Using Hyperbolic Cross Points to Calibrate the Svensson Model to Swap Rates*

Emiliano Sánchez[a,b] and Eva Ferreira[b]

[a]LBBW† (Stuttgart, Germany)

[b]University of the Basque Country, UPV/EHU (Bilbao, Spain), Departamento de

May 15, 2014

Abstract

There are many methods to estimate the forward curve. One of the most popular is based on the Svensson model. However, different parameter estimation procedures can lead to very different results. First, which data (yields vs prices) are used to set the optimization problem is an important issue. On the other hand, the structure of the optimization problem can lead to obtain local optimum, which becomes a computational problem when standard packages are used. Here, we analyze the two optimization problems, either based on yields or on prices. For the first case, a previous interpolation method is needed and for the second case, the optimization problem becomes highly nonlinear. We discuss the desired properties in both situations. Based on that, we propose a new procedure for each case, based on hyperbolic cross points. The method is very easy to implement and does not rely on standard optimization packages. The results from the application part, where the forward curve is estimated from USD swap rates, shows that the direct estimation

*We acknowledge support from MEC (ECO2011-29268 mobility grant), and UPV/EHU (UFI 11/46 BETS). We also thank valuable comments from Prof. Dr. Stute from Giessen University and from participants of the Conference on "Current Topics on Mathematical Finance" held on April 2013 in Vienna. Corresponding author: emiliano.sanchezlopez@gmail.com

†Landesbank Baden-Wuerttemberg. The usual disclaimer with respect to the affiliation applies.

procedure from prices performs best, outperforming existing methods in terms of accuracy, stability and CPU time.

Keywords: Svensson Model; Numerical Optimization; Hyperbolic Cross Points.

1 Introduction and Notation

Many pricing-risk problems starts with the assumption that one has available a robust and reliable estimation of the interest rates term structure in at least one of its forms: i) the discount function ii) the zero-yield curve iii) the forward curve. Normally this is not the case since zero-coupon prices are not directly observable for the whole maturity spectrum. Therefore, they need to be estimated indirectly from coupon-bearing instruments.

A very common task is the computation of a forward rate curve implicit in capital markets, which is far for being a straightforward task. Provided with a sensible data set, one must decide to fit either a parametric, a semi-parametric model or simply interpolate between the knots. Different methods lead to different estimations and each of them present advantages and pitfalls that must be analyzed by the user. According with the objective, the desirable properties of the curve vary. Thus, practitioners can be aware of pricing and hedging and central banks of policy decisions.

Nevertheless, for all purpose there are some common features that any estimation of the term structure should consider: monotonicity of the discount factors that ensures a non-arbitrage positive forward rate curve, smoothness to avoid spurious oscillations, robustness to local perturbations, stability and goodness of fit to market data, or even reproducibility of market quotes for interpolation methods.

Related to the interpolation methods, the usual solutions often presented by software vendors are linear, natural cubic splines and exponential interpolators. These methods present serious drawbacks as lack of stability and robustness. While the linear spline provides good robustness to local perturbations, the result is not even differentiable and therefore, far for being smooth. On the other hand, cubic splines are smoother but very sensible to local perturbations, as we will show. The proposal by Andersen (2007) is a hyperbolic tension spline scheme with a non-parametric formulation. This model provides a combination between the linear and the natural cubic spline. The combination is fixed by the tension parameter selection. According to this, it recovers the linear spline model when the tension parameter is large enough, and the natural cubic spline when the parameter is zero. However, when LeFloc'h (2012) analyzes the difference between sequential and parallel perturbations, it is pointed out that the method does not produce a stable curve in many cases and that tension splines may change drastically when the tension parameter is updated. He instead proposes the use of a certain Hermite spline with derivatives weighted by an harmonic mean. For a detailed survey on interpolation methods, see Hagan and West (2006), where the authors also present their own proposal, named monotone convex method. On the other hand, for a discussion of the out-of-sample performance of these methods we refer to Bliss (1997), Bolder and Gusba (2002) and Hagan and West (2006). In summary, interpolation methods allows us to have the complete zero-yield curve that reproduce the observed data. The question that arises is which interpolation method is more suitable. In particular, the performance under local perturbations is an important issue.

Once the whole curve is available, a very popular method is the iterative extraction method suggested by Fama and Bliss (1987), which is widely known by practitioners as the "bootstrap" method. The bootstrap method in this context does nothing to do with the statistical technique and it is based on an idea that goes back to Caks (1977). Applying this technique, zero-coupon prices are recovered from zero-yields for the same maturities.

However, the interpolation methods usually provide curves hard to interpret and presenting spurious oscillations. Mainly, policy makers rely on methods that provide smoothness, stability and interpretability of the estimated curve. This makes the parametric models more adequate for their purpose. Models proposed by Nelson and Siegel (1987), Svensson (1994), Wiseman (1994) or Björk and Christensen (1999) belong to the this category. According to a survey conducted by the Bank for International Settlements, most central banks rely on parametric models for the approximation of interest rates used for macroeconomic policy analysis. A summary of different countries and the method used by their central banks is presented in Table (1).

Most central banks have adopted parametric methods as the Nelson and Siegel (1987) model and its extension, the Svensson (1994) or nonparametric method like the smoothing splines. The first two are nonlinear parametric methods, being the method provided by Svensson (1994) more general and nowadays considered as the parametric method for reference. It consists on a 6-dimensional parametric function which need to be calibrated against market data by means of a nonlinear optimizer. Nice features of the model are the interpretability of the parameters and the smoothness of the forward rates. A major drawback of this model, besides its acceptance by

Table 1: Estimation procedures by several central banks

Central Bank	**Estimation**	**Minimized Error**
Belgium	Svensson/Nelson-Siegel	Weighted prices
Canada	Exponential Spline	Weighted prices
Finland	Nelson-Siegel	Weighted prices
France	Svensson/Nelson-Siegel	Weighted prices
Germany	Svensson	Yields
Italy	Nelson-Siegel	Weighted prices
Japan	Smoothing spline	Prices
Norway	Svensson	Yields
Spain	Svensson	Weighted prices
Sweden	Smoothing splines	Yields
Switzerland	Smoothing splines	Yields

monetary policy makers, is the fact that two of the parameters are nonlinear and they destroy the convexity of the objective function. Typical least squares solvers like Gauss-Newton or Levenberg-Marquardt may fail and find local minima resulting in what Cairns (2004) calls "catastrophic jumps". That is, the value of a parameter changes abruptly from one calibration to another, loosing its interpretability. A study of different optimizers was done by Bliss, Ahi, Erdogan, and Şener (2012) where a discussion of several algorithms is presented. In particular, they show results for simulated annealing, particle swarm optimization and some benchmark optimization techniques like BGFS and Nelder-Mead.

The best algorithm to solve the minimization problem is still an open problem for both nonlinear proposals. In recent years, we can find different alternatives that also may depend on the data available.

