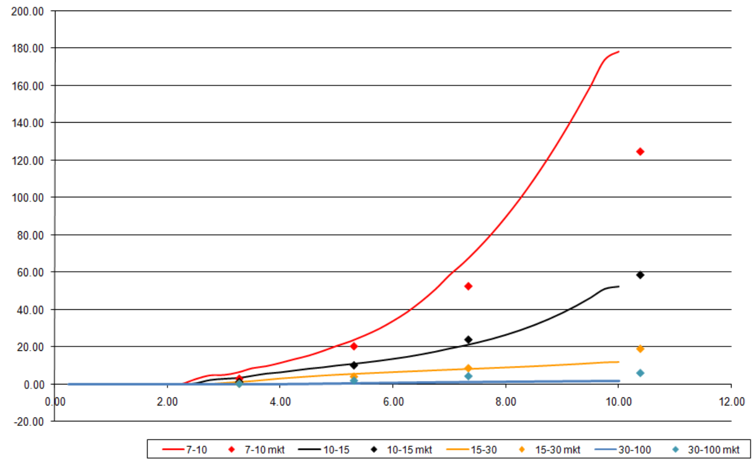


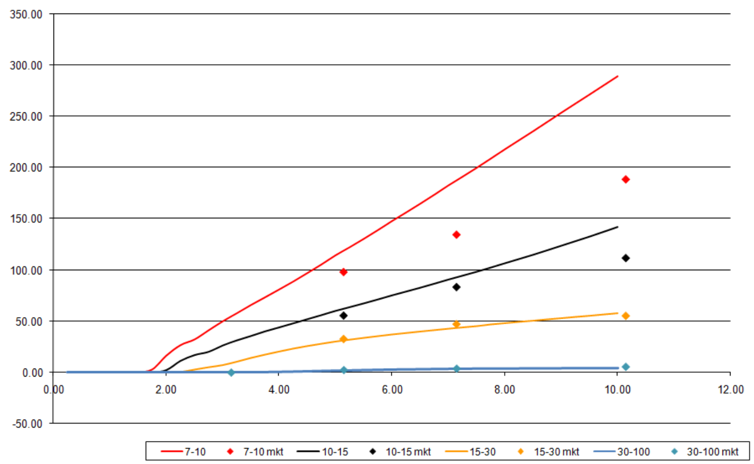




Senior tranche spreads, Apr-06, Mar-07, Oct-07







Summary

✓ Introduction

✓ A Toy Model

✓ Semi-parametric Modeling

✓ Smoothness and Convergence Estimates

✓ Path Dependencies (Abelian, non-Abelian, Moment Methods)

