Convexity conundrums: Pricing cms swaps, caps and floors
1. Introduction. Here we focus on a single class of deals, the constant maturity swaps, caps, and floors. We develop a framework that leads to the standard methodology for pricing these deals, and then use this framework to systematically improve the pricing.

Let us start by agreeing on basic notation. In our notation, today is always \( t = 0 \). We use

\[
Z(t; T) = \text{value at date } t \text{ of a zero coupon bond with maturity } T,
\]

\[
D(T) \equiv Z(0, T) = \text{today's discount factor for maturity } T.
\]

We distinguish between zero coupon bonds and discount factors to remind ourselves that discount factors are not random, we can always obtain the current discount factors \( G(W) \) by stripping the yield curve, while zero coupon bonds \( Z(w, W) \) remain random until the present catches up to date \( w = 0 \).

1.1. Deal definition. Consider a CMS swap leg paying, say, the \( Q \) year swap rate plus a margin \( p \). Let \( w_0 > w_1 > \cdots > w_p \) be the dates of the CMS leg specified in the contract. (These dates are usually quarterly). For each period \( m \), the CMS leg pays

\[
\delta_j (R_j + m) \text{ paid at } w_m \quad \text{for } j = 1, 2, \ldots, m,
\]

where \( R_j \) is the \( Q \) year swap rate and

\[
\delta_j = \text{cvg}(t_{j-1}, t_j, \text{dcb})
\]

is the coverage of interval \( j \). If the CMS leg is set-in-advance (this is standard), then \( R_j \) is the rate for a standard swap that begins at \( t_{j-1} \) and ends \( N \) years later. This swap rate is fixed on the date \( \tau_j \) that is spot lag business days before the interval begins at \( t_{j-1} \), and pertains throughout the interval, with the accrued interest \( \delta_j (R_j + m) \) being paid on the interval’s end date, \( t_j \). Although set-in-advance is the market standard, it is not uncommon for contracts to specify CMS legs set-in-arrears. Then \( R_j \) is the \( N \) year swap rate for the swap that begins on the end date \( t_j \) of the interval, not the start date, and the fixing date \( \tau_j \) for \( R_j \) is spot lag business days before the interval ends at \( t_j \). As before, \( \delta_j \) is the coverage for the \( j \)th interval using the day count basis \( \text{dcb}_{\text{pay}} \) specified in the contract. Standard practice is to use the 30/360 basis for USD CMS legs.

CMS caps and floors are constructed in an almost identical fashion. For CMS caps and floors on the \( N \) year swap rate, the payments are

\[
\delta_j [R_j - K]^+ \text{ paid at } t_j \quad \text{for } j = 1, 2, \ldots, m, \quad \text{(cap)},
\]

\[
\delta_j [K - R_j]^+ \text{ paid at } t_j \quad \text{for } j = 1, 2, \ldots, m, \quad \text{(floor)},
\]

where the \( N \) year swap rate is set-in-advance or set-in-arrears, as specified in the contract.
1.2. Reference swap. The value of the CMS swap, cap, or floor is just the sum of the values of each payment. Any margin payments \( m \) can also be valued easily. So all we need do is value a single payment of the three types,

\[
\begin{align*}
(1.5a) & \quad R_s \quad & \text{paid at } t_p, \\
(1.5b) & \quad [R_s - K]^+ \quad & \text{paid at } t_p, \\
(1.5c) & \quad [K - R_s]^+ \quad & \text{paid at } t_p.
\end{align*}
\]

Here the reference rate \( R_s \) is the par rate for a standard swap that starts at date \( s_0 \), and ends \( N \) years later at \( s_n \). To express this rate mathematically, let \( s_1, s_2, \ldots, s_n \) be the swap’s (fixed leg) pay dates. Then a swap with rate \( R_{fix} \) has the fixed leg payments

\[
\begin{align*}
(1.6a) & \quad \alpha_j R_{fix} \quad & \text{paid at } s_j \quad & \text{for } j = 1, 2, \ldots, n,
\end{align*}
\]

where

\[
(1.6b) \quad \alpha_j = \text{cvg}(t_{j-1}, t_j, \text{dcb}_{sw})
\]

is the coverage (fraction of a year) for each period \( j \), and \( \text{dcb}_{sw} \) is the standard swap basis. In return for making these payments, the payer receives the floating leg payments. Neglecting any basis spread, the floating leg is worth 1 paid at the start date \( s_0 \), minus 1 paid at the end date \( s_n \). At any date \( t \), then, the value of the swap to the payer is

\[
(1.7) \quad V_{sw}(t) = Z(t; s_0) - Z(t; s_n) - R_{fix} \sum_{j=1}^{n} \alpha_j Z(t; s_j).
\]