In many situations, the estimation of the parameters is based on zero-yields, previously computed using some interpolation method. From the computationally point of view, this is a fast method to estimate parameters, since the problem related with the first four parameters become linear. Results of this two-step procedure can be found in Diebold and Li (2006), Gilli, Große, and Schumann (2010) and in Ferstl and Hayden (2010).

However, the results from the optimization problem will highly depend on the selected interpolation method used to obtain the zero-yields, and the discussion above should be taken into account.

An interesting alternative is to obtain the parameters directly from prices, without relying in any previous interpolation. Bolder and Stréliski (1999), Manousopoulos and Michalopoulos (2009), Ferstl and Hayden (2010) and Bliss, Ahi, Erdogan, and Şener (2012) present results from solving a direct optimization problem based on the observable data.

The results are diverse and presented in different schemes. For direct optimization, the function to be minimized is the sum of the squared differences between observed and estimated prices, absolute or relatively measured (mean absolute error (MAE), root mean square error (RMSE)). Bolder and Stréliski (1999) propose to divide the optimization problem into the linear and the nonlinear part. However, their results are not very promising in terms of accuracy RMSE. Manousopoulos and Michalopoulos (2009) compare several methods and recommend to use a global algorithm with a refinement based on the gradient and Hessian functions. Another direct proposal is made by Ferstl and Hayden (2010). Both works report results in terms of accuracy of prices estimation but they do not report CPU time or stability of parameters. The recent work by Bliss, Ahi, Erdogan, and Şener (2012) report detailed results in terms of CPU time and RMSE for 8 alternative algorithms. For the best algorithm, the CPU time goes from 100 to 170 seconds and the RMSE is around 0,20. Also, they report graphs to check the stability of the estimated parameters.

As Manousopoulos and Michalopoulos (2009) discusses, the selection of price or yield error minimization also takes into account other issues. He argues that if a yield curve is used as a monetary policy indicator, the yield error minimization should be used. On the other hand, if a yield curve is used for pricing, then the minimization of price errors is preferable.

Therefore, in this paper, we discuss both class of methods. For the first, based on zero-yields, we first discuss the advantages and pitfalls of different interpolation methods and select the Hermite spline procedure. In both cases, to estimate the Svensson model, the proposed procedure, as in Werner and Ferenczi (2006) and Ferstl and Hayden (2010), split the parameter set in the linear and the nonlinear part. To solve the optimization problem in the nonlinear part, we adapt the idea formulated in Novak and Ritter (2006), and propose an method to search in the hyperbolic cross points (HCP).

Within this framework, the objective of this paper is twofold. On the one hand, we discuss the robustness properties of three interpolation methods, linear, cubic and Hermite splines.

The results enable us to propose the Hermite spline method as the most adequate to be used for interpolation when the Svensson model is estimated using zero-yields. On the other hand, to solve the nonlinear part in the optimization problem, a method based on HPC is developed. The advantage is that the method is not iterative and, therefore, problems related with lack of convergence to convergence a local minimum are avoided.

Therefore, we show the performance of the two different approaches, using the proposed methodologies, and estimate the forward curve using USD swap data from Reuters.

The results show clearly that the direct method using prices outperforms the method based on zero-yields, in terms of accuracy. Therefore, whenever pricing is the objective, direct estimation is preferable. Although direct methods are computationally more expensive, we show how the use of hyperbolic cross points is very suitable for this problem and it makes the search very fast. In fact, the CPU time is quite smaller than those reported in Bolder and Stréliski (1999). In general, the results in terms of accuracy, stability of parameters and CPU time, related to previous works, enable us to propose this methodology as a very competitive procedure to estimate Svensson's parameters.

The rest of the paper is organized as follows. Section 2 discusses the robustness of three interpolation methods, linear, cubic and Hermite splines. Section 3 develops the new proposals to estimate Svensson parameters from the swap prices, either from yields or directly from prices. Section 4 provides the empirical results when the method is applied to USD swap data. Finally, Section 5 summarizes the main conclusions.

2 Robustness analysis of interpolation methods

Definitions and notation

The yield-to-maturity $Y(t,T)$ is the *internal rate of return* (IRR) of stream of cash flows, defined for any pair of time units (t,T), where $t \leq T$. The IRR is the compound rate equating an original disbursement against those cash flows. It could be seen as the root of a polynomial equation. Given a cash flow structure $(C_0,T_0),(C_1,T_1),\ldots,(C_N,T_N)$ the IRR satisfy the following equation: $-C_0 = \sum_{j=1}^{N} C_j \cdot (1+Y(t,T_j))^{-j}$. In order to exist a real positive root, the undiscounted sum of the cash flow, should be bigger than the original disbursement. It is going to be unique as long as there is one and only one change of sign in the cash flow structure. The IRR is dependent on a given cash flow structure and it cannot be used to valuate another streams with different structure. Therefore, in order to represent the *time value of money*, we need to determine IRRs from *zero-coupon bonds*.

The zero-coupon prices are defined as functions $P(t,T)$ such that for any pair (t,T), $P(t,T)$ gives the price in t of a promise to receive a monetary unit at tenor $T \geq t$, where T is measured in years. As a function of T, and under no arbitrage, $P(t,T)$ should be decreasing and monotonous. By definition: $P(t,t) \equiv 1$ and $P(t,\infty) \equiv 0$. Because zero-coupons (discount factors) are the building blocks of financial pricing we require that P should be a *continuous* function, in order to value *any* arbitrary cash flow.

In general, the zero-coupon rate is derived directly from zero-coupon instruments as

$$Z(t,T) = \left(\frac{1}{P(t,T)} - 1\right) \cdot \alpha(0,T), \tag{1}$$

$$Z(t,T) = \frac{1}{P(t,T)^{\alpha(0,T)}} - 1, \tag{2}$$

$$Z(t,T) = -\alpha(0,T)^{-1} \ln P(t,T), \tag{3}$$

for simple, compounded and continuous interest respectively. $\alpha(T_1,T_2)$ is a function that provides the yearly accrual factor, for $T_1 \leq T_2$. From that we obtain an *annuity* factor for a tenor structure $T_1 < T_2 < \cdots < T_N$ as

$$A_N(t) = \sum_{k=1}^{N} P(t,T_{k+1}) \cdot \alpha(T_k,T_{k+1}), \tag{4}$$

and the value of security paying the rate $Y(t,T_N)$ of a bond should be

$$V(t,N) = \frac{Y(t,T_N)}{p} \cdot A_N(t) + P(t,T_N) \cdot \alpha(t,T_N), \qquad N = n \cdot p, \tag{5}$$

where p is the annual coupon frequency of Y and n the time to maturity.

From zero-coupon prices [1], the forward rate $f(t,T_1,T_2)$, is calculated as

$$f(t,T_1,T_2) = \left(\frac{P(t,T_1)}{P(t,T_2)} - 1\right) \cdot \frac{1}{\alpha(T_1,T_2)}, \tag{6}$$

$$f(t,T_1,T_2) = \left(\frac{P(t,T_1)}{P(t,T_2)}\right)^{-\alpha(T_1,T_2)} - 1, \tag{7}$$

$$f(t,T_1,T_2) = \frac{\ln P(t,T_2) - \ln P(t,T_1)}{\alpha(T_1,T_2)}, \tag{8}$$

for simple, compounded and continuous interest rate.