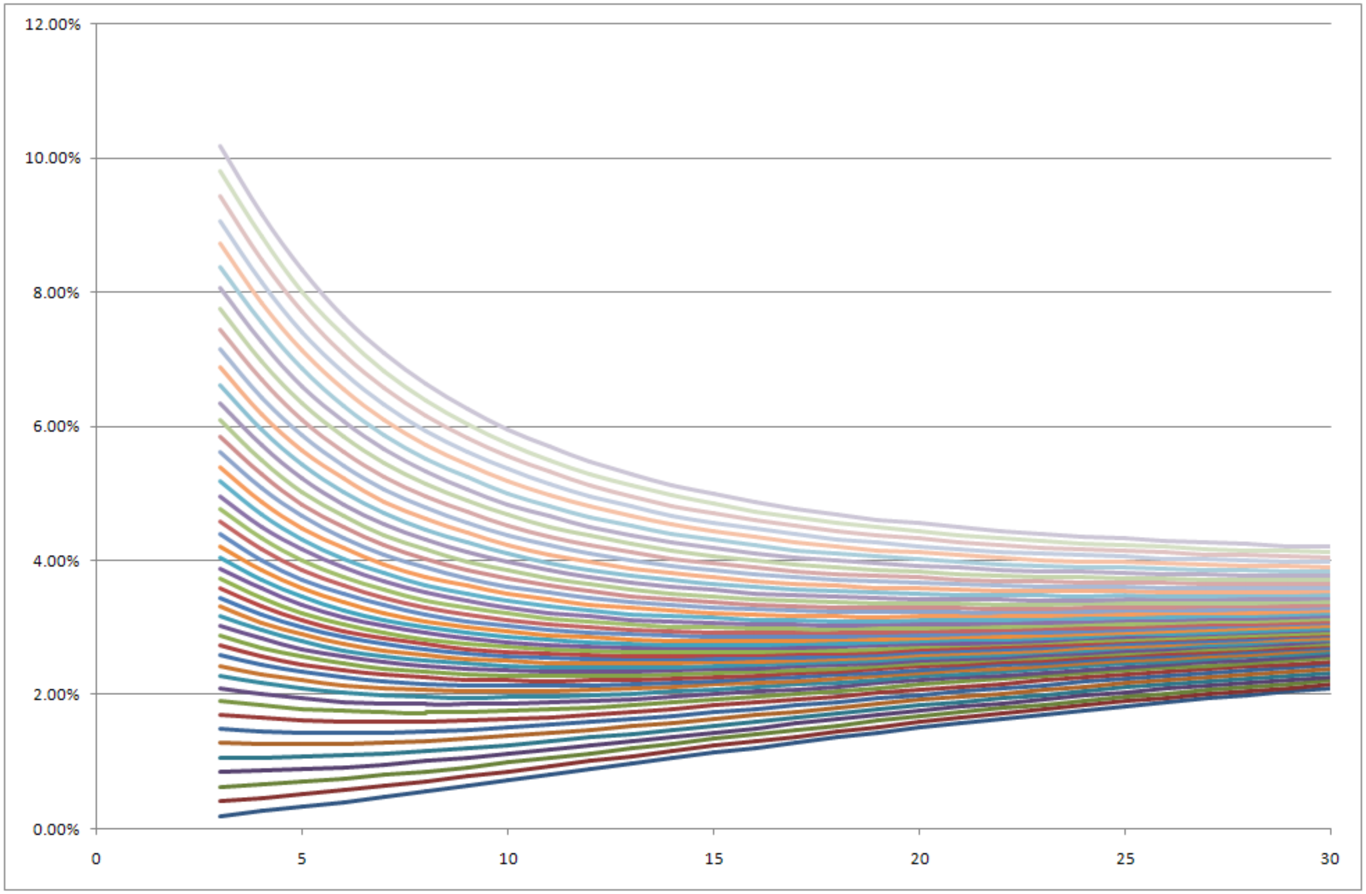
✓ Dynamic Conditioning

✓ Constructive Probability Theory

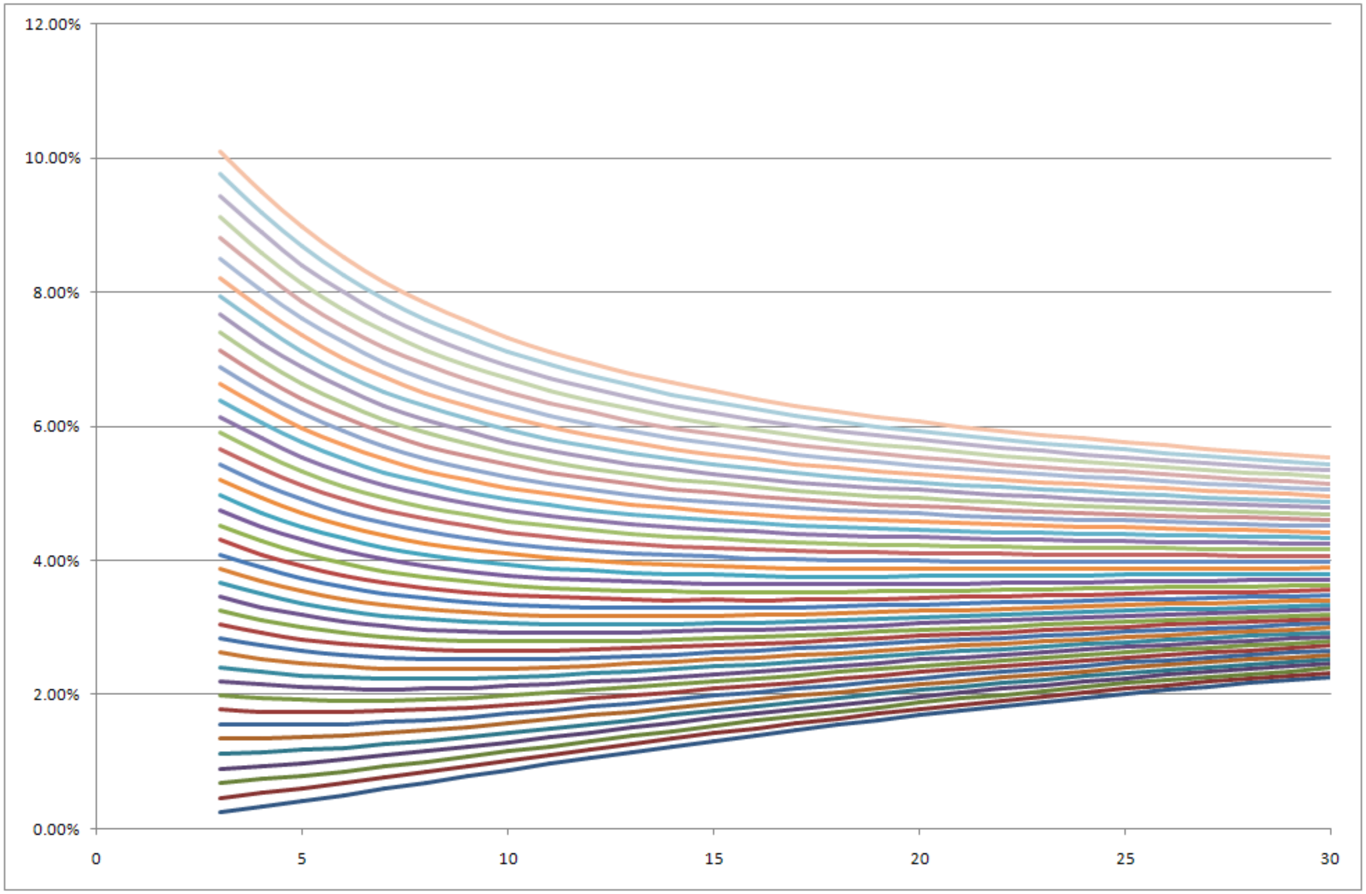
✓ Dynamic Conditioning Model for CDOs

◊ **Stochastic Monetary Policy Model for Interest Rate Derivatives**

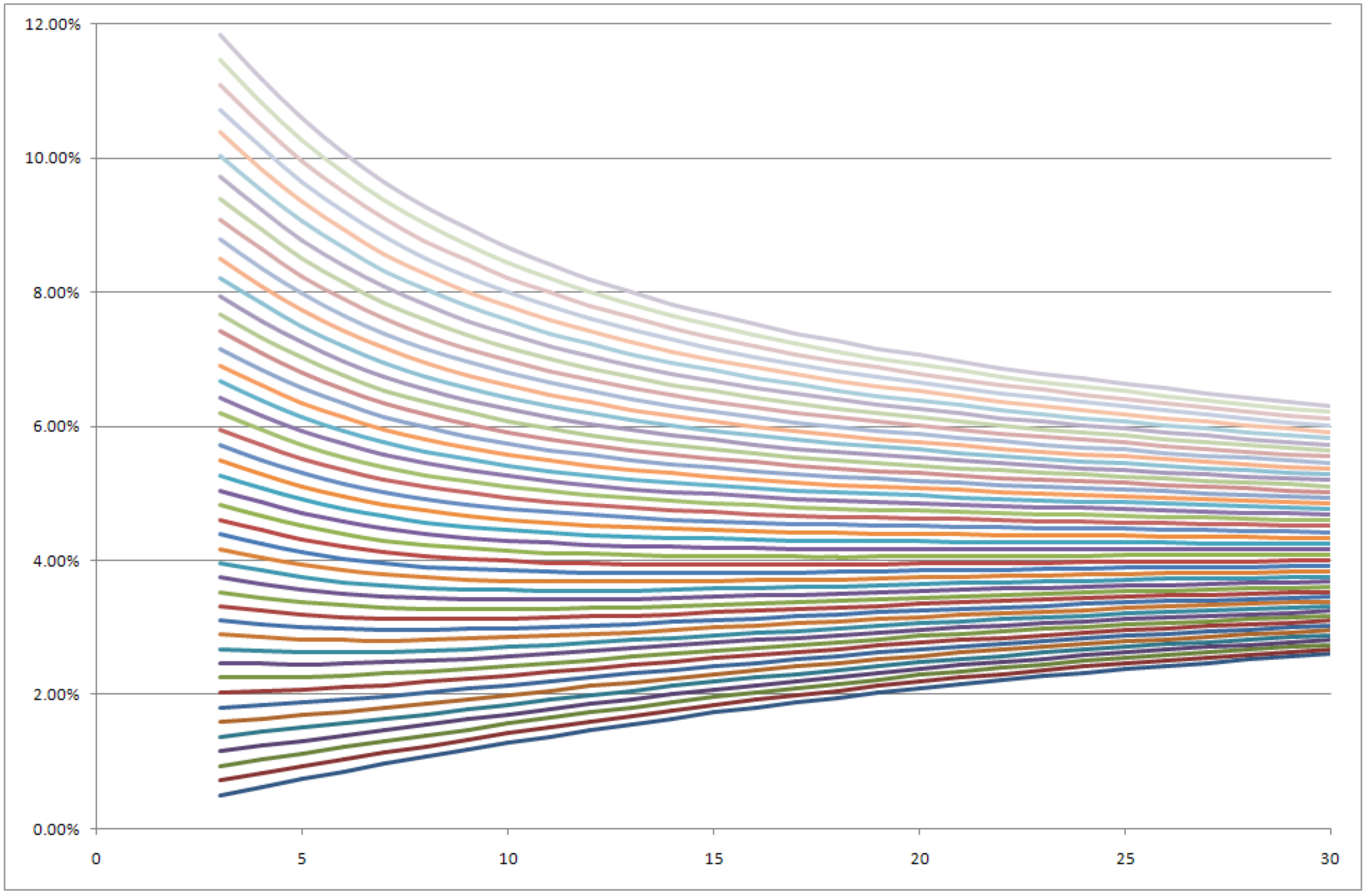
Yield curves in the deflation regime



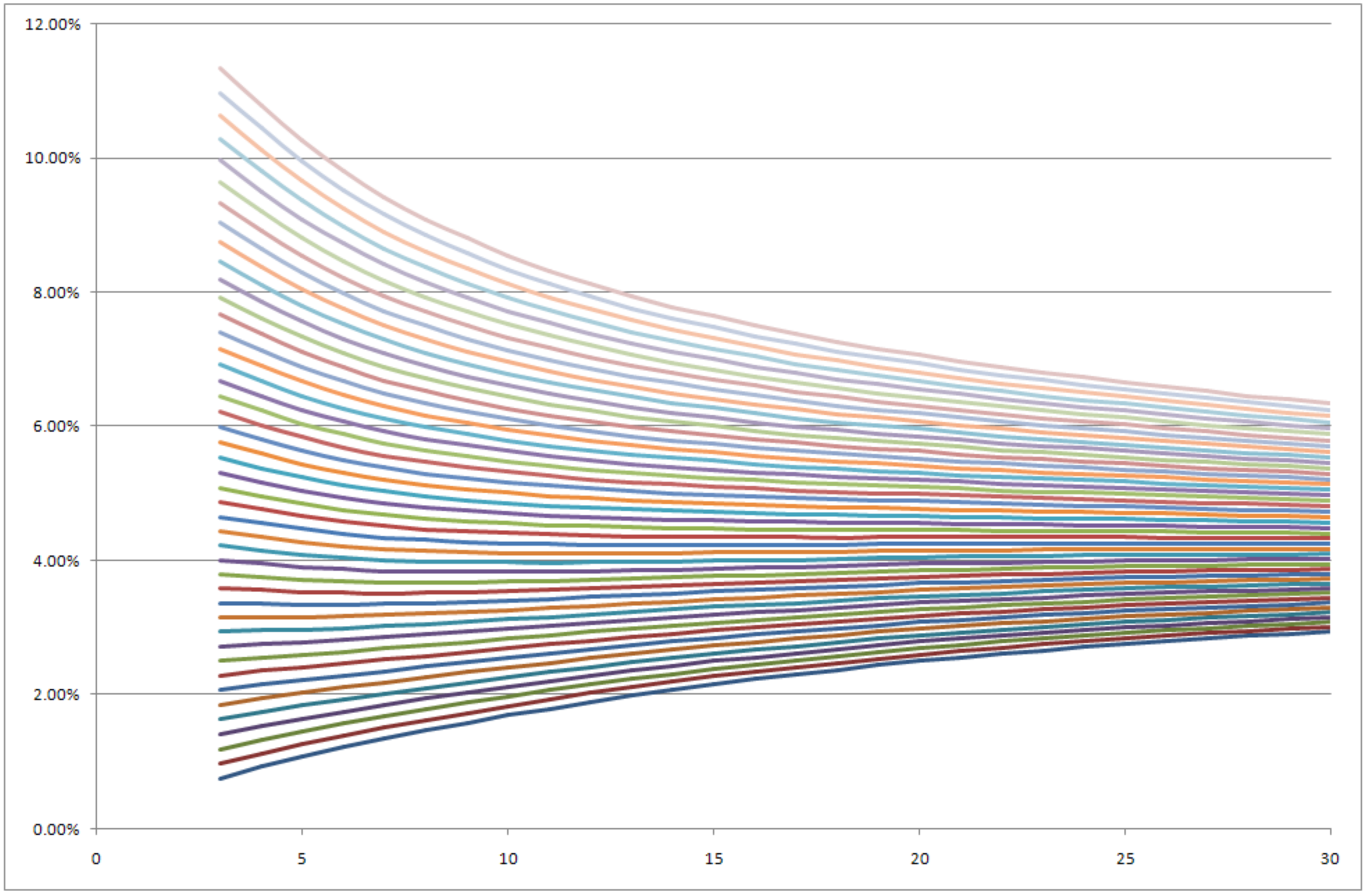
Yield curves with rates falling 75bp/y on average



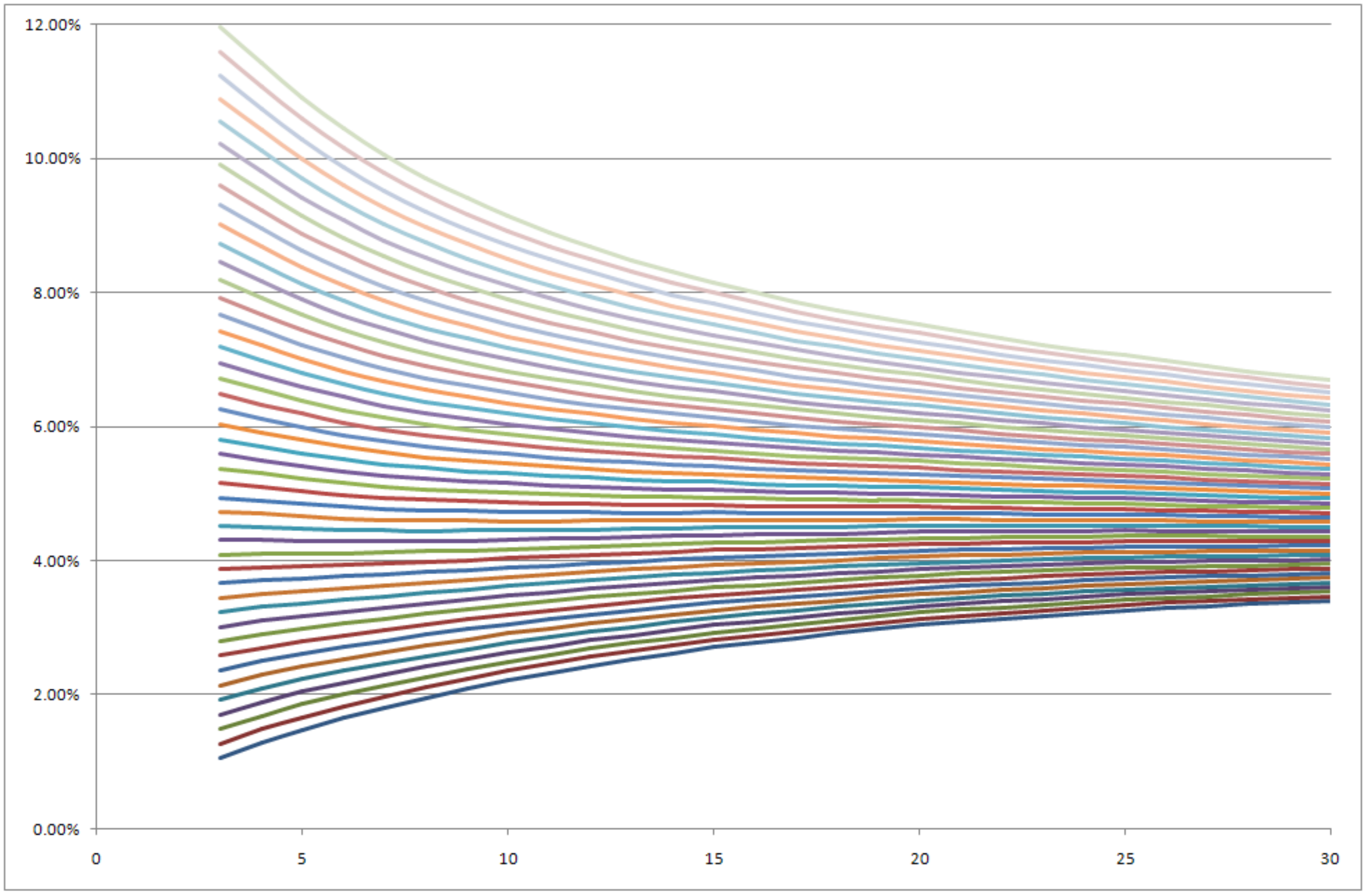
Yield curves with rates falling 50bp/y on average



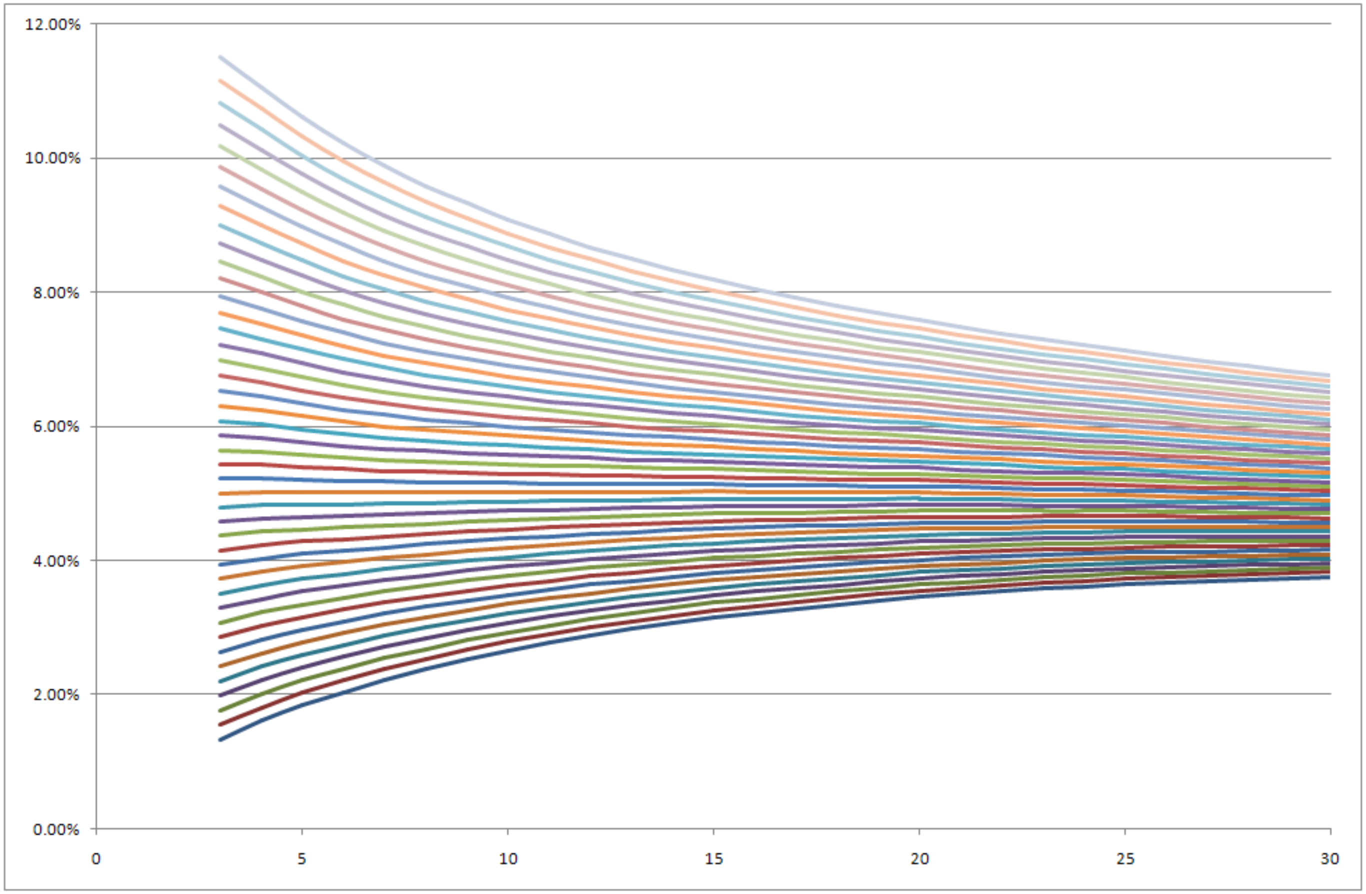
Yield curves with rates falling 25bp/y on average



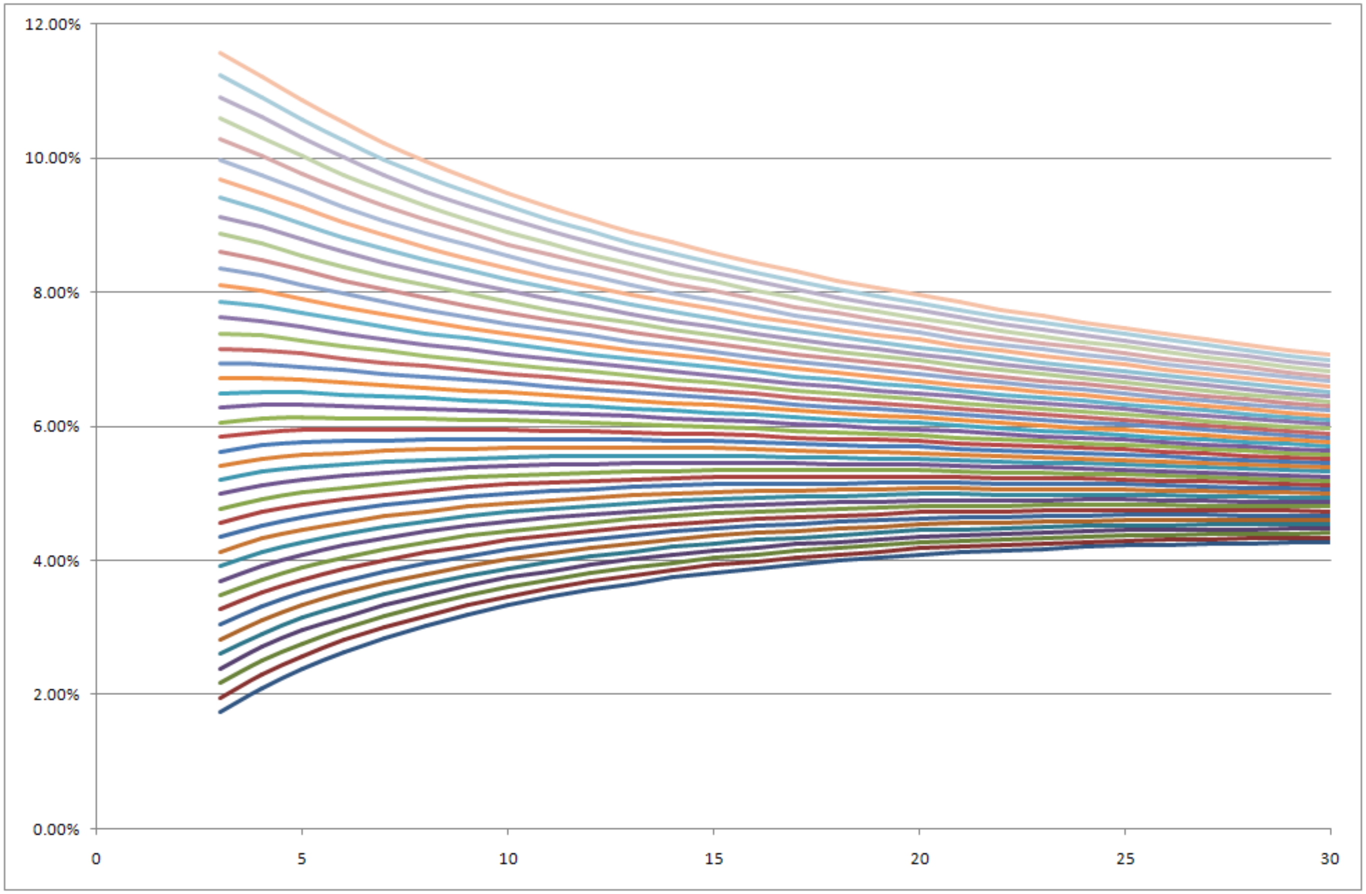
Yield curves with rates stable on average