The level of the swap (also called the annuity, \( PV01 \), \( DV01 \), or numerical duration) is defined as

\[
(1.8) \quad L(t) = \sum_{j=1}^{n} \alpha_j Z(t; s_j).
\]

Crudely speaking, the level \( L(t) \) represents the value at time \( t \) of receiving $1 per year (paid annually or semiannually, according to the swap’s frequency) for \( N \) years. With this definition, the value of the swap is

\[
(1.9a) \quad V_{sw}(t) = [R_s(t) - R_{fix}] L(t),
\]

where

\[
(1.9b) \quad R_s(t) = \frac{Z(t; s_0) - Z(t; s_n)}{L(t)}.
\]

Clearly the swap is worth zero when \( R_{fix} \) equals \( R_s(t) \), so \( R_s(t) \) is the par swap rate at date \( t \). In particular, today’s level

\[
(1.10a) \quad L_0 = L(0) = \sum_{j=1}^{n} \alpha_j D_j = \sum_{j=1}^{n} \alpha_j D(s_j),
\]

and today’s (forward) swap rate

\[
(1.10b) \quad R_0 \equiv R_s(0) = \frac{D_0 - D_n}{L_0}
\]

are both determined by today’s discount factors.
2. Valuation. According to the theory of arbitrage free pricing, we can choose any freely tradeable instrument as our \textit{numeraire}. Examining 1.8 shows that the level $O(w)$ is just the value of a collection zero coupon bonds, since the coverages $\alpha_j$ are just fixed numbers. These are clearly freely tradeable instruments, so we can choose the level $L(t)$ as our numeraire.\footnote{We follow the standard (if bad) practice of referring to both the physical instrument and its value as the “numeraire.”} The usual theorems then guarantee that there exists a probability measure such that the value $V(t)$ of any freely tradeable deal divided by the numeraire is a Martingale. So

$$V(t) = L(t) E \left\{ \frac{V(T)}{L(T)} \bigg| F_t \right\} \quad \text{for any } T > t,$$

provided there are no cash flows between $t$ and $T$.

It is helpful to examine the valuation of a plain vanilla swaption. Consider a standard European option on the reference swap. The exercise date of such an option is the swap’s fixing date $\tau$, which is spot-lag business days before the start date $s_0$. At this exercise date, the payoff is the value of the swap, provided this value is positive, so

$$V_{\text{opt}}(\tau) = [R_s(\tau) - R_{fix}]^+ L(\tau)$$

on date $\tau$. Since the Martingale formula \ref{eq:2.1} holds for any $T > t$, we can evaluate it at $T = \tau$, obtaining

$$V_{\text{opt}}(t) = L(t) E \left\{ \frac{V_{\text{opt}}(\tau)}{L(\tau)} \bigg| F_t \right\} = L(t) E \left\{ [R_s(\tau) - R_{fix}]^+ \bigg| F_t \right\}.$$

In particular, today’s value of the swaption is

$$V_{\text{opt}}(t) = L_0 E \left\{ [R_s(\tau) - R_{fix}]^+ \bigg| F_0 \right\}.$$

Moreover, 1.9b shows that the par swap rate $R_s(t)$ is the value of a freely tradable instrument (two zero coupon bonds) divided by our numeraire. So the swap rate must also a Martingale, and

$$E \{ R_s(\tau) \big| F_0 \} = R_s(0) \equiv R^0,$$

To complete the pricing, one now has to invoke a mathematical model (Black’s model, Heston’s model, the SABR model, . . . ) for how $R_s(\tau)$ is distributed around its mean value $R^0$. In Black’s model, for example, the swap rate is distributed according to

$$R_s(\tau) = R^0 e^{\sigma x \sqrt{\tau} - \frac{1}{2} \sigma^2 \tau},$$

where $x$ is a normal variable with mean zero and unit variance. In the “normal” or “absolute vol” model, the swap rate is distributed according to

$$R_s(\tau) = R^0 + ax \sqrt{\tau},$$

where again $x$ is a normal variable with mean zero and unit variance. One completes the pricing by integrating to calculate the expected value.
2.1. CMS caplets. The payoff of a CMS caplet is

\[ (2.6) \quad [R_s(\tau) - K]^+ \text{ paid at } t_p. \]