[1] We use interchangeably zero-coupon prices or discount factors.

Another important characterization of the term structure are par swaps. By definition a par swap must value par at inception. To determinate the fixed rate which gives the swap a zero present value (i.e the value of the fixed leg equal to the value of the floating leg) we use

$$1 - P(t,T_k)\alpha(T_{k-1},T_k) = Y(t,T_k)\sum_{i=1}^{k-1} P(t,T_i)\cdot\alpha(T_{i-1},T_i), \tag{9}$$

where $T_0 = t$, for each $k = 1, 2, ..., N$ Since the rates $Y(t,T)$ are observable in markets, we can recover from swaps the discount factors $P(t,T)$. We need for that N equations to determine the entire tenor structure of discount factors $T_1 < T_2 < \cdots < T_N$. Namely, we need $N-1$ par swap rates in order to solve for $P(t,T_N)$. As we can't find quotes for all maturities, we need to restore to interpolation schemes (discussed in the next section) to complete the spectrum. The following diagrams represents (assuming market completeness) the discount factors and cash flow structure of par swaps respectively:

$$\mathbf{P} = \begin{pmatrix} \times \\ \times \\ \times \\ \times \\ \vdots \\ \times \end{pmatrix}, \qquad \mathbf{CF} = \begin{pmatrix} \times & & & & & \\ \times & \times & & & & \\ \times & \times & \times & & & \\ \times & \times & \times & \times & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \times & \times & \times & \times & \ldots & \times \end{pmatrix}.$$

Were the symbol $\times$ represents a non-zero entry. Equation (9) in matrix form can be expressed in a more compact form: $\mathbf{1} = \mathbf{CF}\cdot\mathbf{P}$. This can be solve for $\mathbf{P}$ by forward substitution using the following recursive relationship normally referred as "bootstrapping":

$$P(t,T_k) = \frac{1 - Y(t,T_k)\sum_{i=1}^{k-1} P(t,T_i)\cdot\alpha(T_{i-1},T_i)}{1 + Y(t,T_k)\cdot\alpha(T_{k-1},T_k)}, \tag{10}$$

Therefore, to apply the bootstrap methodology and recover the zero-coupon prices, it is necessary to have observations for $Y(t,T_k)$ at all maturities.

Since this is not usually the case, the missing data must be interpolated using an adequate technique. In general, interpolation can be performed on par swap rates, zero rates or discount factors. A wide range of interpolation methods based on splines are available and widely used (linear, exponential, cubic, tension and Hermite splines). For a review of interpolation theory we refer to the works by Lancaster and Salkauskas (1986), De Boor (2000) and Späth (1995). In the survey by Hagan and West (2006) several interpolation schemes are discussed and they outline some desirable features of a curve construction method; that is i) locality of the method

ii) smoothness of forward rates iii) monotonicity (arbitrage-free method) iv) locality of hedges and we add v) goodness of fit to the parametric methods. A more recent work by LeFloc'h (2012) argues that the convex monotone method proposed by Hagan and West (2006) does not always produce stable yield curves. Furthermore, he also suggests that, among the smoothest splines, tension splines present low stability when the tension parameter is updated. Finally he recommend the use of a certain type of (harmonic) Hermite splines.

A numerical comparison of interpolation splines

Within this framework, we suggest Hermite splines as the interpolation method of choice, since it provides smooth curves and better results than the tension splines whiteout the need to consider an extra parameter as with the tension splines. In order to provide more arguments for its use we also study another desirable property of an interpolation method, which is sensitivity to local perturbations.

The basic idea of testing the locality of perturbations is normally referred as the *par-point* approach. It consists on a 1bp perturbation of the par swap rate of each tenor and then compare to that of a parallel perturbation across all maturities. It is nothing but the difference of the (numerical) gradient and the sum of a "sequential" gradient. Ideally, this should be close to zero. Moreover, it is expected that a perturbation (on par swaps) on shorter maturities does not spill over to longer ones on the forward rates.

We now discuss the three splines to be considered and compare the perturbations produced. Since the analysis is valid for any t, this argument will be omitted in this part. So $Y(T) = Y(\cdot, T)$.

Linear Splines. It is computed applying linear interpolation to complete the maturity spectrum needed to "strip" the swap curve

$$Y(T) = Y(T_k) \cdot \lambda + Y(T_{k+1}) \cdot (1 - \lambda), \tag{11}$$

for each $T_k \leq T \leq T_{k+1}$, where λ is defined as $\alpha(T, T_{k+1})/\alpha(T_k, T_{k+1})$. We can make an assessment of the quality of rule in equation (11) with respect to the resulting *locality* and the *smoothness* of the forward rates, The first property is clearly retained since it uses only two grid points. As the derivatives on the grid points T_k and T_{k+1} need not to be the same with this interpolation rule, forward rates will present a jigsaw pattern due to changing signs of the derivatives. Figure (1) depicts the typical pattern of daily forward rates produced by a linear rule for USD swap rates as of 30.06.2013.

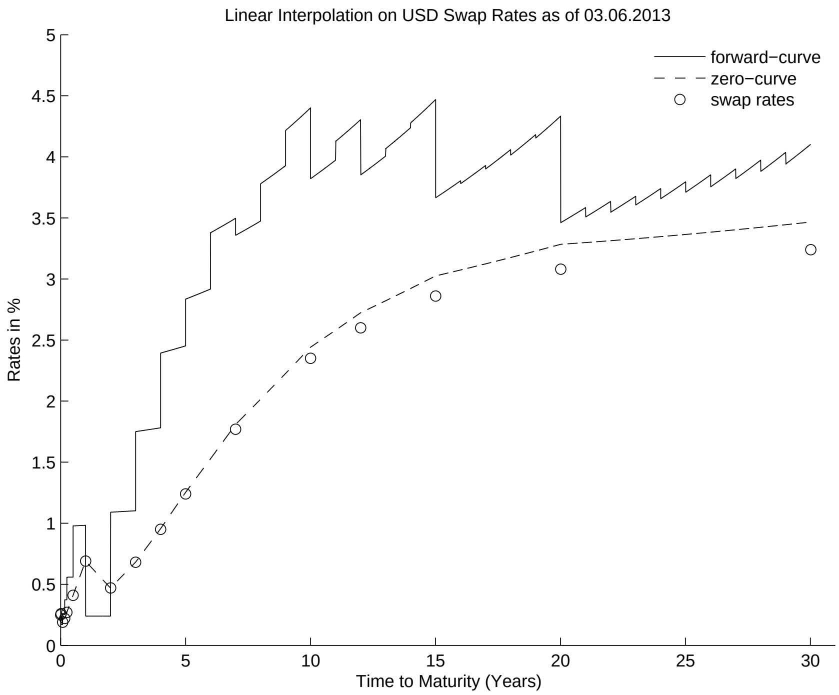

Figure 1: Linear Interpolation on Swap Rates. Zero-coupon rates (dotted line) were obtained stripping swap rates (circles) as a portfolio of zero-coupon bonds. The missing grid points were linearly interpolated. The 1-Day-forward curve (continuous line) presents the typical jigsaw pattern as a result of linear interpolation.