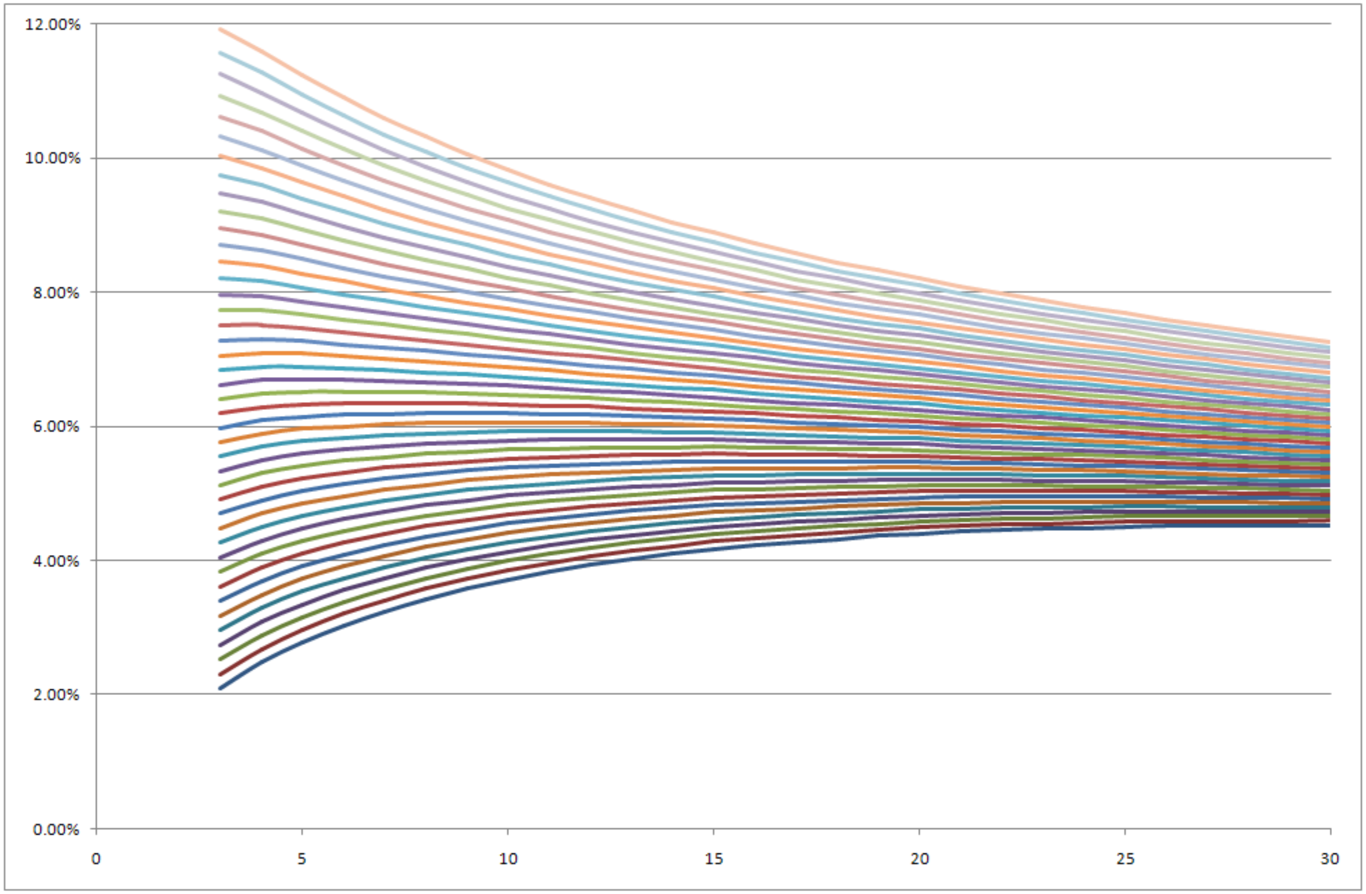
Yield curves with rates rising 25bp/y on average



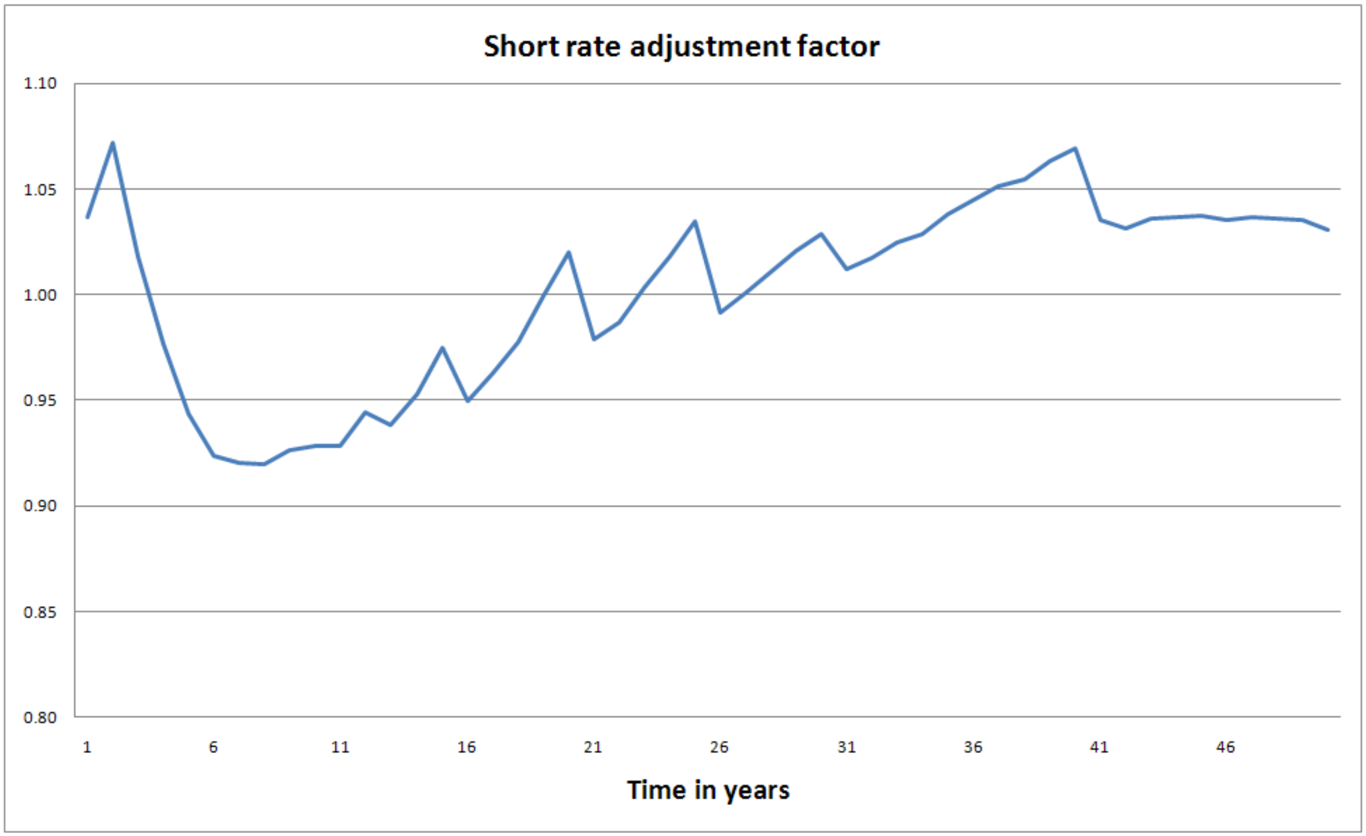
Yield curves with rates rising 50bp/y on average



Yield curves with rates rising 100bp/y on average



Time dependent short rate factor adjustment to match term structure of interest rates



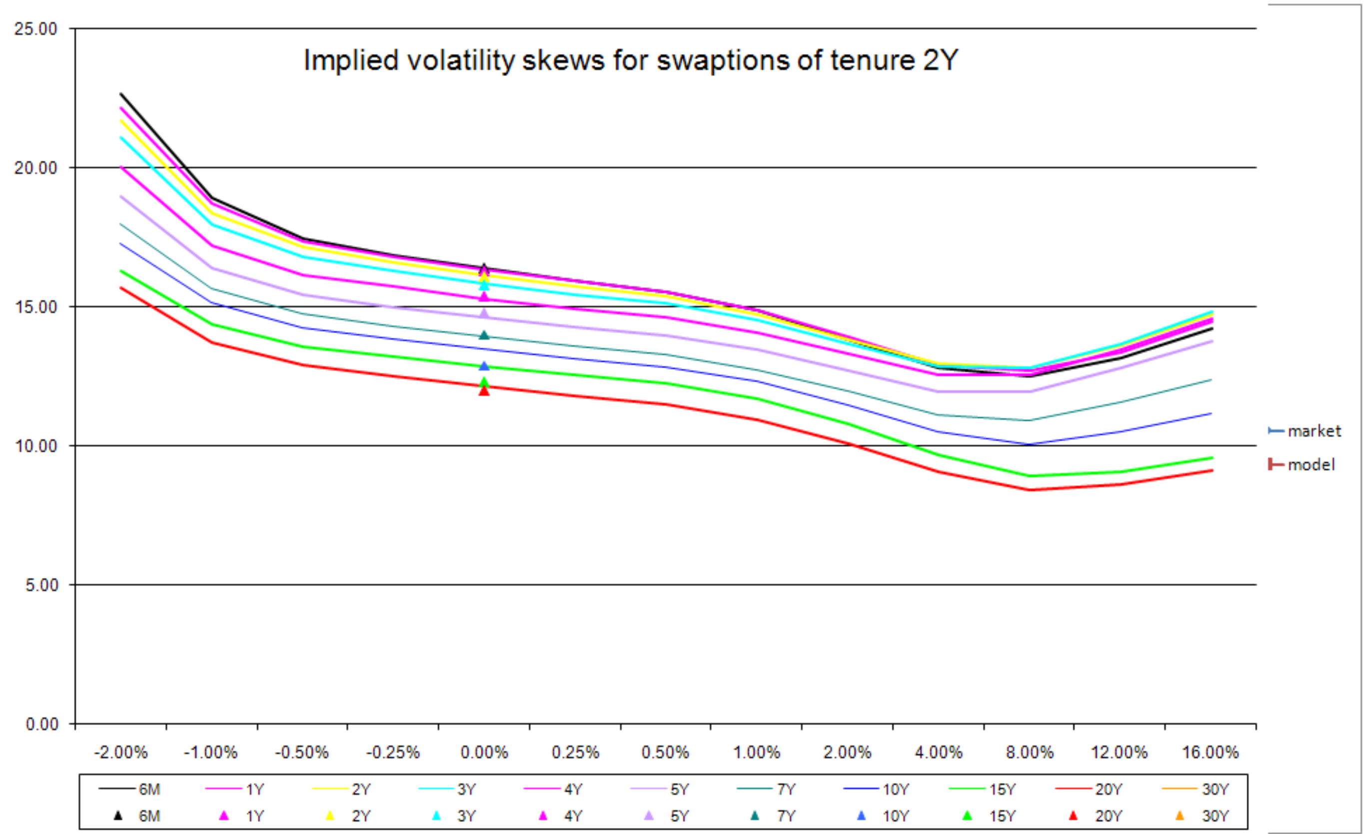
Global fit of the swaption volatility cube

Our model calibrates fairly well to the entire swaption volatility cube. Prior to our work, only the stochastic volatility BGM models (Pieterbarg) succeeded to achieve a global fit of similar quality. However, the economic interpretation of the resulting models is very different and results in substantial discrepancies when the model is applied to some long dated callable swaps.

