Exact Fit to the Swaption Volatility Matrix Using Semidefinite Programming*

Alan Brace† and Robert S. Womersley‡

Working paper presented at
ICBI Global Derivatives Conference
Paris, April 2000

Abstract

This paper addresses the problem of parameterizing jointly to all swaptions in the framework of the lognormal Libor market model (LLMM). First we refine some previous work showing that swaprates are nearly lognormal with a deterministic volatility. Then we formulate the fitting problem in terms of covariance matrices rather than the usual two variable volatility function. Semidefinite programming, in which the variables are positive semidefinite matrices (covariance matrices), minimizes a linear function of the variables subject to linear constraints on the variables. Exact fits to the swaption volatility matrix are imposed through linear constraints on covariance matrices. Various objectives can be used to get close to a historical covariance matrix. The method is illustrated by pricing fixed maturity Bermudan options.

Keywords: lognormal Libor market model, swaption volatility matrix, covariance matrix, semidefinite programming, Bermudan option.

---

*The authors would like to thank National Australia Bank for its support and Paul O’Brien for his help and comments
†National Australia Bank <abrace@ozemail.com.au>
‡School of Mathematics, University of New South Wales <r.womersley@unsw.edu.au>, http://www.maths.unsw.edu.au/~rsw
1 Introduction

Most professionals now regard the so called market models (see [17] for a review) as the best starting point for pricing caps, swaptions and associated exotics. These models are arbitrage free, lognormal, return Black cap or swaption formulae, have two variable (time/maturity) volatility functions, and input today’s yield curve as an initial value. The lognormal Libor market model (LLMM) parameterizes accurately to the cap strip and returns caplet prices that agree with Black values, while the lognormal swaprate model (LSRM) parameterizes accurately to some subsets of the swaption volatility matrix (like the reverse diagonal), and returns some Black swaption values. Outstanding problems that have been considered but not settled are: parameterizing jointly to all swaptions; including caps in the swaption parameterization; and fitting cap and swaption volatility skews [3]. We address the first.

It is important to realize that if all swaptions are to be fitted without compromising the model, then it is prices rather than volatilities that must be fitted, which leads to a highly non-linear optimization problem (the first author has discouraging memories of genetic algorithm routines that took hours and never gave similar answers twice). Clearly, fitting volatilities is simpler, because it cuts out the complexities of the Black
formula. But then, because only a subset of swaption volatilities can be returned without compromising the model, approximation becomes the necessary price of tractability. We choose to make such approximations within the LLMM because it is relatively easy (analytically and computationally), but in principle other market models could be used.

Thus we aim to exactly fit in the LLMM a matrix of approximate swaption volatilities. Our approach involves three main steps:

1. In Section 2 we refine some of the arguments in [6] and [7] which showed that swaprates in the LLMM are virtually lognormal with nearly deterministic volatility. The necessary approximations work because certain low variance martingales are set equal to their start values. We checked these approximations by confirming that our parameterized models returned, via simulation, those prices that are input to the parameterization.

2. Secondly (see Sections 2 and 4) we decided to work with covariance matrices rather than the volatility functions from which they are formed, as that removed another layer of complexity. Except for the case of piecewise constant volatility (which is what we use here), reversing this step to obtain continuous volatility functions is not trivial.

3. Thirdly (see Sections 3 and 4) we realized that given the right volatility structure and optimization objective, the problem was solvable using the recently developed techniques of semidefinite programming (analogous to linear programming except the variables are semidefinite matrices rather than non-negative vectors).

Considerable experimentation led us to volatility functions of the form (with standard notation)

\[ \gamma(t, T) = \phi(t)\xi(T - t), \]

in which \( \phi(\cdot) \) is piecewise constant. These volatility functions can be thought of as being “homogeneous in layers” determined by time \( t \). Swaption volatilities then turn out to be simple linear combinations of covariance matrices, leading to linear equality constraints in the semidefinite framework. Moreover, because \( \xi(\cdot) \) is homogeneous, instantaneous correlation can be sensibly defined, so our objective function can be made to target historic covariance in some way. The resulting parameterizations fit all swaption volatilities with a correlation structure not too different from the historic one. That correlation structure can be thought of as implied correlation.

Of course we parameterized to European swaptions and got back, via simulation, the values input to the parameterization. In addition we decided to test our technique on fixed maturity Bermudans because they tend to be parameterized on a deal-by-deal basis. Specifically, we were interested in whether, with just one parameterization and over a range of data, our method could return similar prices to the standard models for

---

1Our thanks to Marek Musiela for this suggestion.
Bermudans using two significantly different reverse diagonals. The results (see Section 5) surprised us: our Bermudan prices were either roughly equal to, or less than, prices from the standard single factor models.

We are uncertain of what to make of our finding, but take comfort from the fact that Bermudans have been priced with single factor models for a long time, with only a couple (as far as we know) of rumoured instances of significant writeoffs. The point is that, because banks tend to be overall buyers of Bermudans, such writeoffs would likely arise from models that overprice rather than underprice.

2 Swaprate Volatilities in the LLMM

The background to this section is [6], in which Brace et al showed LLMM swaprates were statistically lognormal and derived approximate formulae for their volatilities, and [7] in which Barton et al investigated hedging implications. We show the relative accuracy of these approximations stems from certain low variance martingales being set equal to their start values, and in Theorem 2.4 slightly tighten their approximation ([6], page 9 equation (23)) for swaprate volatility in the LLMM.

For arbitrary time $T$ let $W_T$ and $E_T$ denote Brownian motion and expectation respectively under the forward measure $\mathbb{P}_T$. In particular let $W_0$ and $E_0$ denote Brownian motion and expectation under the arbitrage free measure. Also let $P(t,T)$ denote the time $t$-price of a zero coupon bond maturing at an arbitrary time $T$, and set $T_j = T + j\delta$ for $j = 1, 2, \ldots, n$. Then in the LLMM simple forwards

$$K(t,T) = \frac{1}{\delta} \left\{ \frac{P(t,T)}{P(t,T_1)} - 1 \right\}$$

over arbitrary $\delta$-intervals $[T, T + \delta]$ are lognormal with deterministic volatility $\gamma(t,T)$ and

$$dK(t,T) = K(t,T)\gamma(t,T)dW_{T_1}(t)$$

under forward measures $\mathbb{P}_{T_1}$ at the ends of their intervals. Moreover, see [8], the corresponding stochastic HJM volatility $\sigma(t,T)$ satisfies

$$\int_T^{T_1} \sigma(t,u)du = \frac{\delta K(t,T)}{1 + \delta K(t,T)} \gamma(t,T) = \mu(t,T)\gamma(t,T), \quad (2.1)$$

where

$$\mu(t,T) = \frac{\delta K(t,T)}{1 + \delta K(t,T)}. \quad (2.2)$$

In addition it is easy to show that forward contracts

$$F_T(t,T_j) = \frac{P(t,T_j)}{P(t,T)},$$
expiring at \( T \) and maturing at \( T_j \), satisfy stochastic differential equations SDEs like

\[
dF_T(t, T_j) = -F_T(t, T_j) \int_T^{T_j} \sigma(t, u) du W_T(t),
\]

(2.3)

\[
\frac{d \left[ F_T(t, T_j) K(t, T_j-1) \right]}{F_T(t, T_j) K(t, T_j-1)} = \left\{ \gamma(t, T_j-1) - \int_T^{T_j} \sigma(t, u) du \right\} dW_T(t).
\]

Let \( K(t, T, r) \) be the forward over the interval \([T, T + \delta r]\) defined by

\[
F_T(t, T + \delta r) = P(t, T + \delta r) = \frac{1}{1 + r \delta K(t, T, r)}.
\]

The time \( t \) value of an \( r \)-period payer swap in which the floating forward \( K(t, T_j-r, r) \) over the interval \([T_j-r, T_j]\) is swapped in arrears against fixed \( \kappa \) at \( n/r \) intervals of \( r \delta \), can be expressed as the weighted sum of the simple forwards \( K(t, T_j) \) like

\[
P_{\text{swap}}(t, T, r, n) = E_0 \left\{ \sum_{j=1}^{n} \text{roz}(j, r) \frac{\beta(t)}{\beta(T_j)} \delta [K(t, T_j-r, r) - \kappa] \right| \mathcal{F}_t, \right.
\]

where \( \beta(\cdot) \) is the HJM bank account and \( \text{roz}(\cdot) \) is the \( r \) or zero function

\[
\text{roz}(j, r) = \begin{cases} r & \text{if } j \mod r = 0, \\ 0 & \text{otherwise}. \end{cases}
\]

Introducing the corresponding \( r \)-period swap rate

\[
\omega(t, T, r, n) = \frac{\sum_{j=1}^{n} F_T(t, T_j) K(t, T_j-1)}{\sum_{j=1}^{n} \text{roz}(j, r) F_T(t, T_j)},
\]

(2.4)

this payer swap can be rewritten

\[
P_{\text{swap}}(t, T, r, n) = \delta \sum_{j=1}^{n} \text{roz}(j, r) P(t, T_j) [\omega(t, T, r, n) - \kappa].
\]

A payer swaption exchanges in arrears at \( n/r \) consecutive intervals of \( r \delta \) the time \( T \) swaprate \( \omega(T, T, r, n) \) against the strike \( \kappa \) when \( \omega(t, T, r, n) \geq \kappa \), and so has value

\[
P_{\text{swpn}}(t, T, r, n) = E_0 \left\{ \sum_{j=1}^{n} \text{roz}(j, r) \frac{\beta(t)}{\beta(T_j)} \delta [\omega(t, T, r, n) - \kappa]^+ \right| \mathcal{F}_t, \right.
\]

where \( \mathcal{F}_t \) is the filtration.