Cubic Splines. Another popular interpolation method is the natural cubic spline. The term natural is due to physical splines. The bending energy between two knots comes "naturally" to a minimum. As the bending energy depends on the curvature, dropping some terms with small contributions we can approximate it as the integral of the second derivative of the interpolant between two consecutive knots. If we minimize over the set of knots the sum of the defined integral we arrive to a unique solution: the natural cubic spline.

In order to build the cubic spline system of equations it is enough to impose the condition that the a cubic function is twice differentiable and that the second derivative is a linear spline interpolating the second derivatives at the knots, adding the boundary condition $Y''(T_0) = Y''(T_N) = 0$. From that, the system of equations becomes the tridiagonal system

$$Y(T) = \frac{\alpha(T, T_{k+1})^3}{6\alpha_k} \cdot Y''_k + \frac{\alpha(T_k, T)^3}{6\alpha_k} \cdot Y''_{k+1} + \alpha(T, T_{k+1}) \left(\frac{Y_k}{\alpha_k} - \frac{\alpha_k}{6} Y''_k \right) + \alpha(T_k, T) \left(\frac{Y_{k+1}}{\alpha_k} - \frac{\alpha_k}{6} Y''_{k+1} \right), \qquad \forall T \in [T_k, T_{k+1}). \tag{12}$$

The tridiagonal structure of the system (12) shows that locality is lost since now the interpolation scheme depends on the whole set of knots. The consequence of this is clearly as follows: an individual perturbation of a grid point will propagate (at a decreasing rate) to further maturities in the tenor structure producing spurious oscillations. Obviously the nice feature of the cubic splines is the smoothness of the forward rates.

To use the interpolation formula (12) for any T, we need to obtain the vector of second derivatives Y''. To do that, we write (12) for the knots, which lead to the matrix system

$$\mathbf{B}\mathbf{Y}'' = \mathbf{C}\mathbf{Y} + \mathbf{M}, \tag{13}$$

where $\mathbf{B}$ and $\mathbf{C}$ are of dimension $(N-2) \times (N-2)$, defined as follows:

Index	**B**	**C**
(k,k)	$\frac{\alpha_k+\alpha_{k+1}}{3}$	$-\left(\frac{1}{\alpha_k}+\frac{1}{\alpha_{k+1}}\right)$
$(k,k+1)$	$\frac{\alpha_{k+1}}{6}$	$\frac{1}{\alpha_{k+1}}$
$(k,k-1)$	$\frac{\alpha_k}{6}$	$\frac{1}{\alpha_k}$

$\mathbf{M} = \left(\frac{Y_1}{\alpha_1}, 0, \ldots, 0, \frac{Y_N}{\alpha_{N-1}}\right)^\top$ is a $(N-2)$-dimensional vector containing the boundary conditions. In order to solve the system stated in (13) we first compute the right hand side and then by some numerical scheme we invert matrix $\mathbf{B}$.

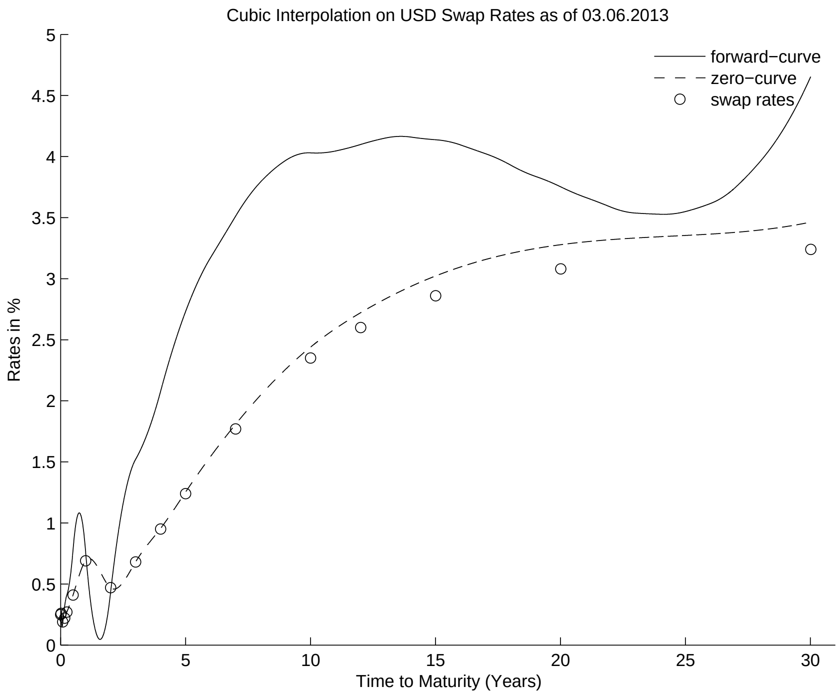

Figure 2: Cubic Interpolation on Swap Rates. Zero-coupon rates (dotted line) were obtained stripping swap rates (circles) as a portfolio of zero-coupon bonds. The missing grid points were with cubic splines interpolated. The 1-Day-forward curve (continuous line) "blows up" on the long end as a result of cubic interpolation.

Hermite Splines. Hermite splines preserve locality at the expense of loosing smoothness. Normally in Hermite interpolation not only function values need to be prescribed but also the slopes. Since we only have the ordinates, an easy way to circumvent this is to use finite difference to determine the slopes. One option is to use the so-called *Catmull-Rom* spline.

They are a family of cubic splines formulated in such a way that the slope (tangent) at each interior point T_k depends on a T_{k-1} and T_{k+1}

$$Y'(T_k) = \frac{Y_{k+1} - Y_{k-1}}{T_{k+1} - T_{k-1}}, \qquad \forall k = 2, \ldots, N-1.$$

On the boundary, forward and backward differences are being used instead for the first and last tangent respectively

$$Y'_1 = \frac{Y_2 - Y_1}{T_2 - T_1}, \qquad Y'_N = \frac{Y_N - Y_{N-1}}{T_N - T_{N-1}}.$$

Nevertheless this type of spline can be written in a derivative free way[2]

$$(14) \qquad Y(T_k) = \mathbf{D}_i(T_k)^\top \cdot \mathbf{A}_i \cdot \begin{pmatrix} Y_{k-1} & Y_k & Y_{k+1} & Y_{k+2} \end{pmatrix}^\top \qquad \forall T \in [T_k, T_{k+1}),$$

[2]For a derivation of the formula see Andersen and Piterbarg (2010, chapter 6).