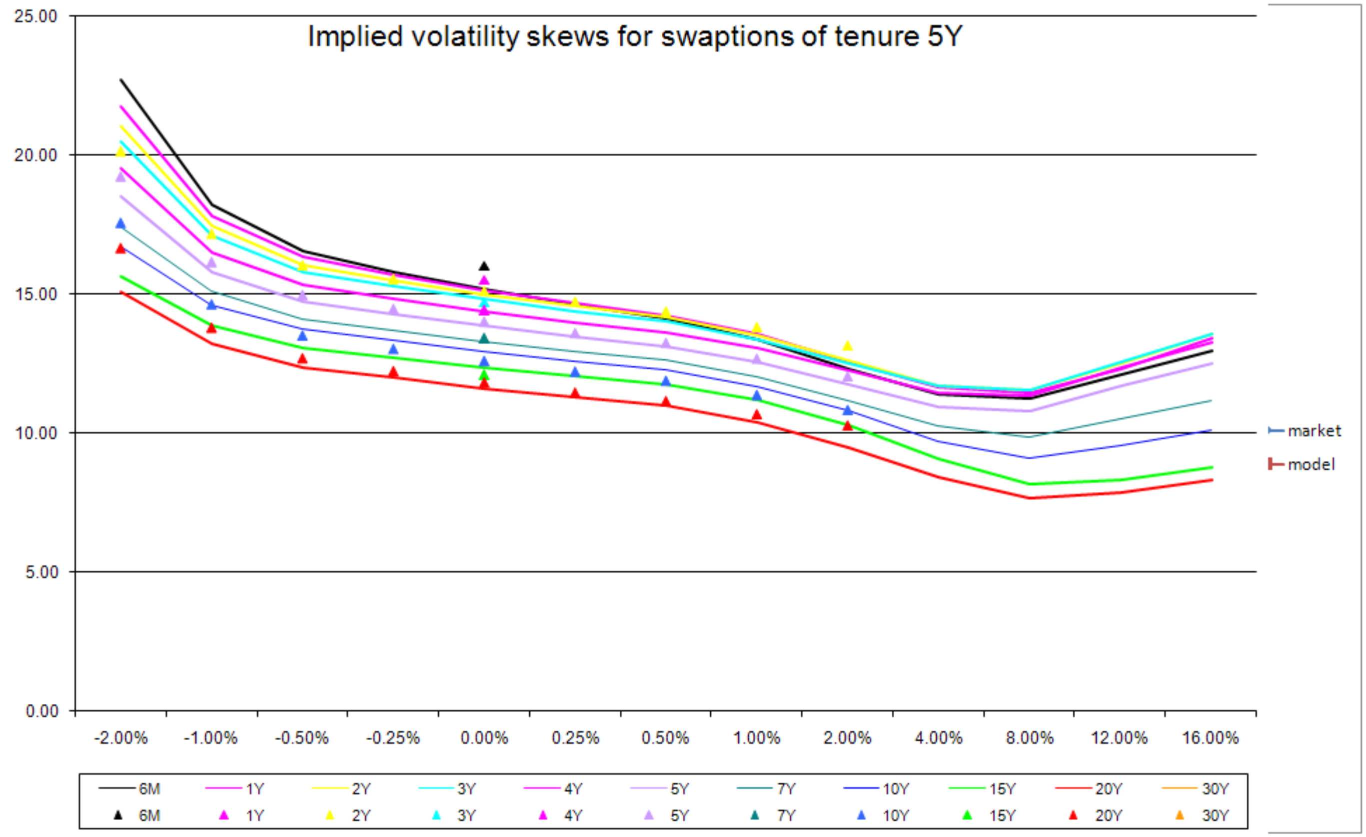
Calibration Errors for ATM swaptions

ATM Swaptions (vol error)															
	Tenor														
Expiry	0	1	2	3	4	5	6	7	8	9	10	15	20	25	30
	1	-1.21	-0.43	0.14	0.53	0.64	0.50	0.44	0.25	0.14	0.02	-0.26	-0.11	0.23	0.70
	2	-0.54	0.04	0.34	0.42	0.34	0.29	0.22	0.13	0.12	0.00	-0.29	-0.13	0.30	0.76
	3	-0.15	0.20	0.28	0.33	0.26	0.22	0.14	0.14	0.03	0.00	-0.19	0.06	0.49	0.95
	4	0.16	0.21	0.25	0.20	0.21	0.16	0.08	0.08	0.06	0.03	-0.08	0.16	0.58	1.03
	5	0.26	0.14	0.15	0.09	0.08	0.12	0.13	0.12	0.09	0.06	0.03	0.26	0.67	1.11
	6	0.00	-0.04	0.01	0.00	0.02	0.05	0.04	0.07	0.04	0.04	0.04	0.30	0.72	1.16
	7	-0.25	-0.20	-0.12	-0.09	-0.03	-0.03	-0.04	0.03	-0.02	0.02	0.04	0.34	0.76	1.19
	8	-0.29	-0.22	-0.18	-0.12	-0.12	-0.08	-0.10	-0.06	-0.06	-0.01	-0.03	0.27	0.72	1.11
	9	-0.34	-0.24	-0.22	-0.18	-0.21	-0.15	-0.20	-0.17	-0.09	-0.05	-0.08	0.20	0.67	1.04
	10	-0.39	-0.27	-0.29	-0.22	-0.30	-0.24	-0.30	-0.27	-0.14	-0.11	-0.15	0.12	0.63	0.95
	20	0.68	0.28	0.23	0.25	0.25	0.21	0.27	0.34	0.41	0.50	0.43	0.60	1.05	1.58

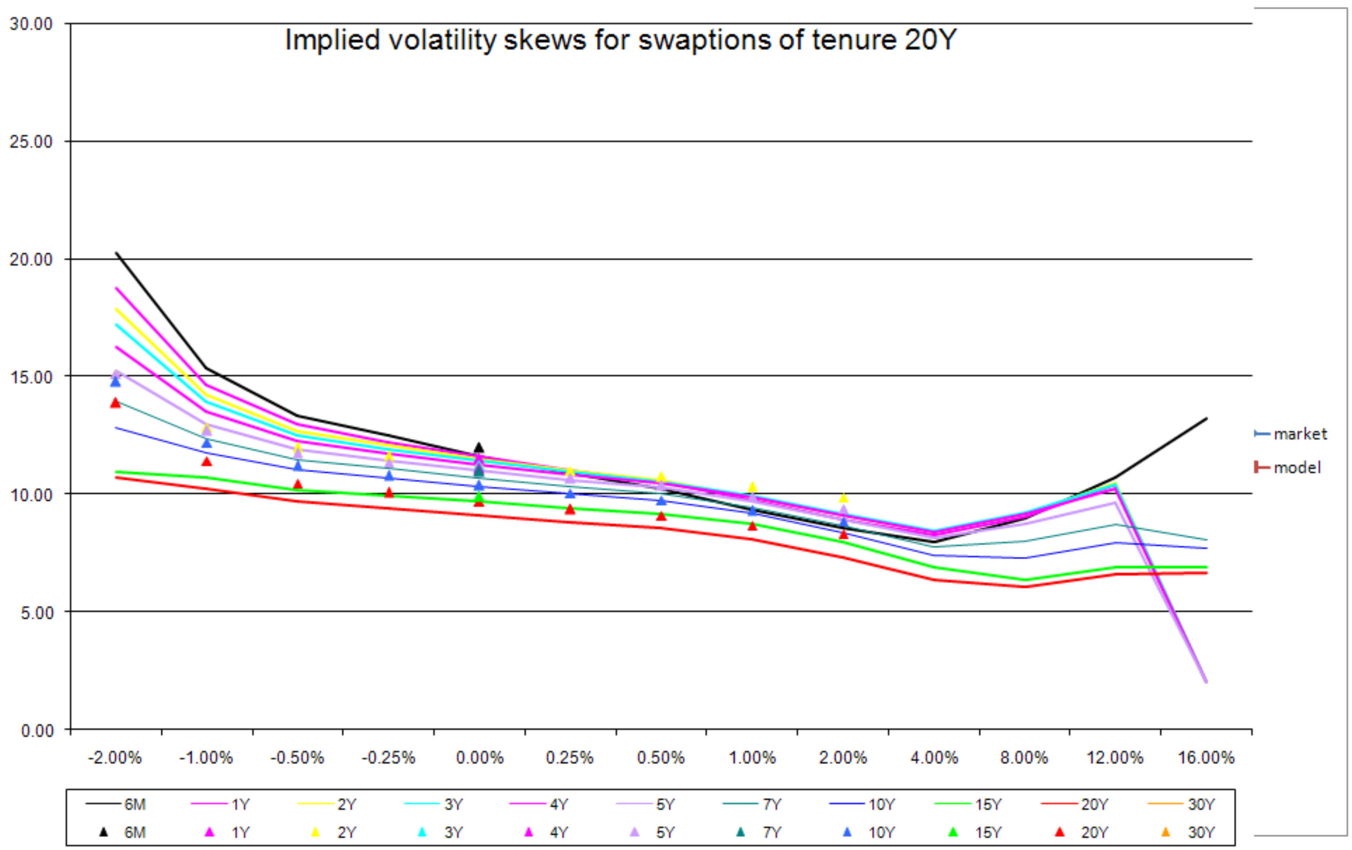
Swaption skews for tenor 2y



Swaption skews for tenor 5y



Swaption skews for tenor 20y



Interpretation of the persistent skew in Black volatilities

From the case of 20 into 30 swaptions implied vols we learn that the inclusion of deflation scenarios in the stochastic drift dynamics is crucial to be able to reproduce implied volatilities for these swaptions in the region of 14% for this particular dataset. Without deflation and without spoiling the other regions of the swaption volatility cube only volatilities of 7-8% can be obtained.

This example shows that implied Black volatilites are best interpreted as stochastic drift effects as opposed to stochastic volatility effects.

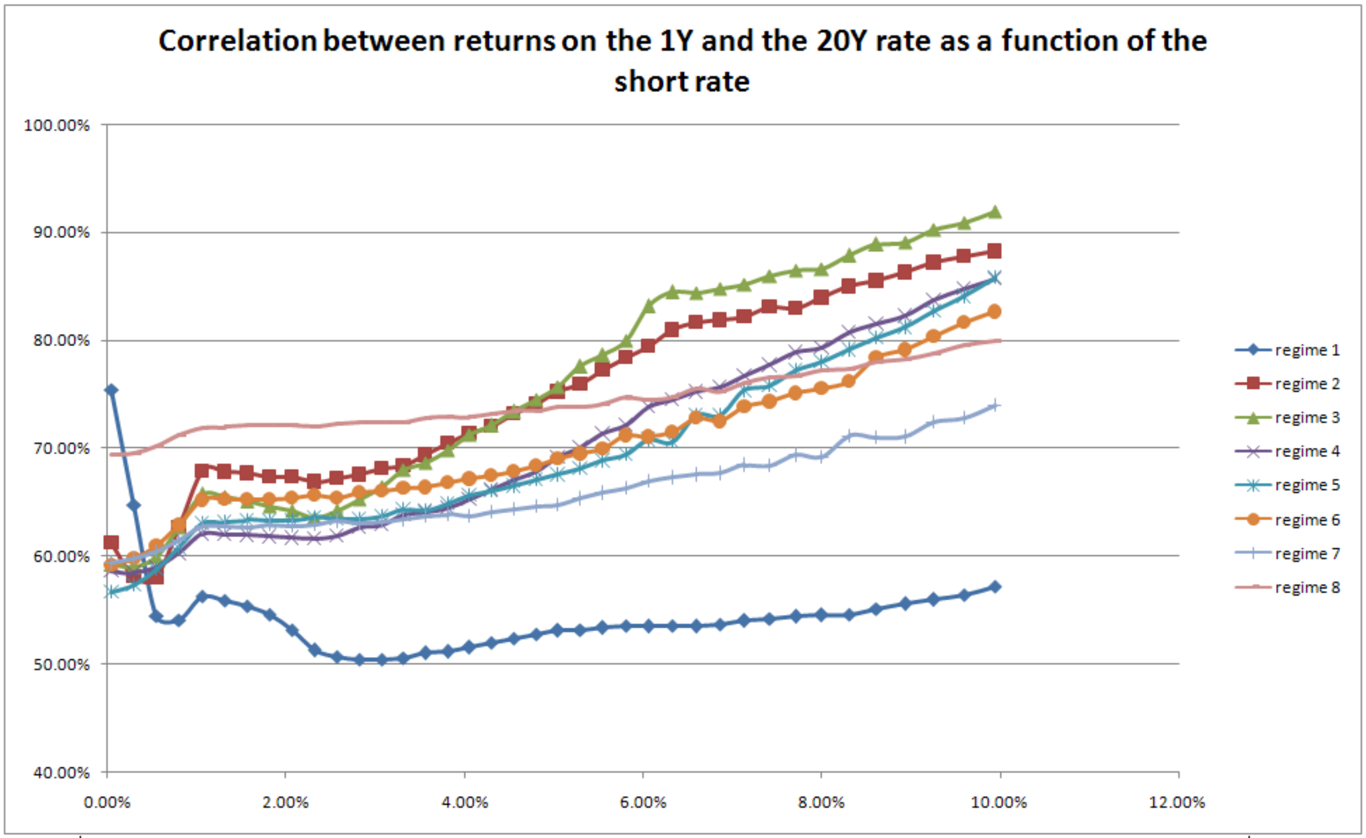
The example also shows that the interpretation of the drift dynamics tends to be economically meaningful.

Correlation matrix

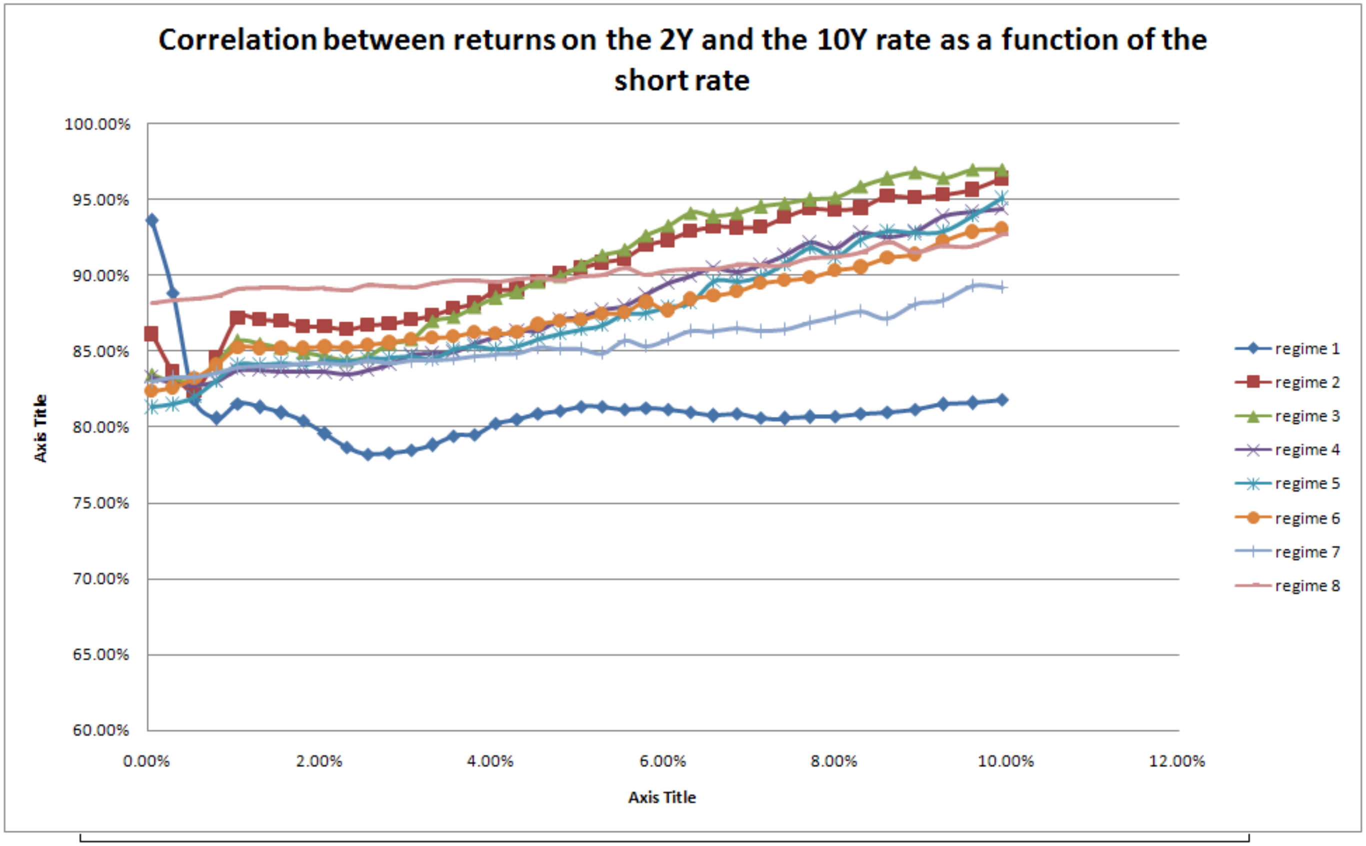
What we notice is that once all implied swaption volatilities are matched, short term correlation also fall in line with historical estimates.