We will tackle this expectation by showing that under a new measure \( \mathbb{P}^{(r)}_T \) equivalent to \( \mathbb{P}_T \), the \( r \)-period swap rate \( \omega(t, T, r, n) \) is close to lognormal.
2.1 Some low variance and other martingales

The first set of (low variance) martingales are the $\mu(t, T)$ defined in (2.1). From (2.3), it is easy to show
\[ \mu(t, T) = 1 - \frac{1}{1 + \delta K(t, T)} = 1 - F_T(t, T_1) \in [0, 1], \]
\[ d\mu(t, T) = \left[1 - \mu(t, T)\right] \mu(t, T) \gamma(t, T) dW_T(t) \]
\[ = \frac{\delta K(t, T)}{[1 + \delta K(t, T)]^2} \gamma(t, T) dW_T(t), \]
demonstrating that $\mu(t, T)$ is a $\mathbb{P}_T$-martingale. The $[1 - \mu(t, T)]$ term in the SDE prevents us viewing $\mu(t, T)$ as an exponential martingale with a stochastic volatility we might find to be small. But in absolute terms, not only is $\mu(t, T)$ bounded in $[0, 1]$, but its absolute volatility is a couple of orders of magnitude ($\delta K(t, T) \approx 1\% - 2\%$) less than the volatilities of the forwards. For this reason, we will approximate $\mu(t, T)$ with its zero value
\[ \mu(t, T) \approx \mu(0, T) = \frac{\delta K(0, T)}{1 + \delta K(0, T)} \tag{2.6} \]
for all maturities $T$.

With the aid of the SDEs (2.3), the following result (which is an easy exercise) helps us analyze the stochastic behaviour of the swaprate (2.4).

**Theorem 2.1** Under some measure $\mathbb{P}$, let $Y(t)$ be an Ito process with SDE
\[ \frac{dY(t)}{Y(t)} = \eta(t) dt + \xi(t) dW(t), \]
and let $X_j, j = 1, \ldots, n$ be exponential martingales with SDEs
\[ \frac{dX_j(t)}{X_j(t)} = \sigma_j(t) dW(t), \]
where the $\eta(t)$, $\xi(t)$ and $\sigma_j(t)$ may be stochastic. Then the variable
\[ Z(t) = \frac{Y(t)}{\sum_{j=1}^{n} \text{roz}(j, r) X_j(t)}, \]
$(1 \leq j \leq n, n \mod r = 0)$ satisfies the SDE
\[ \frac{dZ(t)}{Z(t)} = \eta(t) dt + \left[ \xi(t) - \sum_{j=1}^{n} \text{roz}(j, r) X_j(t) \sigma_j(t) \right] d\tilde{W}^{(r)}(t); \]
\[ d\tilde{W}^{(r)}(t) = dW(t) - \frac{\sum_{j=1}^{n} \text{roz}(j, r) X_j(t) \sigma_j(t)}{\sum_{j=1}^{n} \text{roz}(j, r) X_j(t)} dt. \]
The second set of (low variance) martingales are the weights $u_j(t)$ of the $K(t, T_{j-1})$ in the swaprate formula (2.4), namely
\begin{equation}
  u_j(t) = \frac{F_T(t, T_j)}{\sum_{j=1}^n \text{roz}(j, r)F_T(t, T_j)} \quad j = 1, \ldots, n.
\end{equation}

If $X_j(t) = F_T(t, T_j)$ and $Y(t) = F_T(t, T_j)$ in Theorem 2.1, then using (2.3)
\begin{equation}
  \frac{du_j(t)}{u_j(t)} = \left[ -\int_T^{T_j} \sigma(t, u)du + \sum_{j=1}^n \text{roz}(j, r)u_j(t) \int_T^{T_j} \sigma(t, u)du \right] d\tilde{W}_T^{(r)}(t),
\end{equation}
\begin{equation}
  d\tilde{W}_T^{(r)}(t) = dW_T(t) + \sum_{j=1}^n \text{roz}(j, r)u_j(t) \int_T^{T_j} \sigma(t, u)du dt.
\end{equation}

The weights $u_j(t)$ will all be martingales under the measure $\tilde{P}_T^{(r)}$ induced by $\tilde{W}_T^{(r)}(t)$, and because the plus and minus terms in (2.8) are roughly equal, their volatilities will tend to be small, making their initial values $u_j(0)$ good approximations to subsequent values.

The third set of (not low variance) martingales are the actual terms in the swaprate (2.4) as a whole, namely
\begin{equation}
  v_j(t) = \frac{F_T(t, T_j)K(t, T_{j-1})}{\sum_{j=1}^n \text{roz}(j, r)F_T(t, T_j)} \quad j = 1, \ldots, n.
\end{equation}

If $X_j(t) = F_T(t, T_j)$ and $Y(t) = F_T(t, T_j)K(t, T_{j-1})$ in Theorem 2.1, then using (2.3)
\begin{equation}
  \frac{dv_j(t)}{v_j(t)} = \left[ \gamma(t, T_{j-1}) - \int_T^{T_j} \sigma(t, u)du + \sum_{j=1}^n \text{roz}(j, r)u_j(t) \int_T^{T_j} \sigma(t, u)du \right] d\tilde{W}_T^{(r)}(t);
\end{equation}
\begin{equation}
  d\tilde{W}_T^{(r)}(t) = dW_T(t) + \sum_{j=1}^n \text{roz}(j, r)u_j(t) \int_T^{T_j} \sigma(t, u)du dt.
\end{equation}

Hence the $v_j(t)$ will also be martingales under the measure $\tilde{P}_T^{(r)}$ induced by $\tilde{W}_T^{(r)}(t)$. Note, however, that the volatility of $v_j(t)$ will be approximately $\gamma(t, T_{j-1})$, which is not small.

Because it is convenient to cover them here, we also identify two other types of martingales which will appear later.
A fourth set of (low variance) martingales are the weights in the expression (2.19) below for swaprate volatility, namely
\[
w_j(t) = \frac{F_T(t, T_j)K(t, T_j-1)}{\sum_{j=1}^{n} F_T(t, T_j)K(t, T_j-1)} \quad j = 1, \ldots, n.
\] (2.11)

If \(X_j(t) = F_T(t, T_j)K(t, T_j-1)\) and \(Y(t) = F_T(t, T_j)K(t, T_j-1)\) in Theorem 2.1, then from (2.3)
\[
\frac{dw_j(t)}{w_j(t)} = \left[ \gamma(t, T_j-1) - \int_{T}^{T_j} \sigma(t, u)du \right. \\
- \sum_{j=1}^{n} w_j(t) \left\{ \gamma(t, T_j-1) - \int_{T}^{T_j} \sigma(t, u)du \right\} \left] dW_T^r(t); \quad (2.12)
\]
\[
d\tilde{W}_T^r(t) = dW_T(t) - \sum_{j=1}^{n} w_j(t) \left\{ \gamma(t, T_j-1) - \int_{T}^{T_j} \sigma(t, u)du \right\} dt.
\]

The weights \(w_j(t)\) will all be martingales under the measure \(\tilde{P}_T^r\) induced by \(\tilde{W}_T^r(t)\) and, again because the plus and minus terms in (2.12) are roughly equal, their volatilities will tend to be small, again making their initial values \(w_j(0)\) good approximations to subsequent values.