where

$$\mathbf{A}_k = \begin{pmatrix} -\gamma_k & 2-\beta_k & -2+\gamma_k & \beta_k \\ 2\gamma_k & \beta_k - 3 & 3-2\gamma_k & -\beta_k \\ -\gamma_k & 0 & \gamma_k & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad k = 2, \dots, N-2. \tag{15}$$

for interior points where

$$\gamma_k = \frac{\alpha_k}{\alpha_k + \alpha_{k-1}}, \qquad \beta_k = \frac{\alpha_k}{\alpha_{k+1} + \alpha_k},$$

and for boundary points

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 1-\beta_1 & -1 & \beta_1 \\ 0 & -1+\beta_1 & 1 & \beta_1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \tag{16}$$

$$\mathbf{A}_N = \begin{pmatrix} -\gamma_{N-1} & 1 & -1+\gamma_{N-1} & 0 \\ 2\gamma_{N-1} & -2 & -2-2\gamma_{N-1} & 0 \\ -\gamma_{N-1} & 0 & \gamma_{N-1} & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{17}$$

Figures (1), (2) and (3) show three different forward curve shapes when different splines techniques where applied to interpolate between missing grid points. They illustrate that the Hermite spline is slightly less smooth than the natural cubic spline but still smooth enough to be a good competitor to be selected as a smooth interpolator.

Next, we proceed study the locality of the interpolation methods. The effect of a perturbation of 1bp, applied to the ten years swap rate is analyzed. To show this, we have considered sample of 1000 consecutive trading days (from June 2.009 to June 2.013) from the USD Swap (fix-for-floating) Market.

Table (2), (3) and (4) report the spills over the maturities around ten years and also over longer maturities, when a perturbation of 1bp is made on the 10Y Swap rate.

As expected, the linear spline presents the best performance related with robustness to local perturbation. However, the desirable property of smoothness makes cubic and Hermite splines more attractive for general users. From the results above, we see how the cubic spline is very

Table 2: Linear Spline Interpolation: Perturbation of 1bp on the 10Y Swap Rate. Impact on the Forward Curve as % of Change with Respect to the Original Curve.

	Mean	Median	MAD	Max	Min
8Y	0.69	0.62	0.13	-0.49	-1.26
9Y	0.82	0.77	0.13	-0.60	-1.45
10Y	0.95	0.90	0.13	-0.71	-1.70
11Y	1.19	1.14	0.16	2.03	0.78
12Y	1.42	1.36	0.18	2.38	0.93
13Y	0.01	0.01	0.00	0.01	-0.00
15Y	0.01	0.01	0.00	0.01	-0.00
20Y	0.00	0.00	0.00	0.01	-0.01

Table 3: Cubic Spline Interpolation: Perturbation of 1bp on the 10Y Swap Rate. Impact on the Forward Curve as % of Change with Respect to the Original Curve.

	Mean	Median	MAD	Max	Min
8Y	0.81	0.75	0.14	-0.57	-1.46
9Y	1.13	1.06	0.18	-0.81	-2.03
10Y	0.52	0.48	0.09	-0.38	-0.91
11Y	0.99	0.95	0.13	1.66	0.65
12Y	1.63	1.56	0.21	2.88	1.06
13Y	0.78	0.74	0.10	1.27	0.58
15Y	0.65	0.62	0.08	-0.46	-1.12
20Y	0.60	0.56	0.07	1.14	0.49

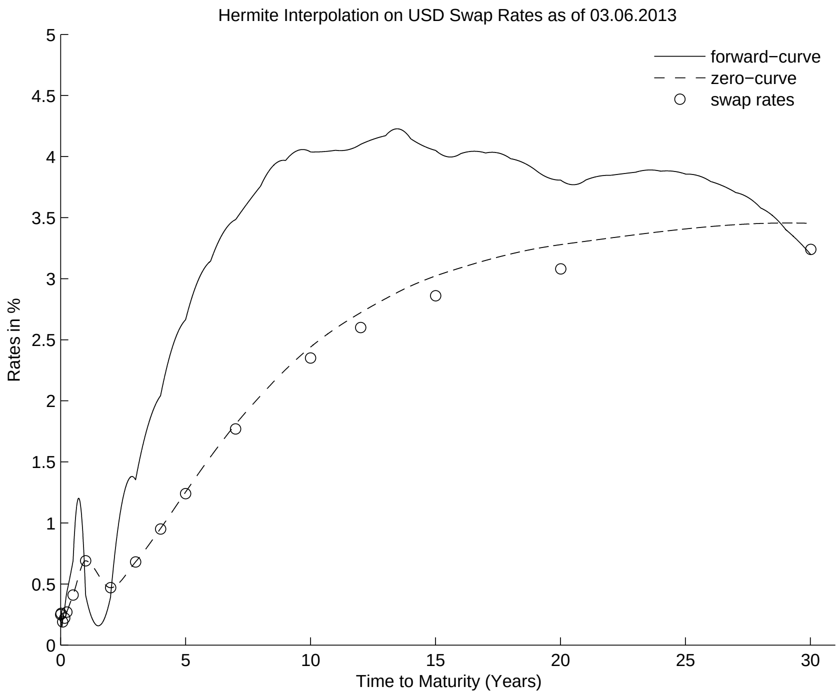

Figure 3: Hermite Interpolation on Swap Rates. Zero-coupon rates (dotted line) were obtained stripping swap rates (circles) as a portfolio of zero-coupon bonds. The missing grid points were with Hermite splines interpolated. The 1-Day-forward curve (continuous line) does not blows up like the cubic spline version at the expense of less smoothness.

sensible to local perturbations and, on the contrary, the Hermite splines are less sensitive to this local perturbation. Therefore, we select the Hermite splines as the interpolation method in the two-step procedure of next section.

3 Two procedures to estimate Svensson parameters

The Svensson model provides in a single-piece function a whole family of yield curves based on six parameters, that extends Nelson and Siegel model with two extra parameters for a better fit of the dynamics.

The specification of Svensson is

$$Z(T;\mathbf{x}) = x_1 + x_2\phi_1(T, x_5) + x_3\phi_2(T, x_5) + x_4\phi_3(T, x_6), \tag{18}$$

Table 4: Hermite Interpolation: Perturbation of 1bp on the 10Y Swap Rate. Impact on the Forward Curve as % of Change with Respect to the Original Curve.

	Mean	**Median**	**MAD**	**Max**	**Min**
8Y	0.83	0.77	0.16	-0.22	-1.65
9Y	1.19	1.11	0.20	-0.70	-2.08
10Y	0.43	0.40	0.13	0.33	-1.44
11Y	1.22	1.17	0.18	2.78	0.69
12Y	1.39	1.33	0.18	2.85	0.90
13Y	0.31	0.29	0.09	1.34	-0.00
15Y	0.17	0.16	0.05	0.01	-0.79
20Y	0.00	0.00	0.00	0.01	-0.00

where

$$\phi_1(T, x_5) = \left(\frac{1-\exp(-T/x_5)}{T/x_5}\right) \tag{19}$$

$$\phi_2(T, x_5) = \left(\frac{1-\exp(-T/x_5)}{T/x_5} - \exp(-T/x_5)\right) \tag{20}$$

$$\phi_3(T, x_6) = \left(\frac{1-\exp(-T/x_6)}{T/x_6} - \exp(-T/x_6)\right), \tag{21}$$

and $\mathbf{x} = (x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6)^{\mathrm{T}}$.