On the other hand, short term volatility and correlations may be right while the vol-cube is far from being correctly reproduced.

Correlation between the 1y and the 20y rate

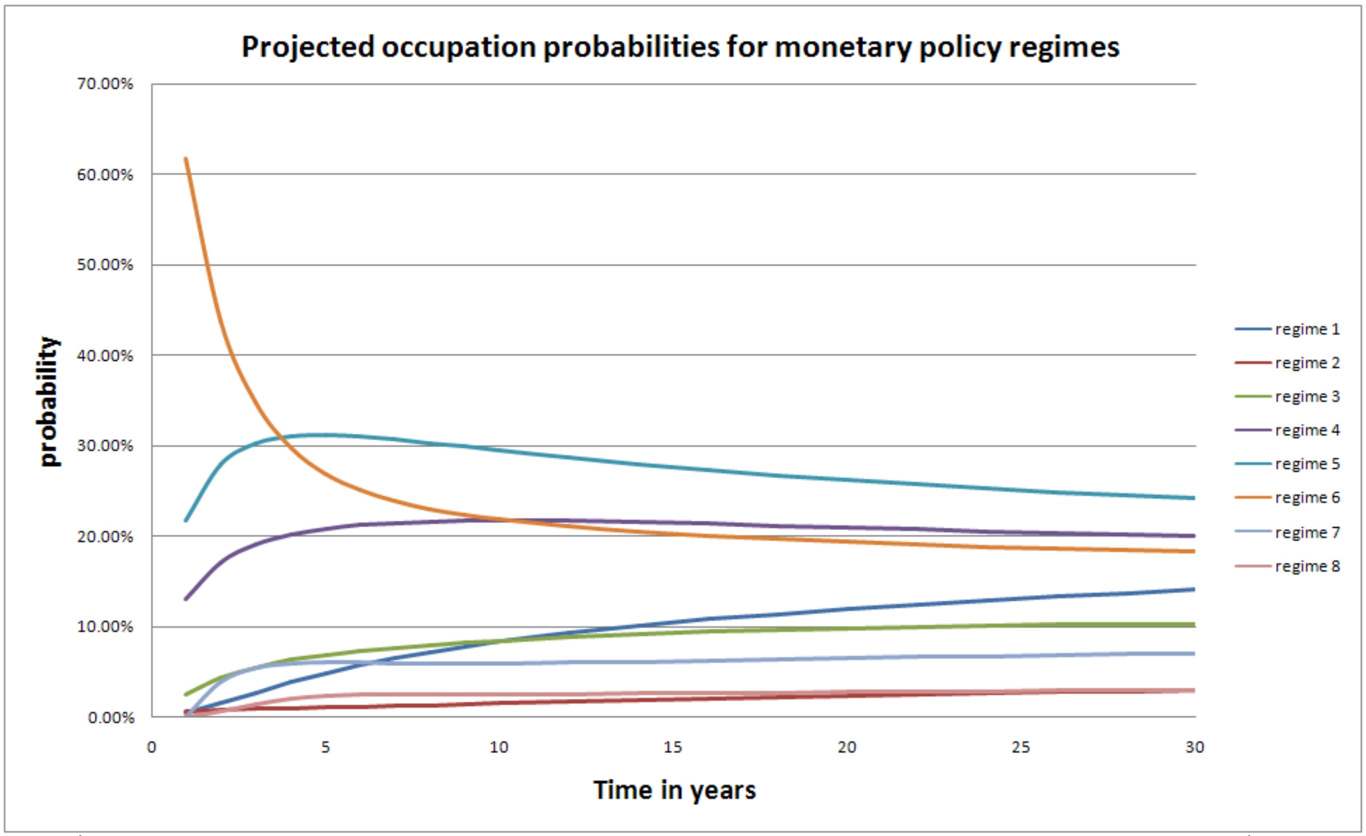


Correlation between the 2y and the 10y rate



Expected Monetary Policy

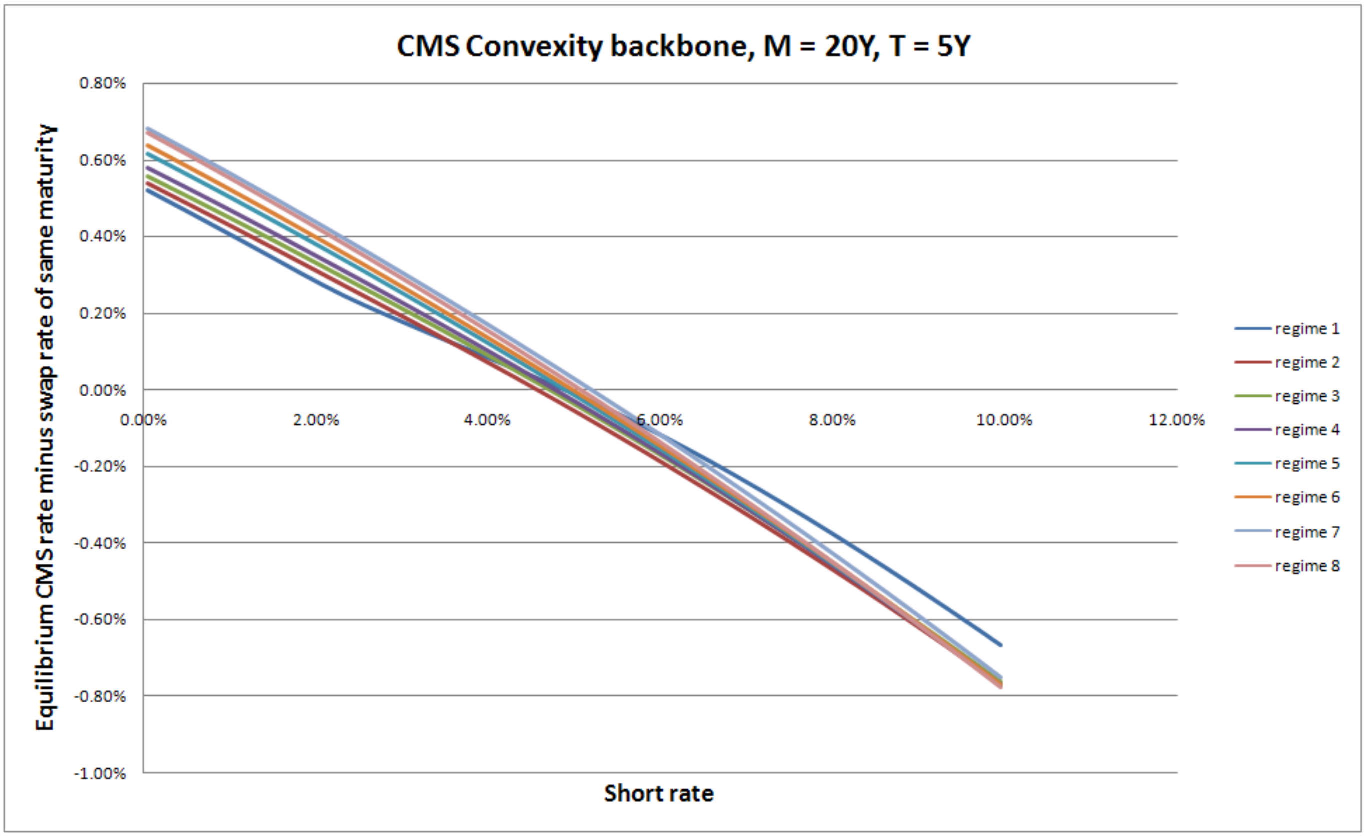
The graph below gives the unconditional expectation of monetary policy regime as a function of time going up to 80 years:



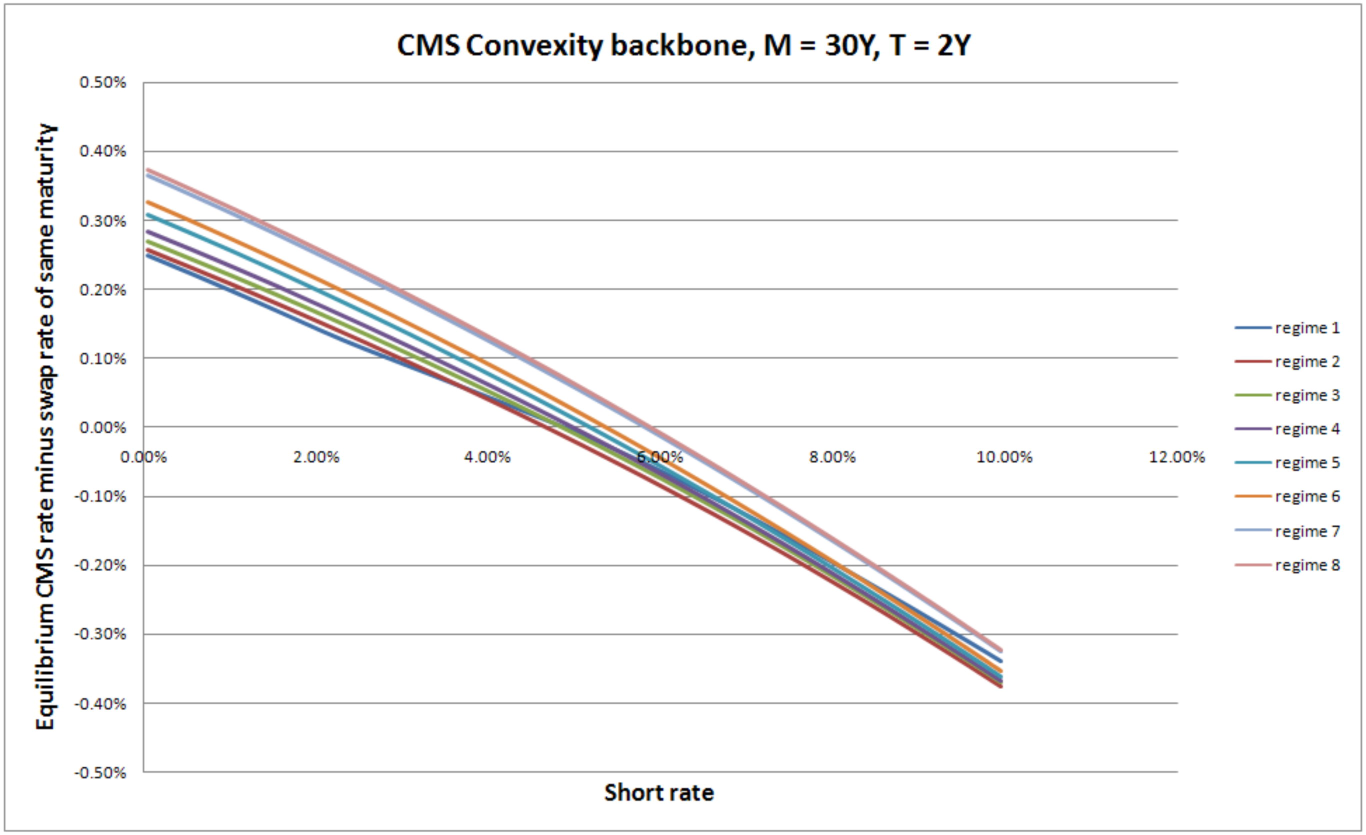
CMS convexity corrections

The CMS convexity correction is the difference between the equilibrium spread for a non-callable CMS swap and the plain vanilla LIBOR swap spread for the same maturity. This convexity corrections depends on monetary policy and the interest rate level.