On the swap's fixing date \( \tau \), the par swap rate \( R_s \) is set and the payoff is known to be \([R_s(\tau) - K]^+ Z(\tau; t_p)\), since the payment is made on \( t_p \). Evaluating 2.1 at \( T = \tau \) yields

\[ (2.7a) \quad V_{cap}^{CMS}(t) = L(t) E \left\{ \frac{[R_s(\tau) - K]^+ Z(\tau; t_p)}{L(\tau)} \bigg| \mathcal{F}_t \right\}. \]

In particular, today's value is

\[ (2.7b) \quad V_{cap}^{CMS}(0) = L_0 E \left\{ \frac{[R_s(\tau) - K]^+ Z(\tau; t_p)}{L(\tau)} \bigg| \mathcal{F}_0 \right\}. \]

The ratio \( Z(\tau; t_p)/L(\tau) \) is (yet another!) Martingale, so it’s average value is today’s value:

\[ (2.8) \quad E \left\{ Z(\tau; t_p)/L(\tau) \big| \mathcal{F}_0 \right\} = D(t_p)/L_0. \]

By dividing \( Z(\tau; t_p)/L(\tau) \) by its mean, we obtain

\[ (2.9) \quad V_{cap}^{CMS}(0) = D(t_p) E \left\{ \left[ R_s(\tau) - K \right]^+ \frac{Z(\tau; t_p)/L(\tau)}{D(t_p)/L_0} \bigg| \mathcal{F}_0 \right\}, \]

which can be written more evocatively as

\[ (2.10) \quad V_{cap}^{CMS}(0) = D(t_p) E \left\{ \left[ R_s(\tau) - K \right]^+ \bigg| \mathcal{F}_0 \right\} \]

\[ + D(t_p) E \left\{ \left[ R_s(\tau) - K \right]^+ \left( \frac{Z(\tau; t_p)/L(\tau)}{D(t_p)/L_0} - 1 \right) \bigg| \mathcal{F}_0 \right\}. \]

The first term is exactly the price of a European swaption with notional \( D(t_p)/L_0 \), regardless of how the swap rate \( R_s(\tau) \) is modeled. The last term is the “convexity correction.” Since \( R_s(\tau) \) is a Martingale and \([Z(\tau; t_p)/L(\tau)]/[Z(t; t_p)/L(t)] - 1\) is zero on average, this term goes to zero linearly with the variance of the swap rate \( R_s(\tau) \), and is much, much smaller than the first term.

There are two steps in evaluating the convexity correction. The first step is to model the yield curve movements in a way that allows us to re-write the level \( L(\tau) \) and the zero coupon bond \( Z(\tau; t_p) \) in terms of the swap rate \( R_s \). (One obvious model is to allow only parallel shifts of the yield curve). Then we can write

\[ (2.11a) \quad Z(\tau; t_p)/L(\tau) = G(R_s(\tau)), \]
\[ (2.11b) \quad D(t_p)/L_0 = G(R_s^0), \]

for some function \( G(R_s) \). The convexity correction is then just the expected value

\[ (2.12) \quad cc = D(t_p) E \left\{ \left[ R_s(\tau) - K \right]^+ \left( \frac{G(R_s(\tau))}{G(R_s^0)} - 1 \right) \bigg| \mathcal{F}_0 \right\}. \]

over the swap rate \( R_s(\tau) \). The second step is to evaluate this expected value.