Motivated by monetary policy analysis, a nice feature of this model is the interpretability of the parameters. Table (5) shows the economic interpretation and the effect of the parameter on the curve.

Since its introduction, several calibration methodologies have been proposed. The first question to calibrate the model is which cost function to minimize. As mentioned in the introduction, the decision might depend on the financial objective and either the function can be written in term of yields or in term of prices. From the estimation point of view, the use of yields assumes that the zero-coupon yields are already given or previously provided by some interpolation method.

However, as discussed in previous section the bootstrapping of the par curve coupled with an interpolation technique has several drawbacks. On the other hand, when the cost function is written in terms of prices directly, we can avoid the need of previous interpolation.

Table 5: Parameters of the Svensson model

Parameter(s)	**Restriction**	**Interpretation**	**Linear**
x_1	≥ 0	Long term rate	Yes
x_2	n.a.	Rate of convergence (slope) of x_1	Yes
$x_1 + x_2$	≥ 0	Short term rate	Yes
x_3	n.a.	Size of the first hump(≥ 0) or dip(≤ 0)	Yes
x_4	n.a.	Size of the second hump(≥ 0) or dip(≤ 0)	Yes
x_5	≥ 0	Location of the first hump/dip	No
x_6	≥ 0	Location of the second hump/dip	No
$x_6 - x_5$	$\geq \kappa > 0$	To avoid interchange of x_5 and x_6	No

The second question is how to deal with the minimization problem. This is a non trivial issue since the model is highly nonlinear in two parameters. To overcome this fact, a common approach is to split the optimization problem in two steps, the nonlinear and the linear part.

To deal with this, Diebold and Li (2006) simply suggest, in the case of the Nelson and Siegel, to fix the nonlinear parameter and perform an OLS estimation. For Svensson's model this implies to fix two parameters, x_5 and x_6 in our notation.

The advantage of using yields is that the linear part is very easy and fast to solve. However, the interpolation problem arises and the nonlinear part still has to be solved. On the contrary, the direct minimization using prices avoids interpolation problems but has a highly nonlinear structure. Because of that, standard packages should not be used since they could provide wrong solutions or even none, as it will be discussed below.

The purpose of the rest of this section is to describe two procedures suitable for the two different optimization problems (yields or prices based). The procedures rely in combining the standard minimization solutions for the first four parameters with a algorithm of optimization for the other two, based on hyperbolic cross points (HCP).

As it will be shown, this combinations takes the best part of each algorithm. For the first four parameters, the dimension is too high to use the HCP method, but the optimization problem becomes a standard least squares problem. Therefore, the problem to estimate the first four parameters can be solved with standard methods. When the problem is set in terms of prices, the algorithm might require a modification of the Hessian matrix to ensure convergence. However, these algorithms present a lot of problems when we approach the second part, to estimate the two extra parameters. For this case, standard methods lead to very complicated expressions for the Hessian matrix and the risk of reaching the wrong points increases. On the other hand, the use of HCP for dimension two is computationally reasonable and, therefore, this method is used for this part. Therefore, the combination ensures the convergence of the method and the guarantee that the solution is the right one.

Hyperbolic cross points

Let $S = [-0.5, 0.5] \times [-0.5, 0.5] \subset \mathbb{R}^2$. We say that any pair $p = (p_1, p_2) \in S$ whose coordinates have a finite binary expansion is a hyperbolic cross point. By finite expansion we mean

$$\sum_{j=1}^{\ell} a_j 2^{-j}, \tag{22}$$

where $a_j \in \{0, 1\}$ and ℓ stands for the *level* of a HCP. The level of 0 is defined as 0. The point (0, 0.5) it is said to have level 1, and the point (0.375, 0) level 2.

Figure (4) depicts a sparse grid containing HCPs of level 8 on S. A straight forward way of reducing the number of points is eliminating the boundaries. In our problem we choose to do so also and coupled with the restriction that x_6 is bigger than x_5 we manage to halve the number of points increasing computational efficiency.

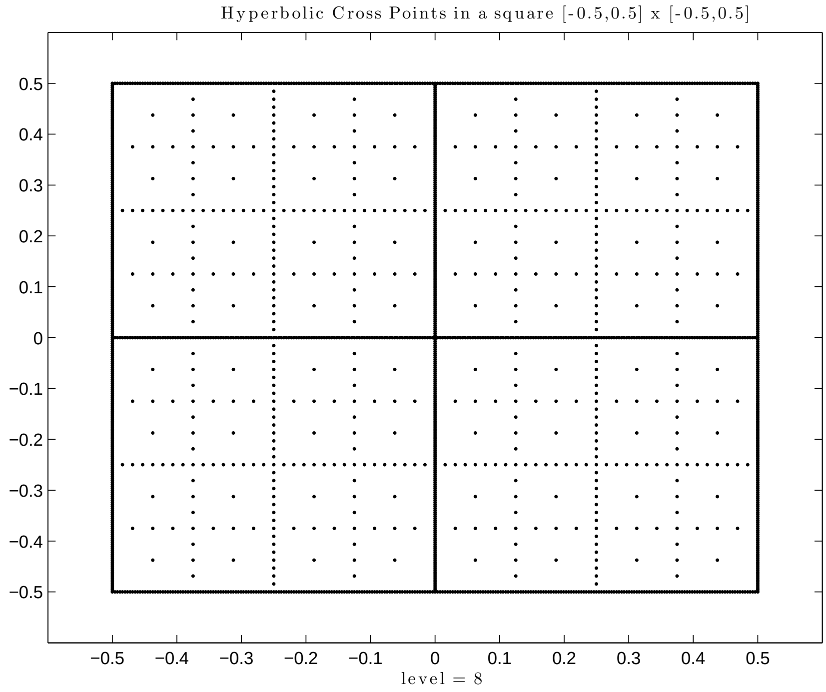

Figure 4: Hyperbolic cross points of level 8.

Optimization procedure based on yields

The parameters are the solution to the following optimization problem

$$\arg\min_{\mathbf{x}\in\mathbb{R}^{\mathbf{6}}} \quad \mathrm{F}(\mathbf{x}) = ||\mathbf{W}\cdot(\mathbf{Z}_{\mathrm{obs}} - \mathbf{Z}(\mathbf{x}))\,||_2^2, \tag{23}$$

where a previous interpolation method is needed to complete the zero-yields. Based on the results from section, we select the Hermite splines for this purpose.

As suggested by Bolder and Stréliski (1999) and Werner and Ferenczi (2006), we split the optimization problem into two parts, linear and nonlinear. The Svensson model requires restrictions on the parameters in order to hold its economic interpretation. As we will chose a splitting strategy to estimate, we divide the restrictions as well.