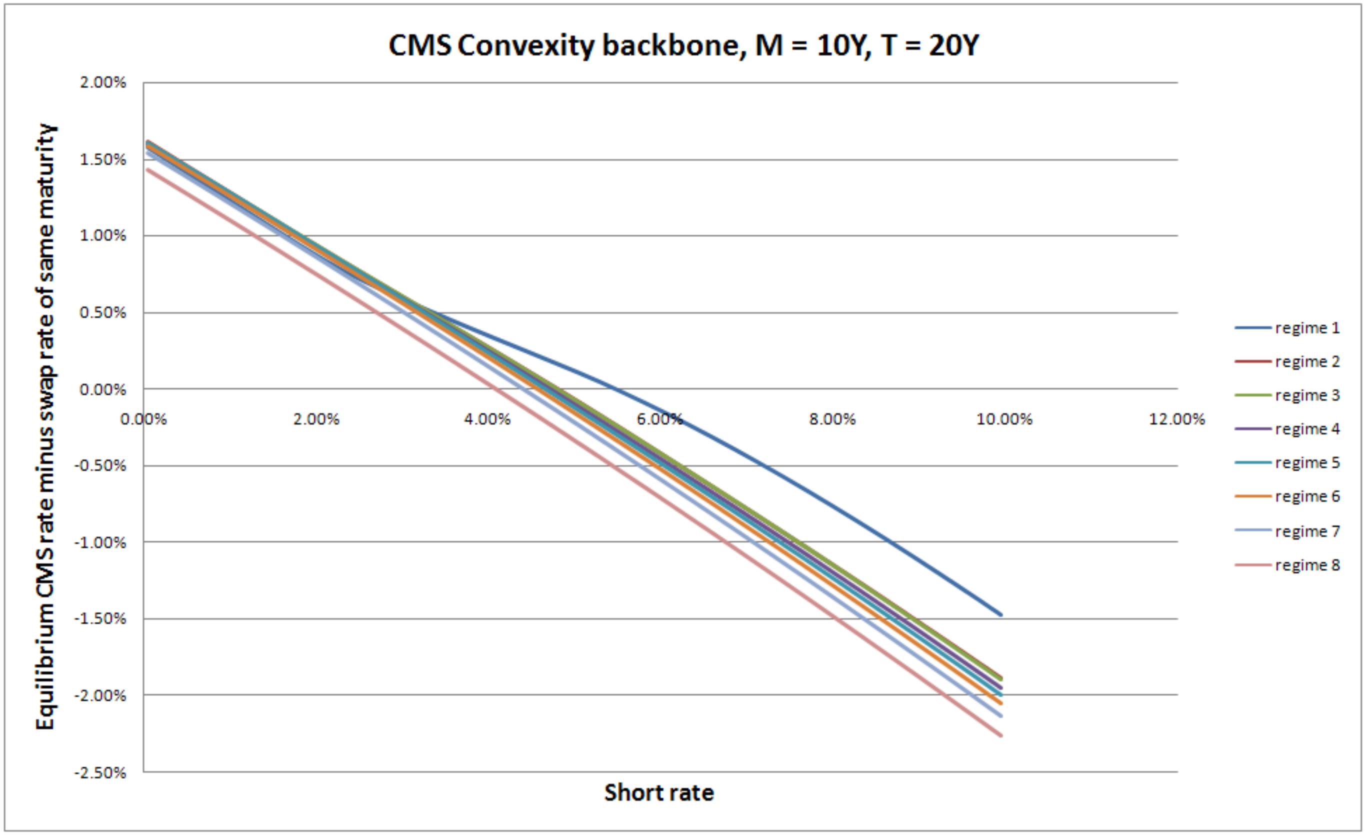
CMS convexity corrections



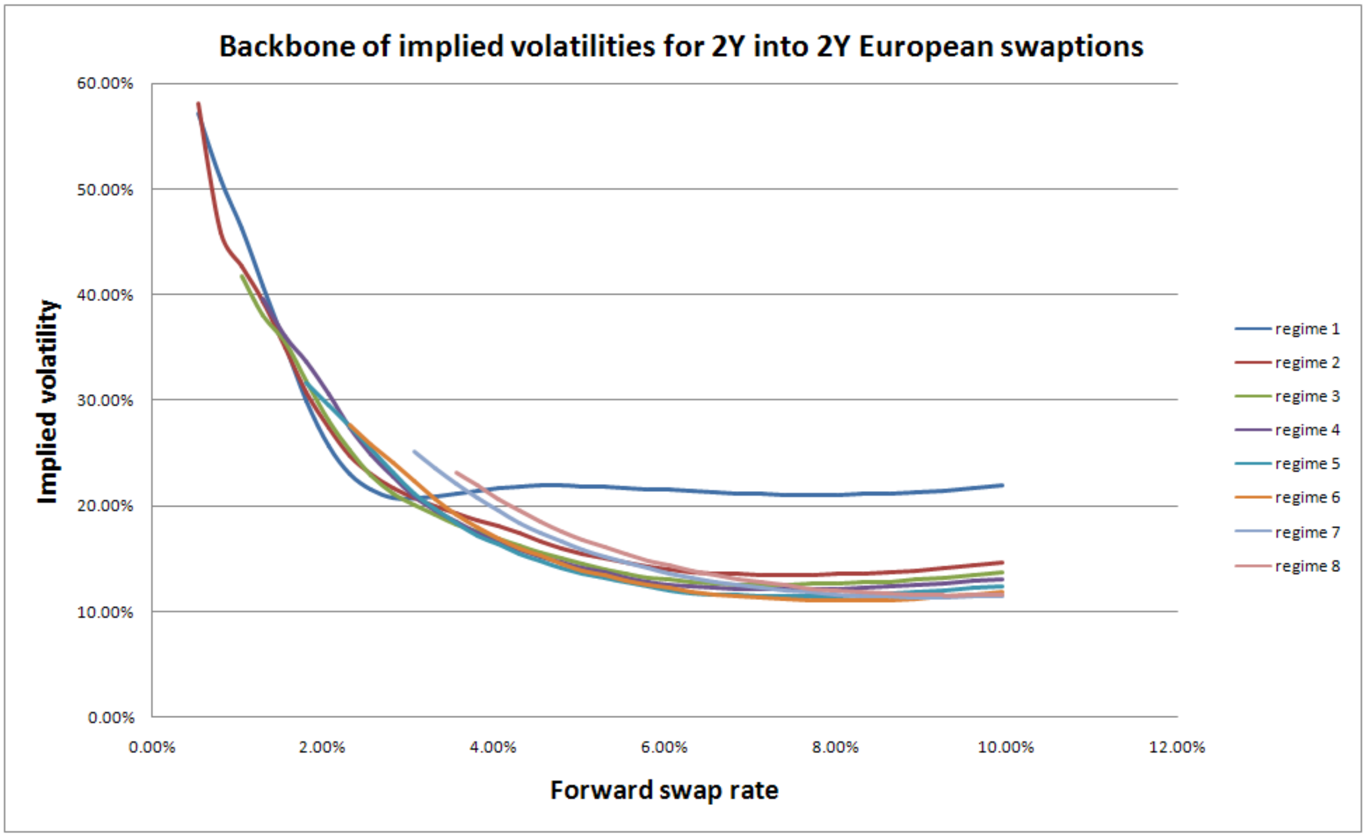
CMS convexity corrections



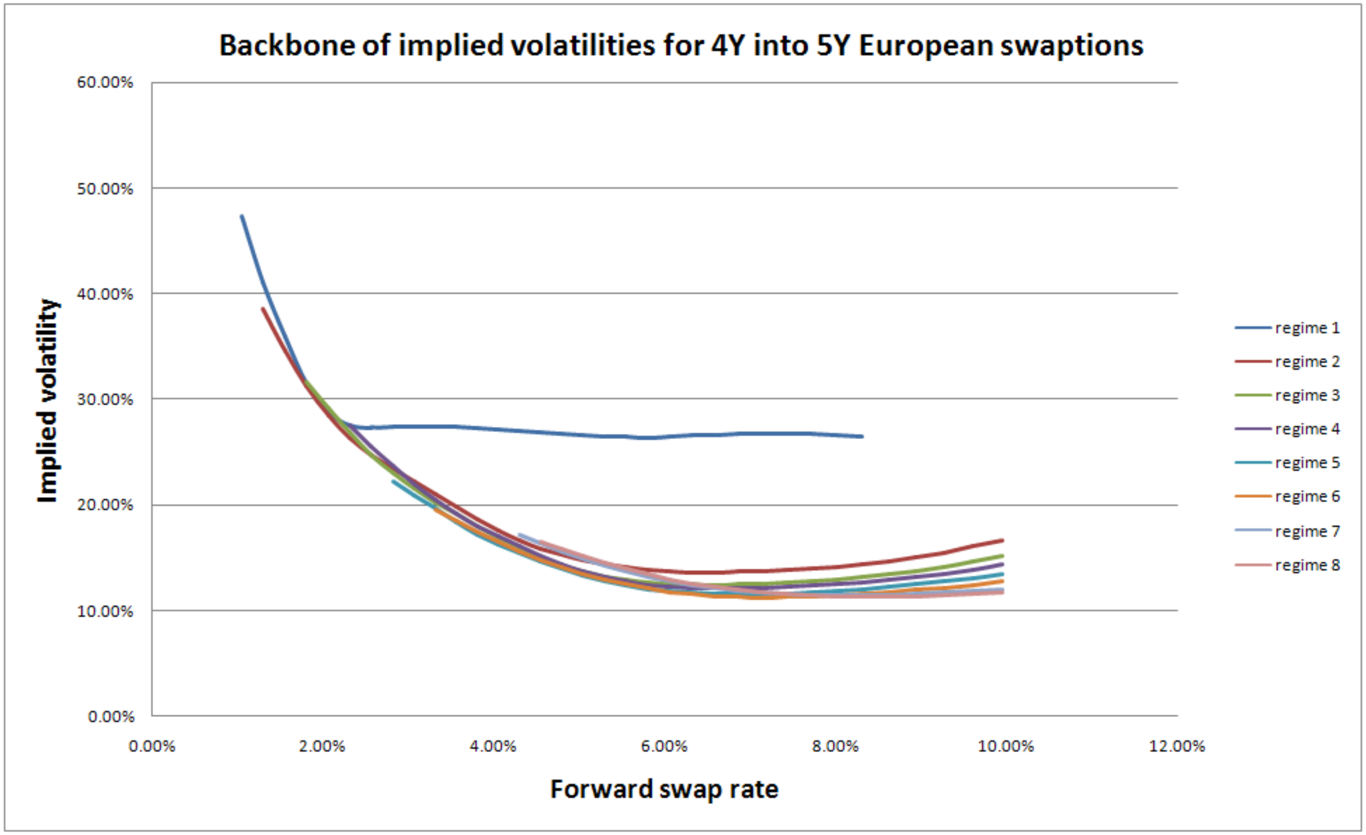
CMS convexity corrections



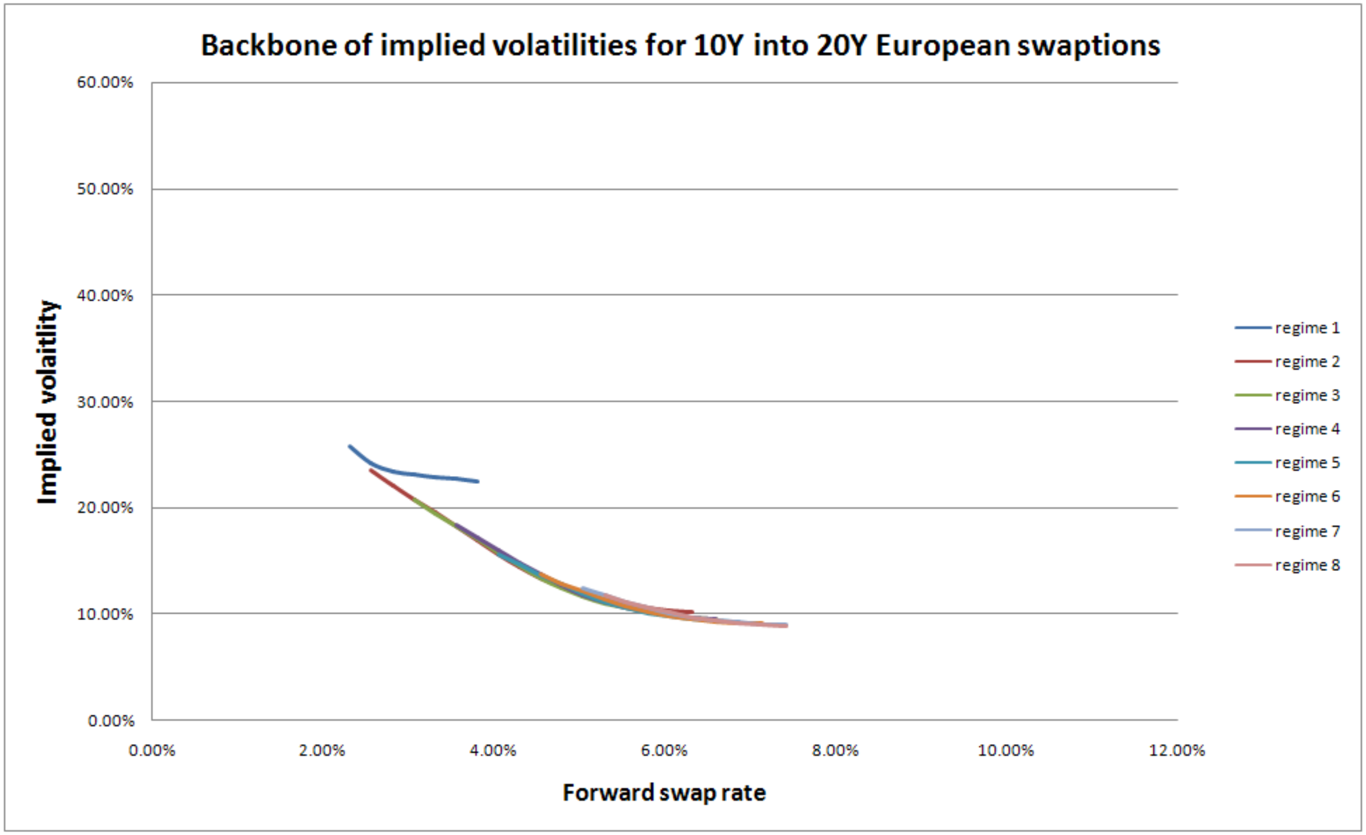
Swaption backbone for 2 into 2Y swaptions