In the appendix we start with the street-standard model for expressing \( L(\tau) \) and \( Z(\tau; t_p) \) in terms of the swap rate \( R_s \). This model uses bond math to obtain

\[ (2.13a) \quad G(R_s) = \frac{R_s}{(1 + R_s/q)^2} - \frac{1}{(1 + R_s/q)^2}. \]
Here \( q \) is the number of periods per year (1 if the reference swap is annual, 2 if it is semi-annual, ...), and

\[
(2.13b) \quad \Delta = \frac{t_p - s_0}{s_1 - s_0}
\]

is the fraction of a period between the swap’s start date \( s_0 \) and the pay date \( t_p \). For deals “set-in-arrears” \( \Delta = 0 \). For deals “set-in-advance,” if the CMS leg dates \( t_0, t_1, \ldots \) are quarterly, then \( t_p \) is 3 months after the start date \( s_0 \), so \( \Delta = \frac{1}{4} \) if the swap is semiannual and \( \Delta = \frac{1}{12} \) if it is annual.

In the appendix we also consider sophisticated models for expressing \( L(\tau) \) and \( Z(\tau; t_p) \) in terms of the swap rate \( R_s \), and obtain increasingly sophisticated functions \( G(R_s) \).

We can carry out the second step by replicating the payoff in 2.12 in terms of payer swaptions. For any smooth function \( f(R_a) \) with \( f(N) = 0 \), we can write

\[
(2.14) \quad f'(K)[R_s - K]^+ + \int_K^\infty [R_s - x]^+ f''(x)dx = \left\{ \begin{array}{ll}
\frac{f(R_s)}{N} & \text{for } R_s > K \\
0 & \text{for } R_s < K
\end{array} \right.
\]

Choosing

\[
(2.15) \quad f(x) \equiv |x - K| \left( \frac{G(x)}{G(R_s)} - 1 \right),
\]

and substituting this into 2.12, we find that

\[
(2.16) \quad cc = D(t_p) \left\{ f'(K)E \{ |R_s(\tau) - K|^+ \mid F_0 \} + \int_K^\infty f''(x)E \{ |R_s(\tau) - x|^+ \mid F_0 \} dx \right\}.
\]

Together with the first term, this yields

\[
(2.17a) \quad V_{cap}^{CMS}(0) = \frac{D(t_p)}{L_0} \left\{ [1 + f'(K)]C(K) + \int_K^\infty C(x)f''(x)dx \right\},
\]

as the value of the CMS caplet, where

\[
(2.17b) \quad C(x) = L_0 E \{ |R_s(\tau) - x|^+ \mid F_0 \}
\]

is the value of an ordinary payer swaption with strike \( x \).

This formula replicates the value of the CMS caplet in terms of European swaptions at different strikes \( x \). At this point some pricing systems break the integral up into 10bp or so buckets, and re-write the convexity correction as the sum of European swaptions centered in each bucket. These swaptions are then consolidated with the other European swaptions in the vanilla book, and priced in the vanilla pricing system. This “replication method” is the most accurate method of evaluating CMS legs. It also has the advantage of automatically making the CMS pricing and hedging consistent with the desk’s handling of the rest of its vanilla book. In particular, it incorporates the desk’s smile/skew corrections into the CMS pricing. However, this method is opaque and compute intensive. After briefly considering CMS floorlets and CMS swaplets, we develop simpler approximate formulas for the convexity correction, as an alternative to the replication method.

2.2. CMS floorlets and swaplets. Repeating the above arguments shows that the value of a CMS floorlet is given by

\[
(2.18a) \quad V_{floor}^{CMS}(0) = \frac{D(t_p)}{L_0} \left\{ [1 + f'(K)]P(K) - \int_{-\infty}^K P(x)f''(x)dx \right\},
\]
where \( f(x) \) is the same function as before (see 2.15), and where

\[
P(x) = L_0 \mathbb{E} \left\{ [x - R_s(\tau)]^+ \mid \mathcal{F}_0 \right\}
\]

is the value of the ordinary receiver swaption with strike \( x \). Thus, the CMS floolets can also be priced through replication with vanilla receivers. Similarly, the value of a single CMS swap payment is

\[
V_{\text{CMS}} \left( \frac{D(t_p)}{L_0} R^0_s \right) + \frac{D(t_p)}{L_0} \left\{ \int_{R^0_s}^{\infty} C(x) f'_{\text{atm}}(x) dx + \int_{-\infty}^{R^0_s} P(x) f''_{\text{atm}}(x) dx \right\},
\]

where

\[
f_{\text{atm}}(x) \equiv |x - R^0_s| \left( \frac{G(x)}{G(R^0_s)} - 1 \right)
\]

is the same as \( f(x) \) with the strike \( K \) replaced by the par swap rate \( R^0_s \). Here, the first term in 2.19a is the value if the payment were exactly equal to the forward swap rate \( R^0_s \) as seen today. The other terms represent the convexity correction, written in terms of vanilla payer and receiver swaptions. These too can be evaluated by replication.