A fifth and final type of martingale helps change measure between \(P_T\) and \(\tilde{P}_T^r\), see Theorem 2.3 below. If \(X_j(t) = F_T(t, T_j)\) and \(Y(t) = 1\) in Theorem 2.1, then using (2.3)
\[
\frac{1}{\sum_{j=1}^{n} \text{roz}(j, r)F_T(t, T_j)} = \left[ \sum_{j=1}^{n} \text{roz}(j, r)u_j(t) \int_{T}^{T_j} \sigma(t, u)du \right] d\tilde{W}_T^r(t).
\] (2.13)

showing the reciprocal of \(\sum_{j=1}^{n} \text{roz}(j, r)F_T(t, T_j)\) is yet another martingale under the measure \(\tilde{P}_T^r\) induced by \(\tilde{W}_T^r(t)\).

### 2.2 Swaprate Analysis

In the LLMM, define the swaprate measure \(\tilde{P}_T^r\), to be the measure equivalent to \(P_T\) induced by the Brownian motion \(\tilde{W}_T^r(t)\) appearing in formulae (2.8) and (2.10) above,
\[
d\tilde{W}_T^r(t) = dW_T(t) + \sum_{j=1}^{n} \text{roz}(j, r)u_j(t) \int_{T}^{T_j} \sigma(t, u)du dt.
\] (2.14)

Note that the \(\tilde{W}_T^r(t)\) can be expressed as the \(u_j(t)\)-weighted sum of the forward Brownian motions \(W_{T_j}(t)\) under the forward measures \(P_{T_j}\),
\[
d\tilde{W}_T^r(t) = \sum_{j=1}^{n} \text{roz}(j, r)u_j(t)dW_{T_j}(t),
\]
which means, intuitively, that the swaprate Brownian motion $\tilde{W}^{(r)}_T(t)$ is not too different from any of the forward Brownian motions $W_{T_j}(t)$.

Using (2.4) and (2.10), after some rearrangement, an SDE for the forward swaprate $\omega(t, T, r, n)$, is

$$
\frac{d\omega(t, T, r, n)}{\omega(t, T, r, n)} = \sum_{j=1}^n \left\{ w_j(t)\gamma(t, T_{j-1}) + [roz(j, r)u_j(t) - w_j(t)] \int_T^{T_j} \sigma(t, u)du \right\} d\tilde{W}^{(r)}_T(t) \tag{2.15}
$$

showing that within the LLMM each swaprate $\omega(t, T, r, n)$, of whatever length $n\delta$, maturity $T$ or roll $r$, will be a martingale under the corresponding swaprate measure $\tilde{P}_T^{(r)}$.

We intend to show that the volatility $\sigma(t, T, r, n)$ is nearly deterministic.

To change measures between $P_T$ and $\tilde{P}_T^{(r)}$ use

**Theorem 2.3** For $0 \leq t \leq T^* \leq T < T_j$

$$
E_T \left\{ f(T^*)|\mathcal{F}_t \right\} = \sum_{j=1}^n roz(j, r)F_T(t, T_j)E^{(r)}_T \left\{ \frac{f(T^*)}{\sum_{j=1}^n roz(j, r)F_T(T^*, T_j)} \right\}\mathcal{F}_t.
$$

**Proof:** From (2.14)

$$
dW_T(t) = d\tilde{W}^{(r)}_T(t) - \sum_{j=1}^n roz(j, r)u_j(t) \int_T^{T_j} \sigma(t, u)du dt,
$$

and so from Girsanov and (2.13), change of measure will be given by

$$
E_T \left\{ f(T^*)|\mathcal{F}_t \right\} = \tilde{E}^{(r)}_T \left\{ E \left\{ \int_T^{T_j} \sum_{j=1}^n roz(j, r)u_j(s) \int_T^{T_j} \sigma(s, u)du d\tilde{W}_T(s) \right\} f(T^*) \right\}\mathcal{F}_t
$$

$$
= \tilde{E}^{(r)}_T \left\{ \frac{\sum_{j=1}^n roz(j, r)F_T(t, T_j)}{\sum_{j=1}^n roz(j, r)F_T(T^*, T_j)} f(T^*) \right\}\mathcal{F}_t
$$

$$
= \sum_{j=1}^n \frac{roz(j, r)F_T(t, T_j)}{\sum_{j=1}^n roz(j, r)F_T(T^*, T_j)} \tilde{E}^{(r)}_T \left\{ \frac{f(T^*)}{\sum_{j=1}^n roz(j, r)F_T(T^*, T_j)} \right\}\mathcal{F}_t,
$$

because $\sum_{j=1}^n roz(j, r)F_T(T^*, T_j)$ is a $\tilde{P}_T^{(r)}$-martingale.  

A change of measure from $\mathbb{P}_T$ to $\tilde{\mathbb{P}}_T^{(r)}$ transforms the swaption formula (2.5) to
\begin{equation}
P_{swpn}(t, T, r, n) = \delta \left[ \sum_{j=1}^{n} \text{roz}(j, r) P(t, T_j) \right] \tilde{\mathbb{E}}_T^{(r)} \left\{ [\omega(T, T, r, n) - \kappa]^+ | \mathcal{F}_t \right\}
\end{equation}
(2.16)
which would lead to the Black swaption formula if $\omega(t, T, r, n)$ were lognormal under $\tilde{\mathbb{P}}_T^{(r)}$.

2.3 Swaprate Volatility Analysis

In [6] it was demonstrated by simulation, that swaprates are statistically lognormal, indicating that the volatility term
\begin{equation}
\sigma(t, T, r, n) = \sum_{j=1}^{n} \left\{ w_j(t) \gamma(t, T_j-1) + [\text{roz}(j, r) u_j(t) - w_j(t)] \int_T^{T_j} \sigma(t, u) du \right\},
\end{equation}
in (2.15) is close to deterministic. Moreover, if in (2.15) we assume
\begin{equation}
\sum_{j=1}^{n} [\text{roz}(j, r) u_j(t) - w_j(t)] \int_T^{T_j} \sigma(t, u) du \approx 0,
\end{equation}
(2.18)
and approximate the $w_j(t)$ by their initial values $w_j(0)$, then we obtain equation (23) on page 9 of [6], which produced swaption values accurate to within $1 - 2\%$.

To get a more exact approximation include the terms (2.18). From (2.2)
\begin{equation}
\int_T^{T_j} \sigma(t, u) du = \sum_{\ell=1}^{j} \mu(t, T_{\ell-1}) \gamma(t, T_{\ell-1}),
\end{equation}
where each $\mu(t, T_{\ell-1})$ is a low variance $\mathbb{P}_{T_\ell}$-martingale. Substituting in (2.17), the instantaneous swaprate volatility $\sigma(t, T, r, n)$ therefore becomes
\begin{equation}
\sigma(t, T, r, n) = \sum_{j=1}^{n} \left\{ w_j(t) \gamma(t, T_j-1) + \right.
\end{equation}
\begin{equation}
\left[ \text{roz}(j, r) u_j(t) - w_j(t) \right] \sum_{\ell=1}^{j} \mu(t, T_{\ell-1}) \gamma(t, T_{\ell-1}) \right\},
\end{equation}
(2.19)
\begin{equation}
= \sum_{j=1}^{n} \gamma(t, T_j-1) \left\{ w_j(t) + \mu(t, T_{j-1}) \sum_{\ell=j}^{n} [\text{roz}(\ell, r) u_{\ell}(t) - w_{\ell}(t)] \right\},
\end{equation}
\begin{equation}
= \sum_{j=1}^{n} A_j(t) \gamma(t, T_j-1)
\end{equation}
where
\begin{equation}
A_j(t) = w_j(t) + \mu(t, T_{j-1}) \sum_{\ell=j}^{n} [\text{roz}(\ell, r) u_{\ell}(t) - w_{\ell}(t)],
\end{equation}
(2.20)
and \( w_j(t), \mu(t, T_{j-1}) \) and \( u_\ell(t) \) are defined by (2.11), (2.2) and (2.7) above.