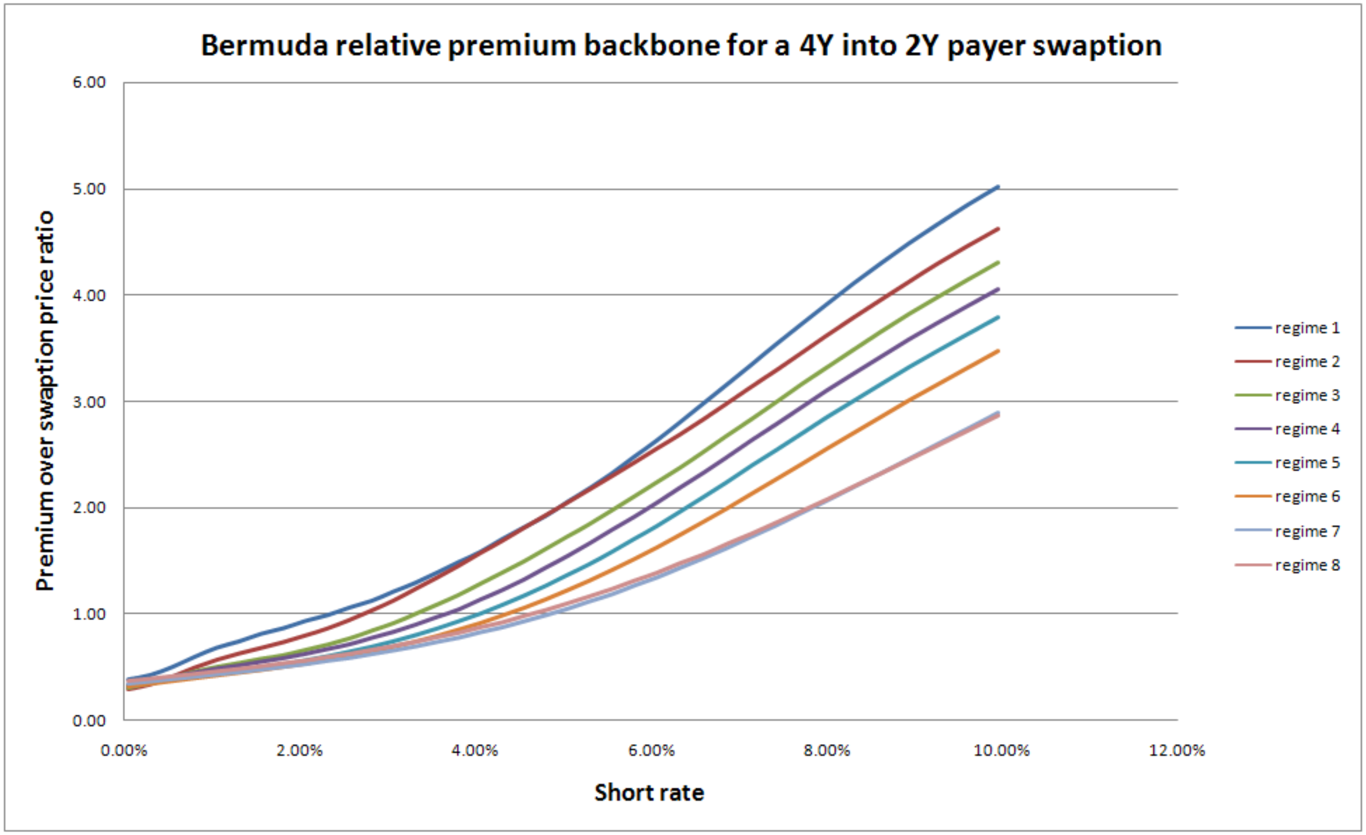
Swaption backbone for 4 into 5Y swaptions



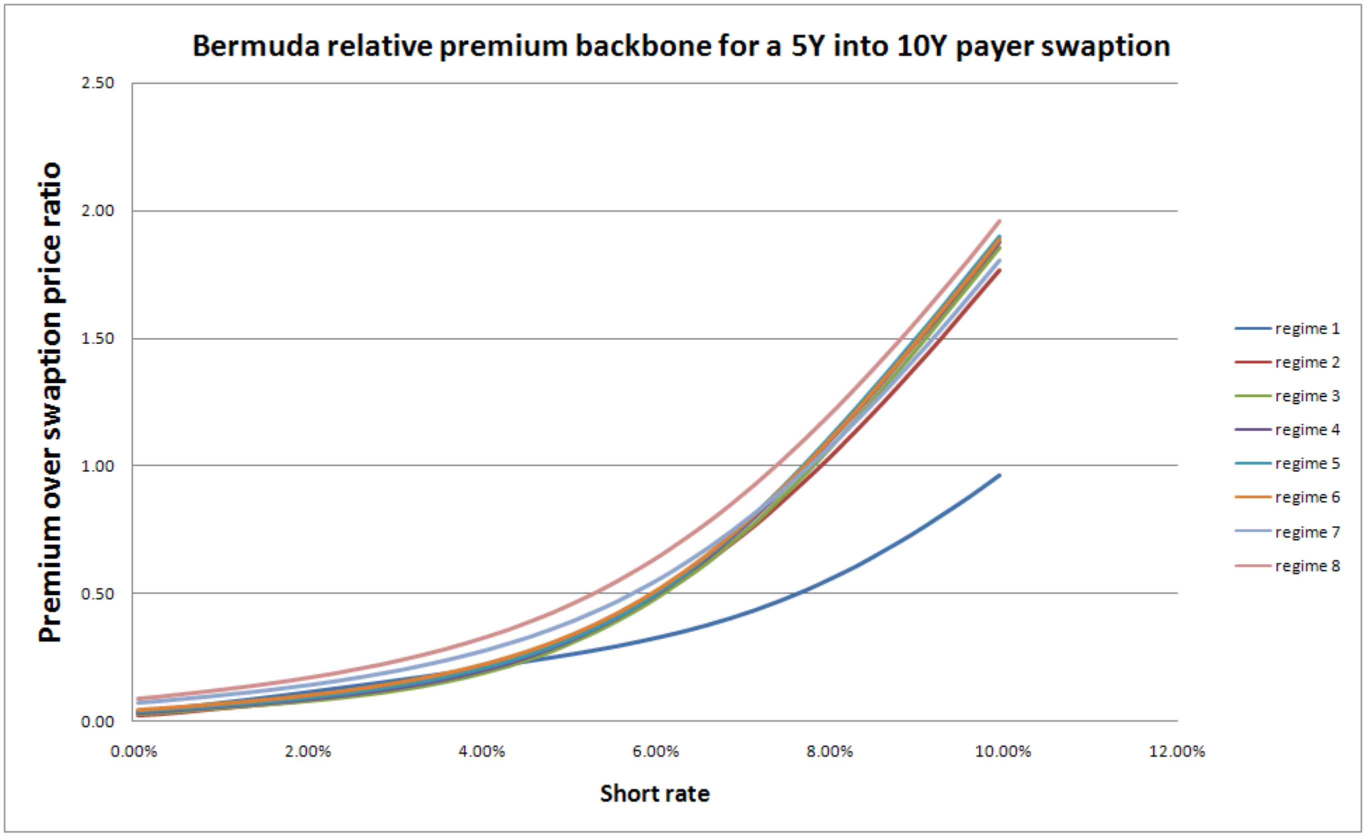
Swaption backbone for 10 into 20Y swaptions



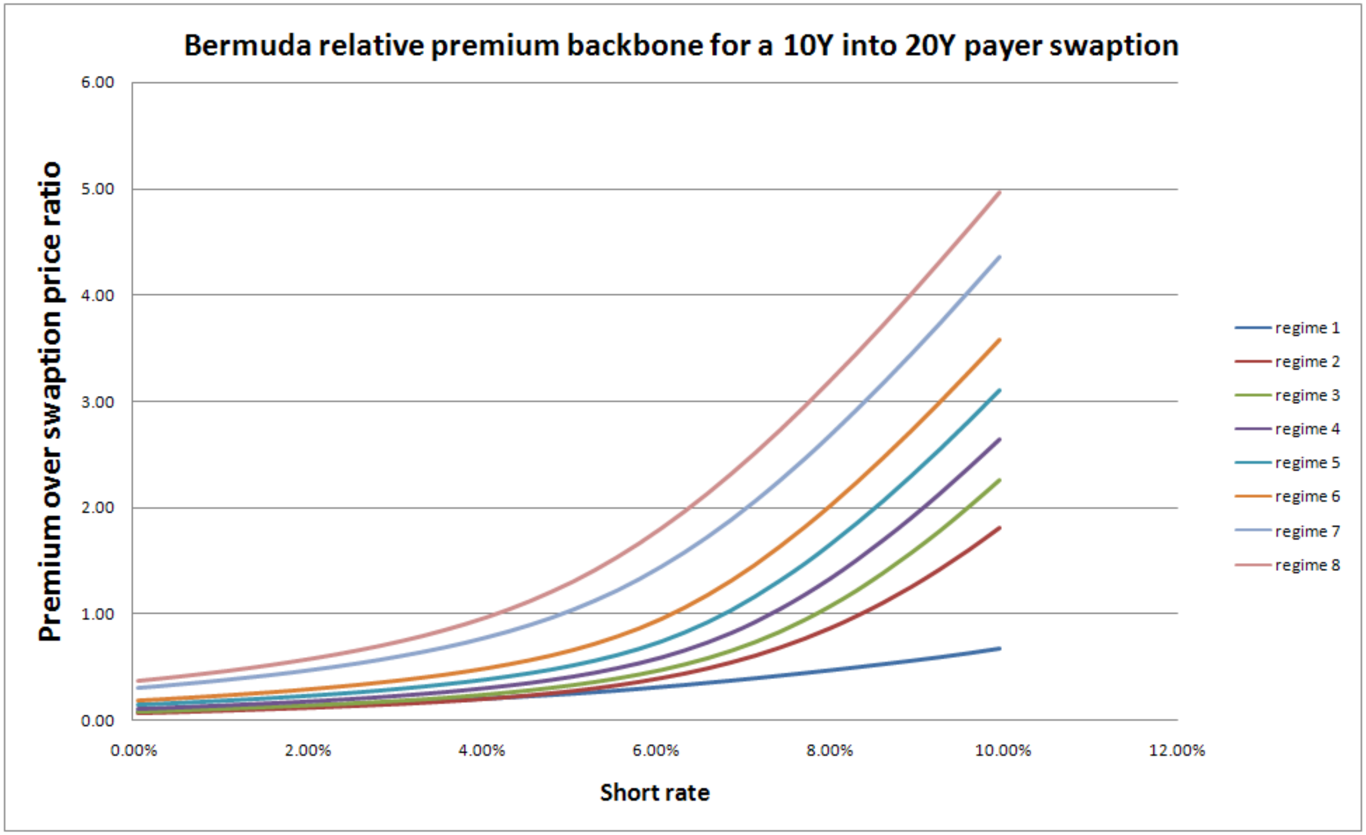
Bermuda backbone for 4 into 2Y swaptions



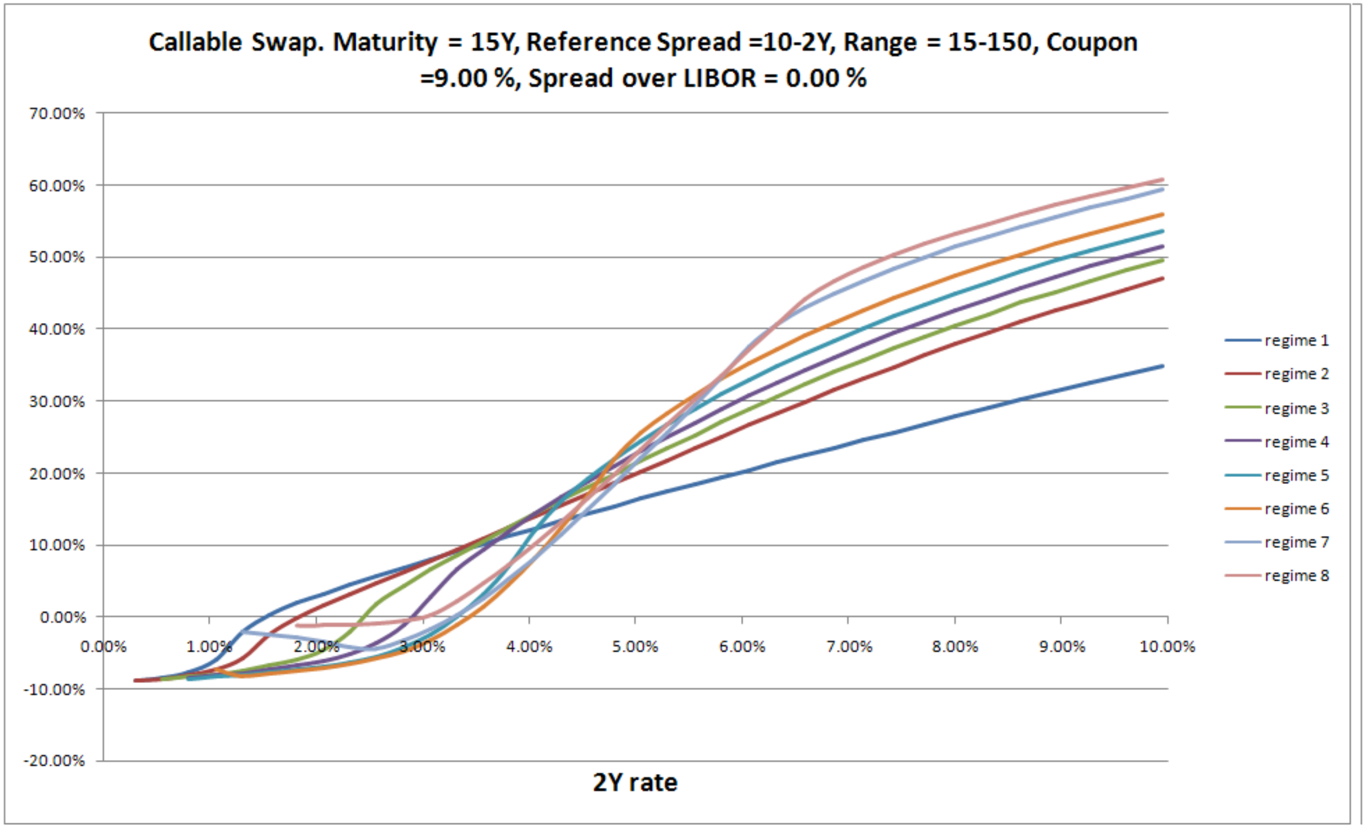
Bermuda backbone for 5 into 10Y swaptions