It should be noted that CMS caplets and floolets satisfy call-put parity. Since

\[
[R_s(\tau) - K]^+ - [K - R_s(\tau)]^+ = R_s(\tau) - K \quad \text{paid at } t_p,
\]

the payoff of a CMS caplet minus a CMS floolet is equal to the payoff of a CMS swaplet minus \( K \). Therefore, the value of this combination must be equal at all earlier times as well:

\[
V_{\text{cap}}^{\text{CMS}}(t) - V_{\text{floor}}^{\text{CMS}}(t) = V_{\text{swap}}^{\text{CMS}}(t) - KZ(t; t_p)
\]

In particular,

\[
V_{\text{cap}}^{\text{CMS}}(0) - V_{\text{floor}}^{\text{CMS}}(0) = V_{\text{swap}}^{\text{CMS}}(0) - KD(t_p).
\]

Accordingly, we can price an in-the-money caplet or floolet as a swaplet plus an out-of-the-money floolet or caplet.

3. **Analytical formulas.** The function \( G(x) \) is smooth and slowly varying, regardless of the model used to obtain it. Since the probable swap rates \( R_s(\tau) \) are heavily concentrated around \( R^0_s \), it makes sense to expand \( G(x) \) as

\[
G(x) \approx G(R^0_s) + G'(R^0_s)(x - R^0_s) + \cdots.
\]

For the moment, let us limit the expansion to the linear term. This makes \( f(x) \) a quadratic function,

\[
f(x) \approx \frac{G'(R^0_s)}{G(R^0_s)} (x - R^0_s)(x - K),
\]

and \( f''(x) \) a constant. Substituting this into our formula for a CMS caplet (2.17a), we obtain

\[
V_{\text{cap}}^{\text{CMS}}(0) = \frac{D(t_p)}{L_0} C(K) + \frac{G'(R^0_s)}{G(R^0_s)} \left\{ (K - R^0_s)C(K) + 2 \int_K^{\infty} C(x) dx \right\},
\]

where we have used \( G(R^0_s) = D(t_p)/L_0 \). Now, for any \( K \) the value of the payer swaption is

\[
C(K) = L_0 \mathbb{E} \left\{ [R_s(\tau) - K]^+ \mid \mathcal{F}_0 \right\},
\]

where \( \mathbb{E} \left\{ \right\} \) is the expectation conditioned on \( \mathcal{F}_0 \).
so the integral can be re-written as

\[ \int_{K}^{\infty} C(x) dx = L_0 \mathbb{E}\left\{ \int_{K}^{\infty} [R_s(\tau) - x]^+ dx \middle| \mathcal{F}_0 \right\} = \frac{1}{2} L_0 \mathbb{E}\left\{ ([R_s(\tau) - K]^+)^2 \middle| \mathcal{F}_0 \right\}. \]

Putting this together yields

\[ V_{CMS}^{cap}(0) = \frac{D(t_p)}{L_0} C(K) + G'(R_s^0) L_0 \mathbb{E}\left\{ [R_s(\tau) - R_s^0] [R_s(\tau) - K]^+ \middle| \mathcal{F}_0 \right\} \]

for the value of a CMS caplet, where the convexity correction is now the expected value of a quadratic “payoff.” An identical arguments yields the formula

\[ V_{CMS}^{floor}(0) = \frac{D(t_p)}{L_0} P(K) - G'(R_s^0) L_0 \mathbb{E}\left\{ [R_s^0, R_s(\tau)] [K - R_s(\tau)]^+ \middle| \mathcal{F}_0 \right\} \]

for the value of a CMS floorlet. Similarly, the value of a CMS swap payment works out to be

\[ V_{CMS}^{swap}(0) = D(t_p) R_s^0 + G'(R_s^0) L_0 \mathbb{E}\left\{ (R_s(\tau) - R_s^0)^2 \middle| \mathcal{F}_0 \right\}. \]