As remarked above, the \( w_j(t), w_j(t) \) and \( \mu(t, T_{j-1}) \) are low variance martingales under their respective measures. So approximating them by their initial values is not unreasonable, and gives the following fundamental and important result:

**Theorem 2.4** In the LLMM, swaprates are almost lognormal, with a volatility that can be closely approximated by the deterministic vector volatility function

\[
\sigma(t, T, r, n) = \sum_{j=1}^{n} A_j \gamma(t, T_{j-1}),
\]

where, for \( j = 1, \ldots, n, \)

\[
A_j = w_j(0) + \mu(0, T_{j-1}) \sum_{\ell=j}^{n} \left[ \text{roz}(\ell, r) u_\ell(0) - w_\ell(0) \right],
\]

\[
w_j(0) = \frac{F_T(0, T_j) K(0, T_{j-1})}{\sum_{j=1}^{n} F_T(0, T_j) K(0, T_{j-1})},
\]

\[
\mu(0, T_{j-1}) = \frac{\delta K(0, T_{j-1})}{1 + \delta K(0, T_{j-1})},
\]

\[
u_j(0) = \frac{F_T(0, T_j)}{\sum_{j=1}^{n} \text{roz}(j, r) F_T(0, T_j)}.
\]

### 2.4 Swaption values

Using (2.16) with the approximation (2.21), the present (time \( t = 0 \)) value of a payer swaption will be

\[
P_{\text{swpn}}(0, T, r, n) \cong \delta \left[ \sum_{j=1}^{n} \text{roz}(j, r) P(0, T_j) \right] \tilde{E}_T^{(r)} \left\{ [\omega(0, T, r, n) e(Z) - \kappa]^+ \right\},
\]

where, under the swaprate measure \( \tilde{P}_T^{(r)} \)

\[
Z \sim N(0, \beta^2 T),
\]

\[
e(Z) = \exp \left( \frac{Z - \frac{1}{2} \text{Var} Z}{T} \right),
\]
and the Black volatility $\beta(T, r, n)$ of the swaption and its swaption zeta $\zeta(T, r, n)$ are given by

\[
\zeta(T, r, n) = \beta(T, r, n)^2 T = \int_0^T |\sigma(s, T, r, n)|^2 ds = \int_0^T \sum_{j=1}^n |A_j \gamma(s, T_{j-1})|^2 ds = \sum_{i=1}^n \sum_{j=1}^n A_i A_j \int_0^T \gamma^*(s, T_{i-1}) \gamma(s, T_{j-1}) ds, = \sum_{i=1}^n \sum_{j=1}^n A_i A_j \Delta_{i,j}
\]

in which

\[
\Delta = (\Delta_{i,j}) = \left( \int_0^T \gamma^*(s, T_{i-1}) \gamma(s, T_{j-1}) \right)
\]

is the swap quadratic variation matrix. A closed formula for the approximate present value of the swaption is therefore

\[
P_{\text{swpn}}(0, T, r, n) \approx \delta \sum_{j=1}^n \text{roz}(j, r) P(0, T_j) B \{ T, r, n, \kappa, \omega(0, T, r, n), \beta(T, r, n) \},
\]

where $B(\cdot)$ is the Black futures formula.

3 Semidefinite programming

Semidefinite programming (SDP) is a relatively recent extension of linear programming in which the variables are positive semidefinite matrices, rather than vector of non-negative variables as in linear programming. General references on semidefinite programming include the review article by Vandenberghe and Boyd [23], the special issue of Mathematical Programming edited by Overton and Wolkowicz [20], and more recently the Handbook of Semidefinite Programming [25].

3.1 Semidefinite/Covariance matrices

Let $\mathbb{S}_n$ denote the space of real symmetric $n \times n$ matrices, that is

\[
\mathbb{S}_n = \{ X \in \mathbb{R}^{n \times n} : X^T = X \}.
\]
The natural inner product on the vector space \( S_n \) is the Frobenius inner product
\[ X \cdot Y = \text{trace}(X^T Y) = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij} Y_{ij}, \quad X, Y \in S_n, \]
which induces the Frobenius norm \( \|X\|_F \) defined by
\[ \|X\|_F^2 = X \cdot X. \]

A matrix \( X \in S_n \) is positive semidefinite (written \( X \succeq 0 \)) if any one of the following equivalent conditions hold:

1. For every vector \( w \in \mathbb{R}^n \), \( w^T X w \geq 0 \).
2. Every eigenvalue of \( X \) is nonnegative: \( \lambda_i(X) \geq 0 \), \( i = 1, \ldots, n \).
3. The Cholesky factorization \( X = LL^T \), where \( L \in \mathbb{R}^n \) is lower triangular, exists.
4. \( X \) has a square root \( Y \in S_n \) such that \( X = YY^T \).
5. \( X \cdot Y \geq 0 \) for every \( Y \in S_n \), \( Y \succeq 0 \).

**Remark 3.1** These are precisely the class of matrices of interest to us because a matrix \( X \in S_n \) is a covariance matrix iff \( X \) is positive semidefinite (see [16] for example).

Semidefinite programming deals with optimization problems of the form
\[
\begin{align*}
\text{find} & \quad X \in S_n \\
\text{to minimize} & \quad C \cdot X \\
\text{subject to} & \quad A_k \cdot X = b_k \quad k = 1, \ldots, m \\
& \quad X \succeq 0,
\end{align*} \tag{3.1}
\]
where \( C \in S_n \), \( A_k \in S_n \) for \( k = 1, \ldots, m \), and \( b_k \in \mathbb{R} \) for \( k = 1, \ldots, m \) are given data. Here the objective function \( C \cdot X \) and constraints \( A_k \cdot X = b_k \) are all linear functions of the variables \( X \). The set \( \{ X \in S_n : X \succeq 0 \} \) is a convex cone, but is not polyhedral, unlike \( \{ x \in \mathbb{R}^n : x \geq 0 \} \) which is a polyhedral convex cone.

The dual problem is
\[
\begin{align*}
\text{find} & \quad y \in \mathbb{R}^m \\
\text{to maximize} & \quad y^T b \\
\text{subject to} & \quad C - \sum_{k=1}^{m} y_k A_k \succeq 0.
\end{align*} \tag{3.2}
\]
Weak duality holds in that if \( X \in S_n \) is feasible for (3.1) and \( y \in \mathbb{R}^m \) is feasible in (3.2), then \( C \cdot X \geq y^T b \). If both the primal problem (3.1) and the dual problem (3.2) have strictly feasible points (a stronger condition than is required in linear programming), then the primal and dual objective values are equal at a solution. The dual formulation is most commonly used in the control theory literature (see [5] for example).
Linear programming is just a special case of semidefinite programming in which the matrices are diagonal $X = \text{diag}(x)$ and $A_k = \text{diag}(a_k)$ where $x, a_k \in \mathbb{R}^n$, so $X \cdot A_k = a_k^T x$. Several different semidefinite matrices $X_1, X_2, \ldots, X_p$ can be accumulated into one larger block diagonal semidefinite matrix $X = \text{diag}(X_1, X_2, \ldots, X_p)$.

We will consider a standard problem which explicitly includes both symmetric positive semidefinite matrix variables and vectors of non-negative variables:

$$\begin{align*}
\text{find} \quad & X \in S_n, x \in \mathbb{R}^n \\
\text{to minimize} \quad & C \cdot X + c^T x \\
\text{subject to} \quad & A_k \cdot X = b_k \quad k = 1, \ldots, m_s, \\
& a_j^T x = \beta_j \quad j = 1, \ldots, m_l, \\
& X \succeq 0, \quad x \geq 0.
\end{align*}$$

(3.3)

Here $X$ and $A_k, k = 1, \ldots, m_s$ are symmetric block diagonal matrices, each with $p$ blocks of size $n_1, \ldots, n_p$, where $n = \sum_{i=1}^p n_i$.

Karmarkar’s 1984 polynomial time algorithm for linear programming lead to an explosion of work in interior point methods for linear programming (see [24] for example). Many of the extensions to more general convex programs were pioneered by Nesterov and Nemirovski [19]. Instead of following edges of the feasible region $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ to a minimum vertex as in the simplex method for linear programming, interior point methods use strictly positive iterates $x > 0$ and follow the “central path” to a solution. In linear programming the nonnegative orthant $x \geq 0$ is a polyhedral convex cone, so the feasible region is the intersection of hyperplanes and a polyhedral convex cone. In semidefinite programming the variables are symmetric matrices $X \in S_n$, which must lie in the convex (but not polyhedral) cone $X \succeq 0$ intersected with hyperplanes $X \cdot A_k = b_k$ for $k = 1, \ldots, m$. Because they move through the interior of the cone constraint, interior point methods have been successfully generalized to provide efficient methods for semidefinite programming (see [1, 22] for example).

### 3.2 Objective functions

Many different problems can be formulated as semidefinite problems (see [23]). In particular semidefinite programming is closely related to eigenvalue optimization problems. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $X$, then

$$I \cdot X = \text{trace}(X) = \sum_{i=1}^n \lambda_i,$$

(where $I$ is the appropriately sized identity matrix). The expression $I \cdot X$ is one of the simplest objective functions to minimize in semidefinite programming, and can be used to “minimize” the size of $X \succeq 0$.