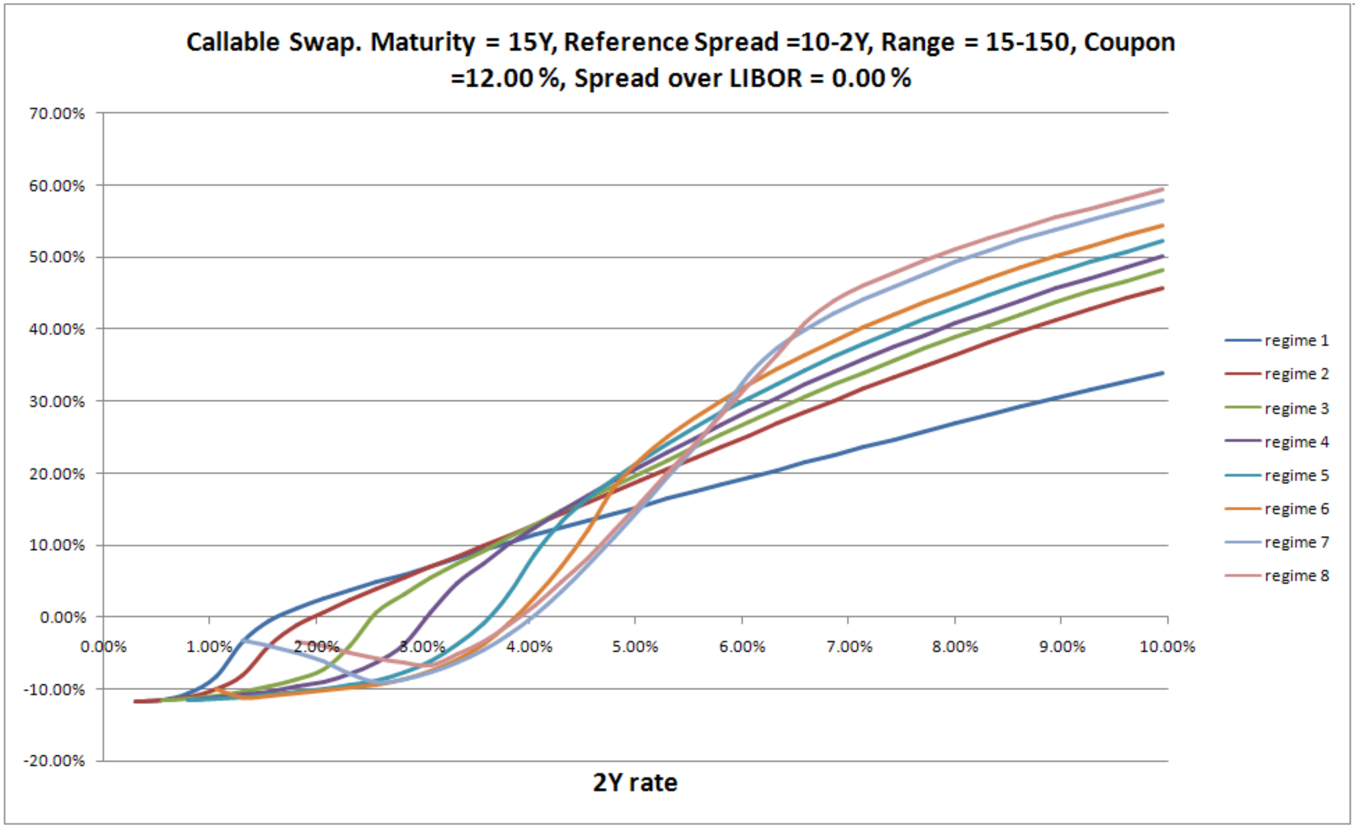
Bermuda backbone for 10 into 20Y swaptions



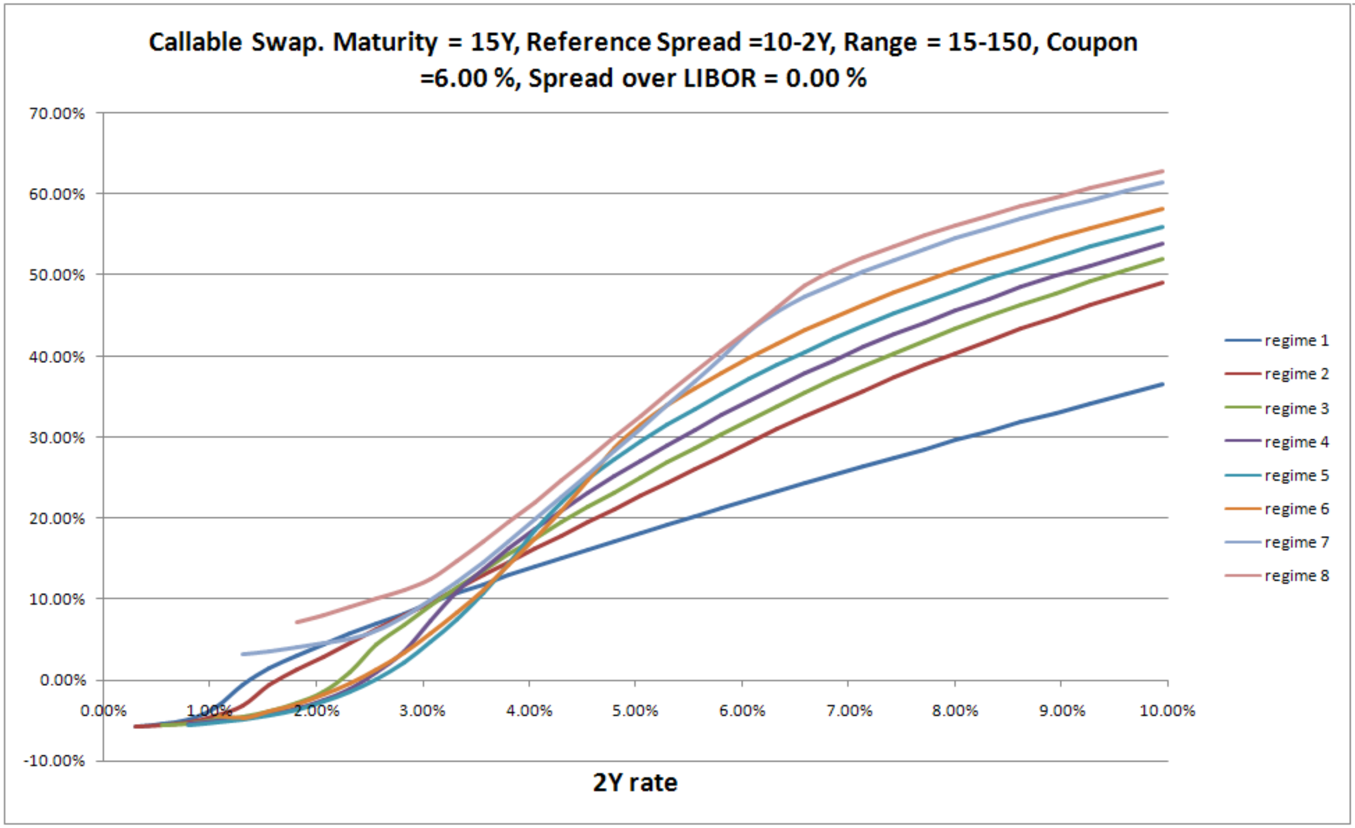
A CMS callable spread range accrual, example 1



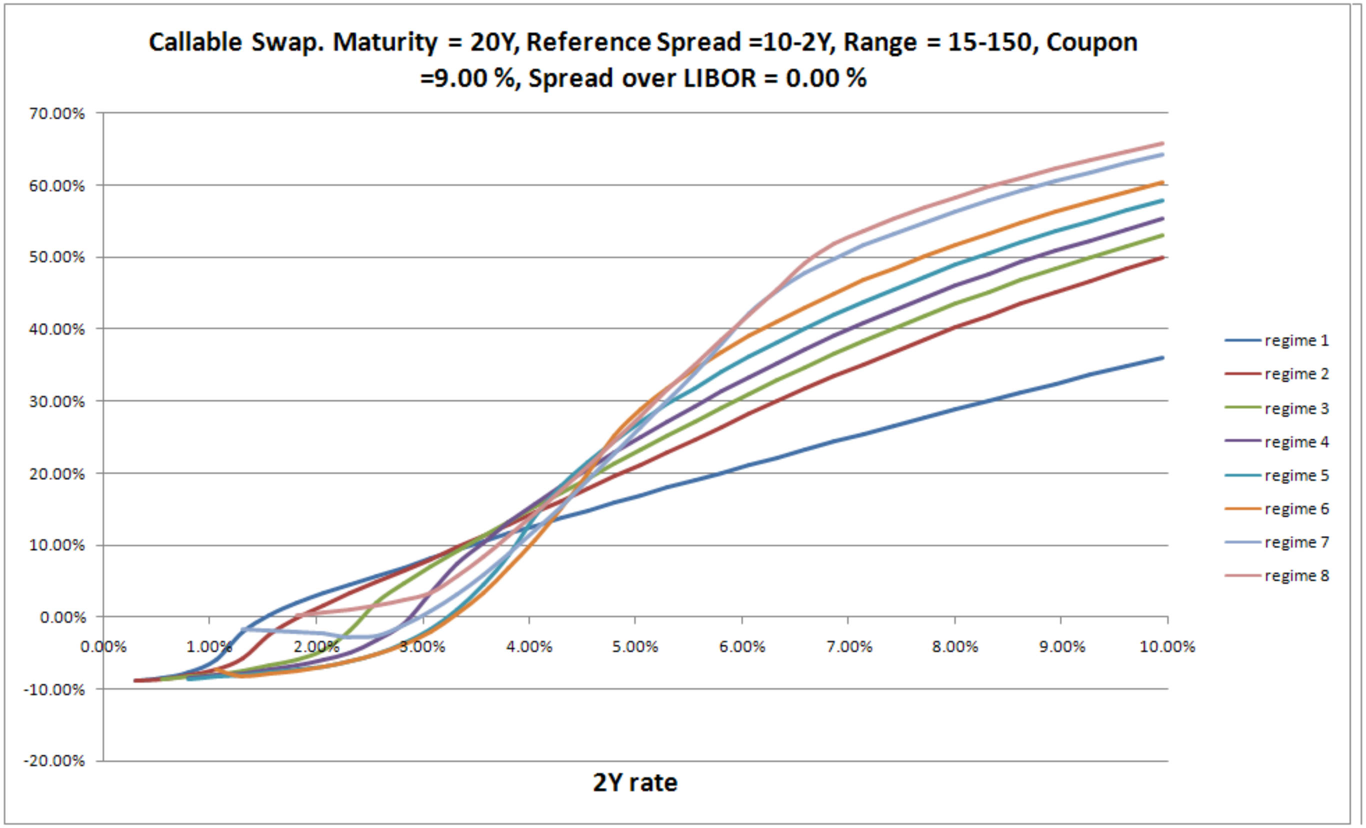
A CMS callable spread range accrual, example 2



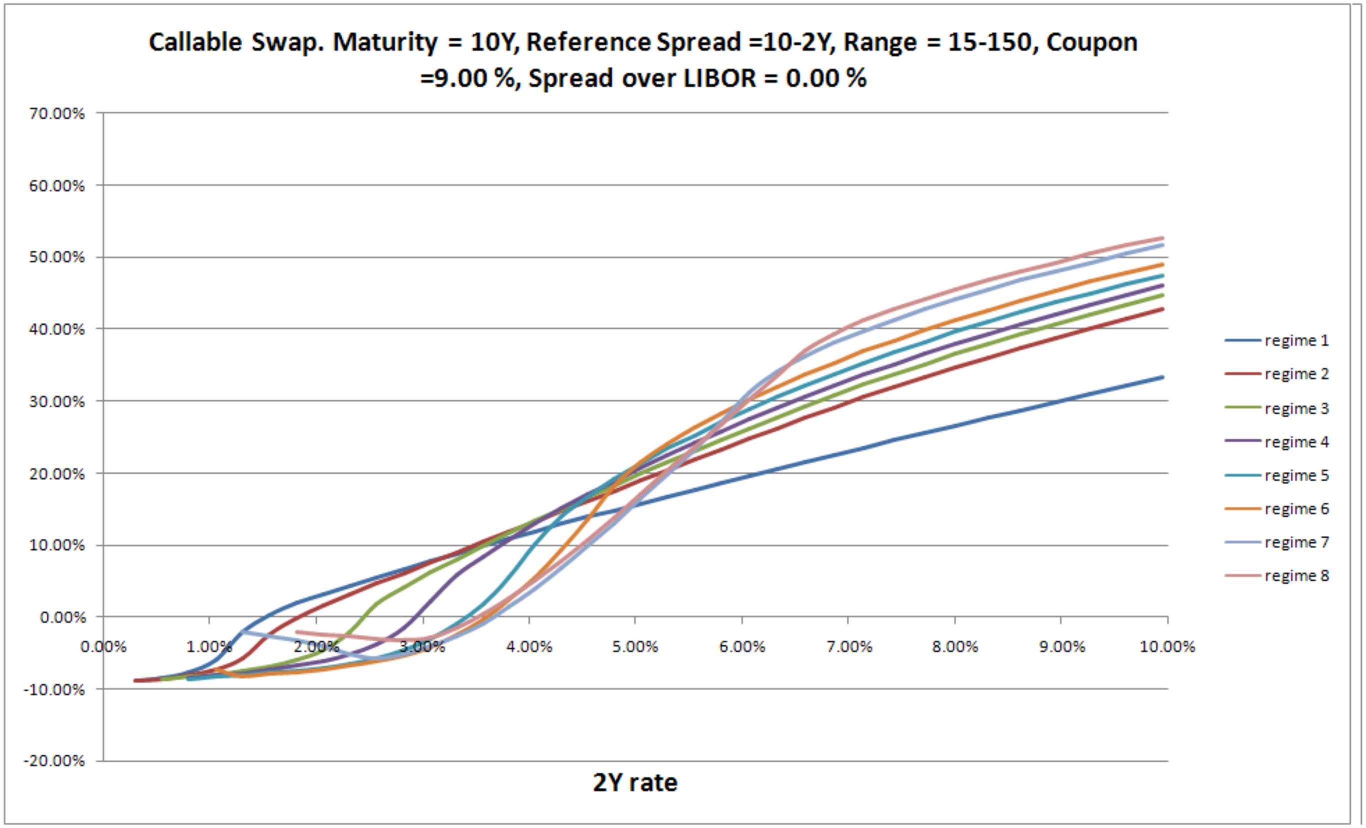
A CMS callable spread range accrual, example 3



A CMS callable spread range accrual, example 4



A CMS callable spread range accrual, example 5



Conclusions

Operator methods are an emerging mathematical framework for Constructive Probability and Mathematical Finance which is suitable for semi-parametric and non parametric modeling.

The practical engineering applications of operator methods rely on fast implementations of matrix-matrix multiplication algorithms, which can nowadays be achieved on massively parallel GPUs.

Theoretical work is still at the beginning and promises to lead to the discovery of new mathematical landscapes.