To finish the calculation, one needs an explicit model for the swap rate \( R_s(\tau) \). There are two simple models one can use. The first is Black’s model, which assumes the swap rate \( R_s(\tau) \) is log normal with a volatility \( \sigma \),

\[ dR_s = \sigma R_s dW; \]

With this model, one obtains

\[ V_{CMS}^{swap}(0) = D(t_p) R_s^0 + G'(R_s^0) L_0 \left( R_s^0 \right)^2 \left[ e^{\sigma^2 \tau} - 1 \right] \]

for the CMS swaplets,

\[ V_{CMS}^{cap}(0) = \frac{D(t_p)}{L_0} C(K) + G'(R_s^0) L_0 \left[ (R_s^0)^2 e^{\sigma^2 \tau} N(d_{3/2}) - R_s^0 (R_s^0 + K) N(d_{1/2}) + R_s^0 K N(d_{-1/2}) \right] \]

for CMS caplets, and

\[ V_{CMS}^{floor}(0) = \frac{D(t_p)}{L_0} P(K) - G'(R_s^0) L_0 \left[ (R_s^0)^2 e^{\sigma^2 \tau} N(-d_{3/2}) - R_s^0 (R_s^0 + K) N(-d_{1/2}) + R_s^0 K N(-d_{-1/2}) \right] \]

for CMS floorlets. Here

\[ d_\lambda = \frac{\ln R_s^0 / K + \lambda \sigma^2 \tau}{\sigma \sqrt{\tau}}. \]

The second model is the normal, or absolute model, which assumes that the swap rate follows

\[ dR_s = adW, \]
where \( a \) is the "absolute" or "normal" vol. This yields

\[
V_{\text{swap}}^{\text{CMS}}(0) = D(t_p) R_s^0 + G'(R_s^0) L_0 a^2 \tau
\]

for the CMS swaplets,

\[
V_{\text{cap}}^{\text{CMS}}(0) = \frac{D(t_p)}{L_0} C(K) + G'(R_s^0) L_0 a^2 \tau N \left( \frac{R_s^0 - K}{a \sqrt{\tau}} \right)
\]

for CMS caplets, and

\[
V_{\text{floor}}^{\text{CMS}}(0) = \frac{D(t_p)}{L_0} P(K) - G'(R_s^0) L_0 a^2 \tau N \left( \frac{K - R_s^0}{a \sqrt{\tau}} \right)
\]

for CMS floorlets. We can obtain the normal vol \( a \) by noting that if \( \sigma \) is the log normal volatility for a swaption with forward rate \( R^0 \) and strike \( K \), then the normal volatility \( a \) of this swaption is

\[
a = \sigma \frac{R_s^0 - K}{\log R_s^0/K} \left[ 1 + \frac{1}{24} \sigma^2 \tau + \frac{1}{5760} \sigma^4 \tau^2 \right]
\]

Near the money \((R_s^0 - K)/K < 20\%\) or so, we can replace this formula with

\[
a = \sigma \sqrt{R_s^0/K} \left[ 1 + \frac{1}{24} \log^2 R_s^0/K + \frac{1}{120} \log^4 R_s^0/K \right] \left[ 1 + \frac{1}{24} \sigma^2 \tau + \frac{1}{5760} \sigma^4 \tau^2 \right]
\]

The key concern with Black’s model is that it does not address the smiles and/or skews seen in the marketplace. This can be partially mitigated by using the correct volatilities. For CMS swaps, the volatility \( \sigma_{\text{ATM}} \) for at-the-money swaptions should be used, since the expected value 3.4c includes high and low strike swaptions equally. For out-of-the-money caplets and floorlets, the volatility \( \sigma_K \) for strike \( K \) should be used, since the swap rates \( R_s(\tau) \) near \( K \) provide the largest contribution to the expected value. For in-the-money options, the largest contributions come from swap rates \( R_s(\tau) \) near the mean value \( R_s^0 \). Accordingly, call-put parity should be used to evaluate in-the-money caplets and floorlets as a CMS swap payment plus an out-of-the-money floorlet or caplet.