The inner problem constraints are stated through equation $\mathbf{C} \cdot \mathbf{x_{1:4}} \geq \mathbf{0}$, where

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{24}$$

For the linear part, we fix x_5, x_6 and write a matrix version of equation (18) as

$$\mathbf{Z}(\mathbf{x_{1:4}}) = \mathbf{A} \cdot \mathbf{x_{1:4}}. \tag{25}$$

where $\mathbf{x_{1:4}}$ denotes the column vector containing the first four parameters and

(26)

$$\mathbf{A} = \mathbf{A}(\mathbf{x_5}, \mathbf{x_6}) = \begin{pmatrix} 1 & \frac{1-\exp(-T_1/x_5)}{T_1/x_5} & \frac{1-\exp(-T_1/x_5)}{T_1/x_5} - \exp(-T_1/x_5) & \frac{1-\exp(-T_1/x_6)}{T_1/x_6} - \exp(-T_1/x_6) \\ 1 & \frac{1-\exp(-T_2/x_5)}{T_2/x_5} & \frac{1-\exp(-T_2/x_5)}{T_2/x_5} - \exp(-T_2/x_5) & \frac{1-\exp(-T_2/x_6)}{T_2/x_6} - \exp(-T_2/x_6) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{1-\exp(-T_n/x_5)}{T_n/x_5} & \frac{1-\exp(-T_n/x_5)}{T_n/x_5} - \exp(-T_n/x_5) & \frac{1-\exp(-T_n/x_6)}{T_n/x_6} - \exp(-T_n/x_6) \end{pmatrix}.$$

Given x_5 and x_6, the problem (23) becomes a linear least squares problem with linear inequality constraints in $\mathbb{R}^4$. Re-writing (23) as

$$\underset{\mathbf{C} \cdot \mathbf{x} \geq \mathbf{0}}{\arg\min} \quad ||\mathbf{W} \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{Z}_{\text{obs}})\,||_2^2, \tag{27}$$

where the weighting matrix $\mathbf{W}$ is diagonal with dimension n. The weighting is done taking into account the duration[3] of each theoretical swap: $w_i = d_i / \sum_{i=1}^{n} d_i$. $\mathbf{x}$ is redefined here as the sub-vector of the previous $\mathbf{x}$, considering only the first four components. To find the solution we apply a method based on a orthogonalization of matrix $\mathbf{A}$ and define the vector of observed prices as $\mathbf{b}$. For a detailed description see Lawson and Hanson (1974). Their technique is based on the transformation of the linear inequality constraints problem to a least distance programming (LDP) problem by previously solving a non-negative least square (NNLS).

[3]The duration of the ith swap is defined as $d_i = \sum_{j=1}^{n} (y_{i,j} * t_j * \delta_j)/100$, where y_i is the ith swap rate and δ_j the jth discount factor. It has the dimension of time.

The main steps can be described as follows:

Algorithm 3.1: LINEAR INEQUALITY PROBLEM (A, b, C)

comment: Compute an orthogonal decomposition

$[\mathbf{Q}, \mathbf{R}, \mathbf{K}] \leftarrow \texttt{qr}(\mathbf{A})$

comment: Transform into a least distance problem

$\mathbf{R} \leftarrow \mathbf{R}\texttt{(1:n,1:n)}$

$\mathbf{b} \leftarrow \mathbf{Q}' * \mathbf{b}$

$\mathbf{K} \leftarrow \mathbf{K}^{-\mathbf{1}}$

$\mathbf{G} \leftarrow \mathbf{C} * \mathbf{K} * \mathbf{R}^{-\mathbf{1}}$

$\mathbf{h} \leftarrow -\mathbf{G} * \mathbf{b}$

comment: Solve a least distance problem

$\mathbf{z}^* \leftarrow \texttt{LDP(G,h)}$

$\mathbf{x}^* \leftarrow \mathbf{K} * \mathbf{R}^{-\mathbf{1}} * (\mathbf{z}^* + \mathbf{b})$

$\mathbf{r} \leftarrow \mathbf{b} - \mathbf{A} * \mathbf{x}^*$

LDP and NNLS algorithms

For a detailed description of the *least distance problem* and the *nonnegative least squares* see Lawson and Hanson (1974). Here we sketch briefly the main steps of the first algorithm.

Algorithm 3.2: LINEAR DISTANCE PROBLEM (G, h)

$\mathbf{E} \leftarrow [\mathbf{G}', \mathbf{h}']$

$\mathbf{f} \leftarrow \texttt{[zeros(n,1);1]}$

comment: Solve a Non-Negative Least Squares problem

$\mathbf{u}^* \leftarrow \texttt{NNLS(E,f)}$

$\mathbf{r} \leftarrow \mathbf{f} - \mathbf{E} * \mathbf{u}^*$

As already discussed, for the estimation of the fifth and sixth components of $\mathbf{x}$ we need to solve a highly nonlinear problem, minimizing a non-convex function. The proposal here is to use a direct search on hyperbolic cross points (HCP), as defined by Novak and Ritter (2006).

Since the dimension of the optimization problem is two, in this case the HCP are defined on the plane. A set of HCP points is defined on a rectangle, which is becoming dense as the level of refinement is increasing.

In concrete, the procedure applies as follows

1. Select a level $\ell = 1, 2, ..$, which defines the degree of refinement for the HCP set
2. Consider all the sequence $\{a^i_j\}$, $i = 1, ..., \#(\ell)$, such that $a^i_j \in \{0, 1\}, j = 1, ..., \ell$. $\#(\ell)$ denotes the number of possible sequences.
3. Define $p_i = \Sigma_{j=1}^{\ell} a^i_j 2^{-j}$, $a^i_j \in \{0, 1\}$.
4. For each $i, k = 1, ..., \#\ell$, define four HCP as $(\pm p_i, \pm p_k)$. This builds a lattice of HCP in the square $[-.5, .5] \times [-.5, .5]$.
5. Define a box $[x_5^{\min}, x_5^{\max}] \times [x_6^{\min}, x_6^{\max}]$.
6. Map the HCP in the benchmark square domain to the defined box.
7. Compute (27) for each HCP satisfying the constraints and find the optimum.

Clearly, the higher the level ℓ, the finer the lattice of points. However, increasing the level exponentially increases the computational cost. As it will be shown in the next section, level 8 performs very well for the considered problem here.

Optimization procedure based on prices

In this case, we consider the prices available in the optimization problem.