The largest eigenvalue of a symmetric matrix can be minimized subject to the con-
Fitting Swaption Volatilities using Semidefinite Programming

strains \( A_k \cdot X = b_k, \ k = 1, \ldots, m \), by adding a further constraint, and solving

\[
\begin{align*}
\text{find} & \quad X \in S_n, \ z \in \mathbb{R} \\
\text{to minimize} & \quad z \\
\text{subject to} & \quad \zeta I - X \succeq 0, \\
& \quad A_k \cdot X = b_k, \ k = 1, \ldots, m. 
\end{align*}
\]

This can be simply justified by noting that the eigenvalues of \( \zeta I - X \) are \( \zeta - \lambda_i(X) \), so that \( \zeta I - X \succeq 0 \) is equivalent to \( \zeta \geq \lambda_i(X) \) for \( i = 1, \ldots, n \). Problem (3.4) can be converted into the form (3.3) by adding a surplus matrix \( S = \zeta I - X \succeq 0 \).

The fit to the swaption volatility matrix is imposed through linear constraints of the form \( A_k \cdot X = b_k \) (see Section 4), while a variety of different objectives can be used. In particular we are interested in making the positive semidefinite covariance matrix “close” to a historical covariance matrix \( G \). The difference \( X - G \) is in general not positive semidefinite, so simple objective like \( \text{trace}(X - G) \) or the largest eigenvalue of \( X - G \) will not work.

We cannot use the Frobenius norm, or its square, directly to force \( X \) close to a target matrix \( G \in S_n \) because \( \|X - G\|_F^2 = (X - G) \cdot (X - G) \) is not linear in \( X \). However a variety of other matrix norms can be used. The 2-norm,

\[
\|X - G\|_2 = \max_{i=1,\ldots, n} |\lambda_i(X - G)|
\]

can be minimized by adding two further constraints and solving

\[
\begin{align*}
\text{find} & \quad X \in S_n, \ z \in \mathbb{R} \\
\text{to minimize} & \quad z \\
\text{subject to} & \quad \zeta I - (X - G) \succeq 0, \\
& \quad \zeta I + (X - G) \succeq 0, \\
& \quad A_k \cdot X = b_k, \ k = 1, \ldots, m, \\
& \quad X \succeq 0, \ z \geq 0.
\end{align*}
\]

This can be done as the constraints \( \zeta I - (X - G) \succeq 0 \) and \( \zeta I + (X - G) \succeq 0 \) imply that \( \zeta \geq \lambda_i(X - G) \) and \( \zeta \geq -\lambda_i(X - G) \), so that

\[
\zeta \geq \max_{i=1,\ldots, n} \{\lambda_i(X - G), -\lambda_i(X - G)\} = \|X - G\|_2.
\]

The constraint \( \zeta \geq 0 \) is in fact implied by the positive semidefinite constraints. From [10] a bound on the two norm of a matrix gives bounds on the maximum difference between elements

\[
\max_{i,j=1,\ldots, n} |X_{ij} - G_{ij}| \leq \|X - G\|_2 \leq n \max_{i,j=1,\ldots, n} |X_{ij} - G_{ij}|.
\]

A much more difficult problem is to find the best covariance matrix \( X \) with \( \text{rank}(X) \leq r \) where the maximal rank \( r, \ 1 \leq r < n \), is specified. Note that if \( X \) is rank-1, so
\[ X = \lambda xx^*, \text{ where } \|x\|_2 = 1, \text{ and } \lambda \geq 0, \text{ then the constraints } A_k \cdot X = b_k \text{ are equivalent to the quadratic constraints } \lambda x^* A_k x = b_k. \]

Semidefinite programming is part of optimization over convex cones. Interior point methods have been developed for problems which include semidefinite, quadratic cone and linear constraints. Although we explicitly use both semidefinite and linear variables we did not directly use the quadratic cone constraints. We used the MATLAB research software SDPPACK available from [1] in our experiments. Another possibility is the package [22].

4 Volatility Structure

4.1 Cap and swaption volatilities

Let \( \gamma(t, T) \) be the volatility function for lognormal forwards over the interval \([T, T + \delta]\). Suppose all caps and swaptions mature at multiples of \( \delta \), and set \( T_i = i\delta \) for \( i = 0, 1, \ldots, N \). We are interested in the volatility \( \sigma(t, m, r, n) \) of the \( r \)-roll forward swap \( \text{Swap}(m, n, r) \) starting at time \( T_m \) (first exchange at \( T_{m+r} \)) and finishing at \( T_{m+n} \) (last of \( \frac{n-m}{r} \) exchanges at \( T_{m+n} \) with \( m + n \leq N \)).

**Remark 4.1** Because caplets are just one exchange swaptions we will treat them as particular members of one general swaption class. This entails no loss of generality, eases the entry of constraints into the SDP framework, and simplifies programming.

Section 2, and in particular Theorem 2.4, imply that the swaprate volatility can be accurately approximated by a linear combination of forward volatilities. In terms of \( \gamma(\cdot) \), the Black volatility \( \beta(m, n, r) \) of a swaption \( \text{Swpn}(m, n, r) \) on the forward swap \( \text{Swap}(m, n, r) \) satisfies

\[
T_m \beta^2(m, n, r) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_i^{(m)} A_j^{(m)} \int_0^{T_m} \gamma^*(t, T_{i+m-1}) \gamma(t, T_{j+m-1}) dt. \tag{4.1}
\]

Part (only part, because there may be extra correlation requirement etc) of the parameterization problem is to find a \( \gamma(\cdot) \) that returns given market cap and swaption volatilities via (4.1).

4.2 Choice of volatility function

We first express the constraints (4.1) as the Frobenius inner product of constraint matrices with a block covariance matrix that we will find using the semidefinite programming.

Let \( \gamma \) be of the form

\[
\gamma(t, T) = \sum_{k=1}^{\infty} \xi^{(k)}(T - t)[t \in ((k-1)\delta, k\delta)], \quad (k-1)\delta < t \leq T, \tag{4.2}
\]

where the vector functions \( \xi^{(k)} \) are homogeneous on the \( k \)th time interval \( t \in ((k-1)\delta, k\delta] \).
Remark 4.2 If the $\xi^{(k)}(\cdot)$ were identical on each layer $k$, then the model would be completely homogeneous and the $\xi^{(k)}(\cdot)$ can be easily estimated from historic data. So the choice (4.2) of $\gamma$ implies a sensible notion of, at least, instantaneous correlation. In contrast there is none for an inhomogeneous $\gamma$, like for example
$$
\gamma(t, T) = \psi(t)\phi(T).
$$

Introduce the $N-1$ semidefinite matrices $\Gamma^{(k)} \in \mathbb{S}_{N-k}$ for $k = 1, 2, \ldots, N-1$ defined by
$$
\Gamma^{(k)} = \left( \Gamma^{(k)}_{i,j} \right) = \left( \frac{1}{\delta} \int_0^{\delta} \xi^{(k)*}(T_{i-1} - t)\xi^{(k)}(T_{j-1} - t)dt \right),
$$
i, j = 2, \ldots, N - k + 1.

Note that $\Gamma^{(1)} \in \mathbb{S}_{N-1}, \ldots, \Gamma^{(N-1)} \in \mathbb{S}_1$. Form the $\frac{1}{2}N(N-1) \times \frac{1}{2}N(N-1)$ block-diagonal semidefinite matrix
$$
\Gamma = \begin{pmatrix}
\Gamma^{(1)} & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & \Gamma^{(k)} & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & \Gamma^{(N-1)}
\end{pmatrix}.
$$

A typical contribution to swaption volatility from the $k^{th}$ layer will be, for $i, j \geq k + 1$,
$$
\int_0^{\delta(k)} \gamma^*(t, T_{i-1})\gamma(t, T_{j-1})dt = \int_0^{\delta(k)} \xi^{(k)*}(T_{i-1} - t)\xi^{(k)}(T_{j-1} - t)dt
$$
$$
= \int_0^{\delta} \xi^{(k)*}(T_{i-k} - t)\xi^{(k)}(T_{j-k} - t)dt
$$
$$
= \delta \Gamma^{(k)}_{i-k+1,j-k+1},
$$
which sums to
$$
\int_0^{T_m} \gamma^*(t, T_{i+m-1})\gamma(t, T_{j+m-1})dt = \delta \sum_{k=1}^m \Gamma^{(k)}_{i+m-k+1,j+m-k+1}.
$$

Remark 4.3 $\Gamma^{(k)}_{i,j}$ was scaled by $\frac{1}{\delta}$ so that the roots of its diagonal elements are forward instantaneous volatilities averaged over a $\delta$-interval, and thus roughly equal to forward volatilities.