4. Conclusions. The standard pricing for CMS legs is given by 3.5b - 3.5e with \( G(R_s) \) given by 2.13a. These formula are adequate for many purposes. When finer pricing is required, one can systematically improve these formulas by using the more sophisticated models for \( G(R_s) \) developed in the Appendix, and by adding the quadratic and higher order terms in the expansion 3.1a. In addition, 3.4a - 3.4b show that the convexity corrections are essentially swaptions with “quadratic” payoffs. These payoffs emphasize away-from-the-money rates more than standard swaptions, so the convexity corrections can be quite sensitive to the market’s skew and smile. CMS pricing can be improved by replacing Black’s model with a model that matches the market smile, such as Heston’s model or the SABR model. Alternatively, when the very highest accuracy is needed, replication can be used to obtain near perfect results.

Appendix A. Models of the yield curve.

A.1. Model 1: Standard model. The standard method for computing convexity corrections uses bond math approximations: payments are discounted at a flat rate, and the coverage (day count fraction) for each period is assumed to be \( 1/q \), where \( q \) is the number of periods per year (1 for annual, 2 for semi-annual, etc). At any date \( t \), the level is approximated as

\[
L(t) = Z(t, s_0) \sum_{j=1}^{n} \alpha_j Z(t, s_j) \approx Z(t, s_0) \sum_{j=1}^{n} \frac{1/q}{1 + R_s(t)/q}.
\]
which works out to

\[ L(t) = \frac{Z(t, s_0)}{R_s(t)} \left[ 1 - \frac{1}{(1 + R_s(t)/q)\Delta} \right]. \]

Here the par swap rate \( R_s(t) \) is used as the discount rate, since it represents the average rate over the life of the reference swap. In a similar spirit, the zero coupon bond for the pay date \( t_p \) is approximated as

\[ Z(t; t_p) \approx \frac{Z(t, s_0)}{(1 + R_s(t)/q)\Delta}, \]

where

\[ \Delta = \frac{t_p - s_0}{s_1 - s_0} \]

is the fraction of a period between the swap's start date \( s_0 \) and the pay date \( t_p \). Thus the standard "bond math model" leads to

\[ G(R_s) = \frac{Z(t; t_p)}{L(t)} \approx \frac{R_s}{(1 + R_s/q)\Delta} \left[ 1 - \frac{1}{(1 + R_s/q)\Delta} \right]. \]

This method a) approximates the schedule and coverages for the reference swaption; b) assumes that the initial and final yield curves are flat, at least over the tenor of the reference swaption; and c) assumes a correlation of 100% between rates of differing maturities.

A.2. Model 2: “Exact yield” model. We can account for the reference swaption’s schedule and day count exactly by approximating

\[ Z(t; s_j) \approx Z(t; s_0) \prod_{k=1}^{j} \frac{1}{1 + \alpha_k R_s(t)}. \]

where \( \alpha_k \) is the coverage of the \( k^{th} \) period of the reference swaption. At any date \( t \), the level is then

\[ L(t) = \sum_{j=1}^{n} \alpha_j Z(t; s_j) = Z(t; s_0) \sum_{j=1}^{n} \alpha_j \left( \prod_{k=1}^{j} \frac{1}{1 + \alpha_k R_s(t)} \right). \]

We can establish the following identity by induction:

\[ L(t) = \frac{Z(t; s_0)}{R_s(t)} \left( 1 - \prod_{k=1}^{n} \frac{1}{1 + \alpha_k R_s(t)} \right). \]

In the same spirit, we can approximate

\[ Z(t; t_p) = Z(t; s_0) \frac{1}{(1 + \alpha_1 R_s(t))\Delta}, \]

where \( \Delta = (t_p - s_0)/(s_1 - s_0) \) as before. Then

\[ G(R_s) = \frac{Z(t; t_p)}{L(t)} \approx \frac{R_s}{(1 + \alpha_1 R_s)\Delta} \left[ 1 - \prod_{k=1}^{n} \frac{1}{(1 + \alpha_k R_s)} \right]. \]