Hence, from (4.1), the swaption Swpn($m, n, r$) has Black volatility $\beta(m, n, r)$ given by
$$
\sum_{i=1}^n \sum_{j=1}^n A^{(m)}_i A^{(m)}_j \int_0^{T_m} \gamma^*(t, T_{i+m-1})\gamma(t, T_{j+m-1})dt = T_m \beta^2(m, n, r),
$$
leading to the equality constraint
\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{1}{m} A^{(m)}_i A^{(m)}_j \Gamma^{(k)}_{i+m-k+1,j+m-k+1} = \beta^2(m, n, r). \tag{4.3}
\]
This constraint can be rewritten as the Frobenius product
\[
u^{(m)} \bullet \Gamma = b^{(m)}, \tag{4.4}
\]
where the \( \frac{1}{2} N(N-1) \times \frac{1}{2} N(N-1) \) constraint matrices \( u^{(m)} \) are formed from the \( A^{(m)}_i \) according to (4.3).

### 4.3 Simplification

To fit a full quarterly volatility spectrum with time to maturity plus length of underlying ranging up to 10 years requires \( N = 40 \), making \( \Gamma \) a \( 780 \times 780 \) matrix. That is a computationally expensive problem for a 300 MHz laptop, despite the block structure of \( \Gamma \). So we “compact” our matrices into annual blocks and hence reduce the scale of the problem by a factor of 4.

Specifically, introduce the nine \( (11-K) \times (11-K) \) matrices \( \Gamma^{(K)} \), for \( K = 1, \ldots, 9 \), defined by
\[
\Gamma^{(k)}_{i,j} = X^{(K)}_{I,J} \quad \text{whenever} \quad i = 4(I-1) + \alpha, \quad j = 4(J-1) + \beta, \quad k = 4(K-1) + \gamma, \\
\text{for} \quad I, J = 1, \ldots, 11 - K, \quad \text{and} \quad \alpha, \beta, \gamma = 1, \ldots, 4. \tag{4.5}
\]

Because both the constraints (4.3) and “compaction” (4.5) are linear, clearly (4.4) can be rewritten in the form
\[
U^{(m)} \bullet X = b^{(m)}, \tag{4.6}
\]
where \( X \) is the (now more computationally tractable) \( 54 \times 54 \) block-diagonal semidefinite matrix defined by
\[
X = \begin{pmatrix}
X^{(1)} & \ldots & 0 & \ldots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \ldots & X^{(K)} & \ldots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & \ldots & X^{(9)}
\end{pmatrix},
\]
and the \( U^{(m)} \) are obtained by “compacting” the \( u^{(m)} \). We can regard each successive block \( X^{(K)} \) as the covariance of the \( K^{th} \) annual layer.
4.4 Implied Correlation

Our aim is to satisfy the swaption volatility constraints (4.6) with an $X$ that embodies, in some sense, a satisfactory implied correlation structure. Our basic assumption reflects the commonly accepted wisdom that the principle components of the historical covariance matrix are stable (although the level of volatility, that is eigenvalues, may change), and about 5 factors are needed to explain movement of the simple forward curve.

Because of the simplification applied in § 4.3, an appropriate historical covariance matrix for us to use will be the one generated by annual simple forwards (we were unable to identify a suitable quarterly estimator). In view of our 10-year horizon, suppose that it is $10 \times 10$, and let $e_1, e_2, \ldots, e_5$ be its first 5 eigenvectors ordered by size of the corresponding eigenvalue. Looking backwards in time, we would not therefore be surprised to find a good model for the forward curve returning a covariance matrix of form

$$\sum_{i=1}^{5} \lambda_i e_i e_i^\ast,$$

where the $\lambda_i$ are constants.

Using MATLAB matrix notation, for $i = 1, \ldots, 5$ set

$$E_i^{(1)} = e_i e_i^\ast,$$

$$E_i^{(K)} = E_i^{(K-1)} (1:10-K, 1:10-K), \quad K = 2, \ldots, 9.$$ 

We make a linear combination of the $E_i^{(K)}$ the target for the for the covariance $X^{(K)}$ in the $K^{th}$ layer. Specifically, in each layer $K$ we seek $X^{(K)}$ and constants $\lambda_i^{(K)} (i = 1, \ldots, 5)$ to minimize

$$\left\| X^{(K)} - \sum_{i=1}^{5} \lambda_i^{(K)} E_i^{(K)} \right\|_2.$$

That can be done, for example, by solving the optimization problem

$$\begin{align*}
\text{find} \quad & X^{(K)}, \lambda_i^{(K)}, i = 1, \ldots, 5; K = 1, \ldots, 9 \\
\text{to minimize} \quad & \sum_{K=1}^{9} \left\| X^{(K)} - \sum_{i=1}^{5} \lambda_i^{(K)} E_i^{(K)} \right\|_2, \\
\text{subject to} \quad & U^{(m)} \bullet X = b^{(m)}, \\
& X^{(K)} \succeq 0, \lambda_i^{(K)} \geq 0.
\end{align*}$$

Various permutations of this setup allows correlations in different layers to be “tied together” resulting in a reasonably uniform distribution of volatility throughout layers and maturities.
Table 1: The downward sloping volatility surface of 2 Feb 1999 considered by Longstaff et al, which has a 3.2% profile over that part of the matrix to which we parameterize.

5 Numerical Results

5.1 Volatility data

In what follows expiry means the time from now to exercise of a European swaption (Euro), tenor means the length in years of the underlying forward swap into which the Euro exercises, and maturity means the time from now to the last exchange of the swap. Thus

\[ \text{Expiry} + \text{Tenor} = \text{Maturity} \]

Euros are described in Expiry/Maturity terms; for example, a 1/5 Euro means it expires in 1 year, matures in 5 years and has tenor 4 years. In a similar vein, a 1/5 fixed maturity Bermudan means the first exercise may occur at 1 year, and it and all later exercises are into swaps maturing at 5 years. This is the notation of Tables 6, 7, 8 and 9 below. The at-the-money strike of the underlying 1/5 Euro is defined to be the at-the-money strike for the 1/5 Bermudan.

Somewhat confusingly, but for good practical reasons, the swaption volatility matrix is expressed in Expiry × Tenor terms. Thus in Tables 1, 2 and 3 the volatility of the 1/5 Euro can be found in the 1 × 4 position. The reverse diagonal of the 1/5 fixed maturity Bermudan is the set of volatilities in the 1 × 4, 2 × 3, 3 × 2 and 4 × 1 positions of the swaption volatility matrix; they respectively comprise the volatilities of the 1/5, 2/5, 3/5 and 4/5 Euros which underly the 1/5 Bermudan.

Most existing production programs to price fixed maturity Bermudans seem to parameterize on a deal-by-deal basis to their reverse diagonal. Moreover, they are one-factor models, like Andersen’s BDT and 1-Factor models [4], or the 1-Factor model of Longstaff et al [15], or our (so-called) Analytic 1-Factor and Diag 1-Factor models. Multi-factor
versions, like the N-Factor model of Longstaff et al [15] or our Diag 5-Factor model, also fit historic correlation. More ambitious programs fit more swaptions, either approximately or using contorted volatility functions, and may involve some measure of implied correlation, but as far as we are aware they are not widely used in practice.

This state of affairs, combined with the importance of Bermudans as a financial tool persuaded us to test our SDP parameterization on them. Specifically, we were interested in whether, with just one parameterization and over a range of data, our SDP method could return similar prices to the standard models for Bermudans off two (significantly different) reverse diagonals.

To compare models we priced five semiannual Bermudans (1/5, 2/5, 1/10, 3/10 and 5/10) with four sets of data:

1. The flat data set used by Andersen [4], in which semiannual yields were 6%, and to calculate the prices of the 1/5 and 2/5 Bermudans (respectively, the 1/10, 3/10 and 5/10 Bermudans) the Black volatility of all European swaptions was 20% (15% respectively).

2. The downward sloping volatility surface of 2 Feb 1999 (Table 1) considered by Longstaff et al [15], in which that part of the volatility surface to which we parameterize has a 3.2% profile (difference between the largest and smallest Black volatilities).

3. The nearly flat volatility surface of 20 Aug 1998 (Table 2) which rises and falls slightly with a 1.4% profile over the part to which we parameterize.

4. The steeply falling volatility surface of 9 Oct 1998 (Table 3) which has an 11.0% profile over the part to which we parameterize. Our experience has been that surfaces like this are particularly demanding of parameterization routines.