This approximates the yield curve as flat and only allows parallel shifts, but has the schedule right.
A.3. Model 3: Parallel shifts. This model takes into account the initial yield curve shape, which can be significant in steep yield curve environments. We still only allow parallel yield curve shifts, so we are approximate

\[(A.9) \quad \frac{Z(t; s_j)}{Z(t; s_0)} \approx \frac{D(s_j)}{D(s_0)} e^{-(s_j-s_0)x} \quad \text{for } j = 1, 2, \ldots, n\]

where \(x\) is the amount of the parallel shift. The level and swap rate \(R_s\) are given by

\[(A.10a) \quad \frac{L(t)}{Z(t; s_0)} = \sum_{j=1}^{n} \alpha_j \frac{D(s_j)}{D(s_0)} e^{-(s_j-s_0)x}\]

\[(A.10b) \quad R_s(t) = \frac{D(s_0) - D(s_n)e^{-(s_n-s_0)x}}{\sum_{j=1}^{n} \alpha_j D(s_j)e^{-(s_j-s_0)x}}\]

Turning this around,

\[(A.11a) \quad R_s \sum_{j=1}^{n} \alpha_j D(s_j)e^{-(s_j-s_0)x} + D(s_n)e^{-(s_n-s_0)x} = D(s_0)\]

determines the parallel shift \(x\) implicitly in terms of the swap rate \(R_s\). With \(x\) determined by \(R_s\), the level is given by

\[(A.11b) \quad \frac{L(R_s)}{Z(t; s_0)} = \frac{D(s_0) - D(s_n)e^{-(s_n-s_0)x}}{D(s_0)R_s}\]

in terms of the swap rate. Thus this model yields

\[(A.12a) \quad G(R_s) = \frac{Z(t; t_p)}{L(t)} = \frac{R_s e^{-(t_p-s_0)x}}{1 - \frac{D(s_n)e^{-(s_n-s_0)x}}{D(s_0)}}\]

where \(x\) is determined implicitly in terms of \(R_s\) by

\[(A.12b) \quad R_s \sum_{j=1}^{n} \alpha_j D(s_j)e^{-(s_j-s_0)x} + D(s_n)e^{-(s_n-s_0)x} = D(s_0)\].

This model’s limitations are that it allows only parallel shifts of the yield curve and it presumes perfect correlation between long and short term rates.

A.4. Model 4: Non-parallel shifts. We can allow non-parallel shifts by approximating

\[(A.13) \quad \frac{Z(t; s_j)}{Z(t; s_0)} \approx \frac{D(s_j)}{D(s_0)} e^{-[h(s_j)-h(s_0)]x},\]

where \(x\) is the amount of the shift, and \(h(s)\) is the effect of the shift on maturity \(s\). As above, the shift \(x\) is determined implicitly in terms of the swap rate \(R_s\) via

\[(A.14a) \quad R_s \sum_{j=1}^{n} \alpha_j D(s_j)e^{-[h(s_j)-h(s_0)]x} + D(s_n)e^{-[h(s_n)-h(s_0)]x} = D(s_0).\]
Then

\[
L(R_s) = \frac{Z(t; s_0)}{Z(t; s_0)} = \frac{D(s_0) - D(s_n)e^{-[h(s_n)-h(s_0)]x}}{D(s_0)R_s}
\]
determines the level in terms of the swap rate. This model then yields

\[
G(R_s) = \frac{Z(t; t_p)}{L(t)} \approx \frac{R_se^{-[h(t_p)-h(s_0)]x}}{1 - \frac{D(s_n)}{D(s_0)}e^{-[h(s_n)-h(s_0)]x}},
\]
where \(x\) is determined implicitly in terms of \(R_s\) by

\[
R_s \sum_{j=1}^{n} \alpha_j D(s_j)e^{-[h(s_j)-h(s_0)]x} + D(s_n)e^{-[h(s_n)-h(s_0)]x} = D(s_0).
\]

To continue further requires selecting the function \(h(s_j)\) which determines the shape of the non-parallel shift. This is often done by postulating a constant mean reversion,

\[
h(s) - h(s_0) = \frac{1}{\kappa} \left[ 1 - e^{-\kappa(s-s_0)} \right].
\]
Alternatively, one can choose \(h(s_j)\) by calibrating the vanilla swaptions which have the same start date \(s_0\) and varying end dates to their market prices.