<table>
<thead>
<tr>
<th>20 Aug 1998</th>
<th>3mo</th>
<th>6mo</th>
<th>9mo</th>
<th>12mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical %</td>
<td>5.688</td>
<td>5.719</td>
<td>5.719</td>
<td>5.731</td>
</tr>
<tr>
<td>SwapRates %</td>
<td>5.705</td>
<td>5.755</td>
<td>5.804</td>
<td>5.844</td>
</tr>
<tr>
<td>Black Vol %</td>
<td>5.705</td>
<td>5.755</td>
<td>5.804</td>
<td>5.844</td>
</tr>
</tbody>
</table>

Table 2: The nearly flat volatility surface of 20 Aug 1998, which rises and falls slightly with a 1.4% profile over the part to which we parameterize.
Tables 1, 2, 3 contain input data from the US taken off Bloomberg for the three days 2 Feb 99, 20 Aug 98 and 9 Oct 98. Swap rates are semi-annual, the daycount for both cash and swap rates is Actual/360, and volatilities are percentage Black semiannual swaption volatilities.

Generally volatility surfaces either rise and fall or just fall, rarely do they uniformly rise. Moreover the 1/5, 2/5 and 1/10, 3/10, 5/10 reverse diagonals are fairly widely separated. So we feel that, while more comparisons might be better, our choice of data and Bermudan maturities ought to constitute a robust and revealing test.

The models used by Andersen and Longstaff et al are described in their papers [4] and [15] respectively. The volatility function in our Analytic 1-Factor and Diag 1-Factor models depends only on maturity; the former splines values on nodes and numerically integrates down through the exercise times (hence the name “Analytic”), while the latter combines the simulation techniques of Glasserman et al [9] with the Bermudan techniques of Longstaff et al [14]. The volatility function in the Diag 5-Factor model depends mainly on maturity, but also incorporates a homogeneous multi-factor component to carry historical correlation.

### 5.2 Covariance approximation

Table 4 specifies our covariance matrix. Annual forwards were used because we felt their covariance best corresponded to the way we “compacted” our matrices and volatility functions from quarterly to annual periods in order to get a computationally simple enough problem for our SDP software to tackle (we were also unable to find a better estimator).

Targeting actual historic correlation via the SDP objective function is not possible because the problem is non-linear. So we fell back on what seems to be part of the folk
Fitting Swaption Volatilities using Semidefinite Programming

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.392</td>
<td>2.696</td>
<td>2.564</td>
<td>2.365</td>
<td>2.189</td>
<td>2.038</td>
<td>1.913</td>
<td>1.810</td>
<td>1.729</td>
<td>1.668</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.760</td>
<td>3.834</td>
<td>3.634</td>
<td>3.413</td>
<td>3.229</td>
<td>3.086</td>
<td>2.975</td>
<td>2.889</td>
<td>2.825</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.745</td>
<td>3.641</td>
<td>3.526</td>
<td>3.433</td>
<td>3.373</td>
<td>3.341</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.590</td>
<td>3.522</td>
<td>3.461</td>
<td>3.421</td>
<td>3.397</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.518</td>
<td>3.510</td>
<td>3.497</td>
<td>3.480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.547</td>
<td>3.561</td>
<td>3.555</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
<th>6 yr</th>
<th>7 yr</th>
<th>8 yr</th>
<th>9 yr</th>
<th>10 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.596</td>
<td>3.603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: 100 times the annualized historical covariance matrix for annual forwards, computed from data running from May 1994 to Feb 2000.

<table>
<thead>
<tr>
<th>Total</th>
<th>&lt; 20%</th>
<th>&lt; 30%</th>
<th>&gt; 30%</th>
<th>&gt; 50%</th>
<th>&gt; 70%</th>
<th>&gt; 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat 15%</td>
<td>371</td>
<td>371</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2 Feb 1999</td>
<td>323</td>
<td>365</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20 Aug 1998</td>
<td>363</td>
<td>371</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9 Oct 1998</td>
<td>279</td>
<td>313</td>
<td>58</td>
<td>26</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5: Number of elements, out of a total of 371, whose absolute percentage differences between implied and target correlations lie in the given ranges, for the four data sets.

The wisdom of financial mathematics (we have no references), that the shapes of the largest (ranked by eigenvalue) eigenvectors of the historical covariance matrix are invariant. In our SDP routine we minimize the matrix 2-norm of the difference between the covariance matrix we are seeking, and a target covariance matrix formed by any linear combination of the first five eigenvectors times their transposes, of the historic covariance.

Thus the correlations implied by our fitted covariances approximate the correlations determined by the target covariance, which in turn approximate actual historic correlations. Experiments, in which we changed eigenvalues but not eigenvectors, convinced us that absolute differences of up to 30% between corresponding elements of correlation matrices related in this way, generally make little difference in practice. So we believe a good covariance fit simultaneously gets all elements of the implied correlation matrices to within 30% of corresponding correlations derived from the target, and pointedly ignore the actual historic correlation matrix. Our dilemma arises, of course, from the existence of two appealing but incompatible notions of “good correlation fit”, only one of which suits the SDP framework. The results of our covariance fit are contained in Table 5.

We have not analyzed implied correlation extensively, but despite the many unanswered questions that occur to us, some comments can be made. Some 371 comparisons are possible between elements in the implied correlation matrices and corresponding el-
Table 6: Bermudan and Euro prices in the flat case. Numbers which strike us as interesting have been rendered bold.

Table 5: Numbers of elements in the target correlation matrices. Listed in Table 5 are the numbers of elements whose absolute percentage difference falls in the indicated category. By our criteria, clearly the Flat and 20 Aug 98 fits can be regarded as good, the 2 Feb 99 as satisfactory, and the 9 Oct 98 fit as poor. All efforts to significantly improve the last fit failed, so we feel that the result reflects reality: when the volatility surface slopes steeply down, implied correlation moves sharply away from historical correlation.

### 5.3 Bermudan prices

Our Analytic and Diag models were parameterized on a deal-by-deal basis to the reverse diagonal of the Bermudan being considered. To price the 1/5 and 2/5 Bermudans we parameterized to the 1/5, 2/5, 3/5 and 4/5 Euros, whose volatilities are boxed in Tables 1, 2 and 3. To price the 1/10, 3/10 and 5/10 Bermudans we parameterized to 3/10, 5/10 and 7/10 Euros (volatilities boxed), and also to the 1/10, 2/10, 4/10, 6/10, 8/10 and 9/10 Euros by linearly interpolating their unspecified volatilities off the tables. Our SDP routine parameterized to those 31 Euros, with maturity less than or equal to 10 years.
2 Feb 1999 | Exp/Mat | 1/5 | 2/5 | 1/10 | 3/10 | 5/10 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>European</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer</td>
<td>Black</td>
<td>127.1</td>
<td>134.4</td>
<td>249.8</td>
<td>323.2</td>
<td>270.9</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>127.3</td>
<td>134.8</td>
<td>251.5</td>
<td>321.6</td>
<td>269.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6)</td>
<td>(0.7)</td>
<td>(1.2)</td>
<td>(1.6)</td>
<td>(1.4)</td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer</td>
<td>Analytic 1-Factor</td>
<td>181.6</td>
<td>158.0</td>
<td><strong>474.3</strong></td>
<td>418.4</td>
<td>307.9</td>
</tr>
<tr>
<td></td>
<td>Longstaff 1-Factor</td>
<td>179.7</td>
<td>157.0</td>
<td><strong>466.1</strong></td>
<td>411.3</td>
<td>304.1</td>
</tr>
<tr>
<td></td>
<td>Diag 1-Factor</td>
<td>180.6</td>
<td>157.7</td>
<td><strong>463.7</strong></td>
<td>415.0</td>
<td>309.4</td>
</tr>
<tr>
<td></td>
<td>(0.3)</td>
<td>(0.4)</td>
<td>(0.7)</td>
<td>(0.8)</td>
<td>(0.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Longstaff N-Factor</td>
<td>189.7</td>
<td>164.9</td>
<td><strong>500.4</strong></td>
<td>439.1</td>
<td>322.3</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>182.4</td>
<td>159.5</td>
<td><strong>473.9</strong></td>
<td>421.8</td>
<td>313.5</td>
</tr>
<tr>
<td></td>
<td>(0.3)</td>
<td>(0.4)</td>
<td>(0.8)</td>
<td>(0.8)</td>
<td>(0.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td><strong>183.8</strong></td>
<td><strong>161.2</strong></td>
<td><strong>460.5</strong></td>
<td><strong>407.2</strong></td>
<td><strong>314.0</strong></td>
</tr>
<tr>
<td></td>
<td>(0.7)</td>
<td>(0.7)</td>
<td>(1.6)</td>
<td>(1.6)</td>
<td>(1.4)</td>
<td></td>
</tr>
<tr>
<td><strong>European</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiver</td>
<td>Black</td>
<td>127.1</td>
<td>134.4</td>
<td>249.8</td>
<td>323.2</td>
<td>270.9</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>126.8</td>
<td>134.8</td>
<td>252.5</td>
<td>322.6</td>
<td>271.2</td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(0.6)</td>
<td>(1.1)</td>
<td>(1.4)</td>
<td>(1.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiver</td>
<td>Analytic 1-Factor</td>
<td>175.4</td>
<td>154.8</td>
<td>409.8</td>
<td>383.6</td>
<td>295.8</td>
</tr>
<tr>
<td></td>
<td>Longstaff 1-Factor</td>
<td>175.5</td>
<td>154.3</td>
<td>412.7</td>
<td>386.7</td>
<td>295.1</td>
</tr>
<tr>
<td></td>
<td>Diag 1-Factor</td>
<td>174.8</td>
<td>155.0</td>
<td>407.9</td>
<td>382.2</td>
<td>299.2</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.5)</td>
<td>(0.6)</td>
<td>(0.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Longstaff N-Factor</td>
<td>181.9</td>
<td>159.6</td>
<td>431.1</td>
<td>401.5</td>
<td>305.3</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>175.6</td>
<td>155.7</td>
<td>409.8</td>
<td>383.9</td>
<td>300.1</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.5)</td>
<td>(0.6)</td>
<td>(0.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>176.1</td>
<td>156.6</td>
<td>404.0</td>
<td>374.3</td>
<td>301.0</td>
</tr>
<tr>
<td></td>
<td>(0.6)</td>
<td>(0.5)</td>
<td>(1.3)</td>
<td>(1.3)</td>
<td>(1.1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Bermudan and Euro prices for the downward sloping volatility surface considered by Longstaff et al. Numbers which strike us as interesting have been rendered bold.

years, which appear above the lower stepped line in the tables.

Our results are contained in Tables 6, 7, 8 and 9 which list the prices of at-the-money semi-annual European and Bermudan options in points–per–million ($100 units per $1,000,000 face value) for each of our four data sets.

In a separate exercise to back check each parameterization (not tabulated), at-the-money values of the 31 fitted Euros were simulated; in all cases values returned were within two standard deviations of the corresponding Black value. Those 31 Euros include the 1/5, 2/5, 3/10, 5/10 Euros, so our simulated European prices in Tables 6–9 ought to agree with the Black prices. They do, with the exception of the 3/10 European Payer in Table 9 which we regard as a statistically acceptable anomaly. So our SDP parameterization successfully returns fitted options.

To our surprise, Bermudan prices produced by the SDP parameterizations were either roughly equal to, or less than, prices from the single factor models. In particular, prices of the “real Bermudan” 1/10s (the 5/10s are “half vanilla”) were almost uniformly lower, and in the case of the 1/10 Payer of 9 Oct 1998 significantly so (40-50 points or around 9% of value). This finding contrasts with the view of Longstaff et. al [15] who believe
Fitting Swaption Volatilities using Semidefinite Programming

20 Aug 1998

<table>
<thead>
<tr>
<th></th>
<th>Exp/Mat</th>
<th>1/5</th>
<th>2/5</th>
<th>1/10</th>
<th>3/10</th>
<th>5/10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>European</strong></td>
<td>Black</td>
<td>99.0</td>
<td>110.9</td>
<td>192.9</td>
<td>258.5</td>
<td>225.2</td>
</tr>
<tr>
<td>Payer</td>
<td>SDP 5-Factor</td>
<td>99.2</td>
<td>110.2</td>
<td>190.3</td>
<td>257.7</td>
<td>225.1</td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td>Analytic 1-Factor</td>
<td>147.8</td>
<td>130.8</td>
<td><strong>373.6</strong></td>
<td>332.4</td>
<td>253.6</td>
</tr>
<tr>
<td>Payer</td>
<td>Diag 1-Factor</td>
<td>147.0</td>
<td>130.5</td>
<td><strong>366.6</strong></td>
<td>329.4</td>
<td>253.6</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>148.1</td>
<td>132.0</td>
<td><strong>373.5</strong></td>
<td>334.7</td>
<td>257.3</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td><strong>150.2</strong></td>
<td><strong>129.2</strong></td>
<td><strong>371.7</strong></td>
<td><strong>329.3</strong></td>
<td><strong>254.7</strong></td>
</tr>
<tr>
<td><strong>European</strong></td>
<td>Black</td>
<td>99.0</td>
<td>110.9</td>
<td>192.9</td>
<td>258.5</td>
<td>225.2</td>
</tr>
<tr>
<td>Receiver</td>
<td>SDP 5-Factor</td>
<td>99.3</td>
<td>110.1</td>
<td>190.8</td>
<td>257.9</td>
<td>226.0</td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td>Analytic 1-Factor</td>
<td>136.1</td>
<td>126.5</td>
<td>328.3</td>
<td>315.8</td>
<td>248.0</td>
</tr>
<tr>
<td>Receiver</td>
<td>Diag 1-Factor</td>
<td>135.4</td>
<td>126.6</td>
<td>327.1</td>
<td>314.7</td>
<td>249.9</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>136.1</td>
<td>127.2</td>
<td>329.2</td>
<td>316.4</td>
<td>250.4</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>137.6</td>
<td>124.7</td>
<td>327.3</td>
<td>312.4</td>
<td>249.1</td>
</tr>
</tbody>
</table>

Table 8: Bermudan and Euro prices for a fairly flat volatility surface. Numbers which strike us as interesting have been rendered bold.

single factor models underprice Bermudans.

We are uncertain just what to make of our results. Our experience with the Diag 5-Factor model was that one could shift prices 20–30 points by massaging correlation; tight fits of forwards to cash and swap rates produced lumpy unstable eigenvectors and higher prices, while looser fits produced smoother more stable eigenvectors and lower prices. Perhaps there is some roughly similar effect, not yet clear to us, that works in the opposite direction for our SDP parameterizations.

Nevertheless, we take comfort from the fact that Bermudans have been priced with single factor models for a long time, with only a couple (as far as we know) of rumoured instances of significant writeoffs. Because banks tend to be overall buyers of Bermudans, such writeoffs would likely arise from models that overprice rather than underprice. These thoughts combine to favour our findings.

References


Table 9: Bermudan prices for the steeply falling volatility surface. Numbers which strike us as interesting have been rendered bold.

<table>
<thead>
<tr>
<th>9 Oct 1998</th>
<th>Exp/Mat</th>
<th>1/5</th>
<th>2/5</th>
<th>1/10</th>
<th>3/10</th>
<th>5/10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>European</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer</td>
<td>Black</td>
<td>158.3</td>
<td>159.3</td>
<td>290.9</td>
<td>327.0</td>
<td>263.4</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>159.1</td>
<td>158.9</td>
<td>294.9</td>
<td>330.6</td>
<td>264.3</td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer</td>
<td>Analytic 1-Factor</td>
<td>243.9</td>
<td>194.4</td>
<td>550.4</td>
<td>425.7</td>
<td>302.7</td>
</tr>
<tr>
<td></td>
<td>Diag 1-Factor</td>
<td>243.2</td>
<td>195.2</td>
<td>542.4</td>
<td>423.4</td>
<td>303.0</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>247.2</td>
<td>198.0</td>
<td>558.0</td>
<td>431.2</td>
<td>307.0</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>237.9</td>
<td>194.5</td>
<td>506.4</td>
<td>413.2</td>
<td>312.9</td>
</tr>
<tr>
<td><strong>European</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiver</td>
<td>Black</td>
<td>158.3</td>
<td>159.3</td>
<td>290.9</td>
<td>327.0</td>
<td>263.4</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>157.9</td>
<td>160.4</td>
<td>297.5</td>
<td>326.7</td>
<td>265.1</td>
</tr>
<tr>
<td><strong>Bermudan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiver</td>
<td>Analytic 1-Factor</td>
<td>195.4</td>
<td>175.2</td>
<td>410.2</td>
<td>379.4</td>
<td>286.9</td>
</tr>
<tr>
<td></td>
<td>Diag 1-Factor</td>
<td>195.4</td>
<td>177.0</td>
<td>410.3</td>
<td>378.4</td>
<td>288.6</td>
</tr>
<tr>
<td></td>
<td>Diag 5-Factor</td>
<td>195.3</td>
<td>177.6</td>
<td>411.3</td>
<td>380.5</td>
<td>289.3</td>
</tr>
<tr>
<td></td>
<td>SDP 5-Factor</td>
<td>188.2</td>
<td>173.7</td>
<td>386.9</td>
<td>370.2</td>
<td>295.1</td>
</tr>
</tbody>
</table>

results”, *SIAM Journal on Optimization*, 8, 746–768.


Fitting Swaption Volatilities using Semidefinite Programming