Let $\mathbf{F} : \mathbb{R}^6 \to \mathbb{R}$ be a function mapping a six-dimensional vector of parameters to a scalar. The set of parameters $\mathbf{x}^*$ that makes $\mathbf{F}(\mathbf{x}^*)$ minimum is given by

$$\arg\min_{\mathbf{x}\in\mathbb{R}^6} \quad \mathrm{F}(\mathbf{x}) = ||\mathbf{W} \cdot (\mathbf{P}_{\text{obs}} - \mathbf{P})\,||_2^2, \tag{28}$$

where $\mathbf{W} \in \mathbb{R}^{m\times m}$ is a weighting matrix (as defined in the in the previous subsection), $\mathbf{P}_{\text{obs}} \in \mathbb{R}^m$ is a vector of observed prices and $\mathbf{P} \in \mathbb{R}^m$ is a vector of theoretical prices, namely, a discounted cash flow given by:

$$\mathbf{P} = \mathbf{CF} \cdot \exp(-\mathbf{Z}(\mathbf{x}) \cdot \mathbf{T}), \tag{29}$$

where $\mathbf{CF} \in \mathbb{R}^{m \times n}$ and $\mathbf{Z}$ and $\mathbf{T}$ have the proper dimensions. Recall that $\mathbf{CF}$ is a cash flow matrix with a number of rows equal to number of par rates or bonds and a number of columns equal to the number of cash flows of the par swap with largest maturity.

As before, we split the problem into two parts. However, for this case the first part, that optimizes for the first four parameters, now leads to a nonlinear least squares problem. Here, since the problem is 4-dimensional and the use of a global method would be computationally very costly. Therefore the complete proposal for the optimization based on prices will be:

1. A modified Newton method algorithm for the first part, using Gill-Murray factorization to avoid indefinite or semi-positive definite Hessians.

2. An active set strategy stated in Gill, Murray, and Wright (1981) to handle the restrictions.

3. A global search of HCP for the second part.

4 Empirical Application using Swap Rates for USD

We use a data set composed of fix-for-floating (1 month) interest rate swaps (IRS) denominated in USD. The data was collected on a daily basis starting from June 2.009 to June 2.013 (1000 observations) from Thomson Reuters. The day count convention is ACT/360, the coupon frequency is annual and the business day convention is the so called modified following rule. The maturities quoted and their corresponding Thomson Reuters Classification are detailed in table (6). The short end of the curve (tenor < 12 month) can be readily obtained from LIBOR rates. This rates are calculated by Reuters on behalf of the british banker's association[4] (BBA) and are published on a daily basis.

Figure (5) depicts the market evolution for USD swap rates for the following maturities: 1Y, 5Y, 30Y and the so called Fed rate. The difference between the short and long rates gives us a hint about the shape of the term structure. That is, "inverted" when 30Y < 1Y, roughly for the period 2.006-2.009 and "normal" from January 2.009 on. From the graphic we can also appreciate that although the key rate is a main driver of the shape of the term structure, it doe not has the final word when it comes to long rates (the correlation drops significantly as a function of time to maturity). Figure (6) shows three possible types of shapes of the term structure for the USD economy.

[4]See `http://www.bbalibor.com` for an explanation about the methodology.

Table 6: Par Swap Rates for USD

#	**Maturity**	**Reuters IC**
1	1Y	USD1YFSR=,LAST
2	2Y	USDAM3L2Y=,ASK
3	3Y	USDAM3L3Y=,ASK
4	4Y	USDAM3L4Y=,ASK
5	5Y	USDAM3L5Y=,ASK
6	7Y	USDAM3L7Y=,ASK
7	10Y	USDAM3L10Y=,ASK
8	12Y	USDAM3L12Y=,ASK
9	15Y	USDAM3L15Y=,ASK
10	20Y	USDAM3L20Y=,ASK
11	30Y	USDAM3L30Y=,ASK

In this section we present the results from calibrating the Svensson model to par swaps with the methods proposed in Section 2. Thus, we apply the procedure based on yields and the method based directly on prices. For the latter, to illustrate the convergence problems discussed in the previous section, we also report the results when a Gauss-Newton algorithm is applied instead of a direct search.

To discuss and compare the accuracy of the methods, two error measures are considered: the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE), for the difference between observed and estimated prices.

RMSE and MAE

To measure the goodness-of-fit of the algorithms two performance statistics were reported:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - \hat{P}_i)^2}. \tag{30}$$

Because we have calibrated to par swaps, by definition: $P_i \equiv 100,\ \forall\ i = 1, \ldots, n$.

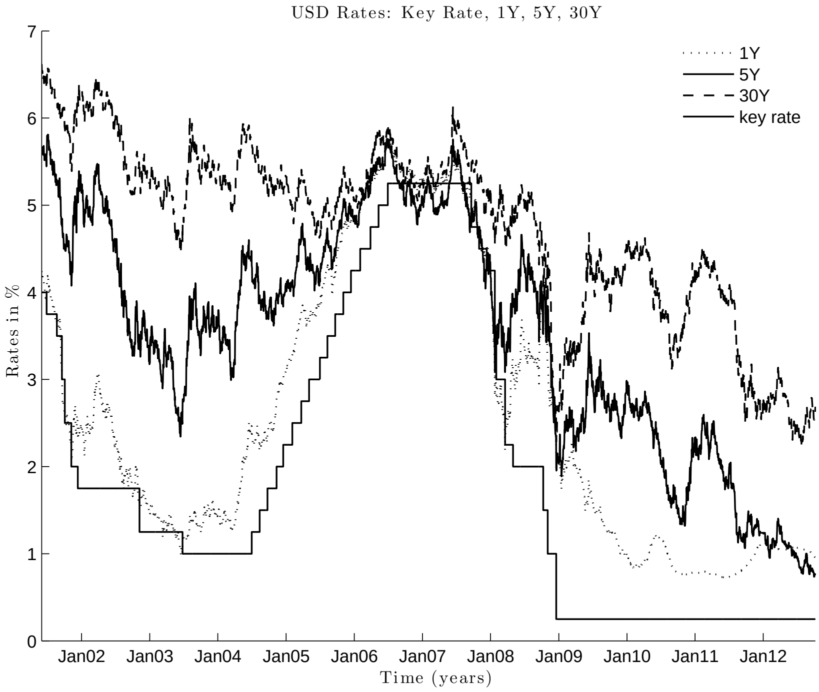

Figure 5: Market data for USD from June 2.001 to October 2.012.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |P_i - \hat{P}_i|. \tag{31}$$

It is useful to notice that RMSE $\leq \sqrt{n}$ MAE, in order to relate both quantities.

This comparison enables to show the importance of the interpolation when pricing is the objective. Also the average maximum and minimum fitted price are reported. To measure the computational effort of the algorithms, we report the CPU time.

To apply the HCP we have to decide the level of refinement, which means the amount of points in which the algorithm searches for the minimum. In Figure (7) we observe the convergence rate if we refine the mesh by increasing the level of the HCP. The improvement by successive refinement is not compensated beyond level 12. Moreover, if we take into account the CPU time reported on table (7) we conclude that between 8 and 10 the extra computational cost does not compensate the improvement on the RMSE. In table (7) RMSE and CPU are reported for each level of HCP. Clearly a trade-off should be made to have an acceptable level of accuracy and speed. Around level 8 one gets a reasonable accuracy at a considerable speed.

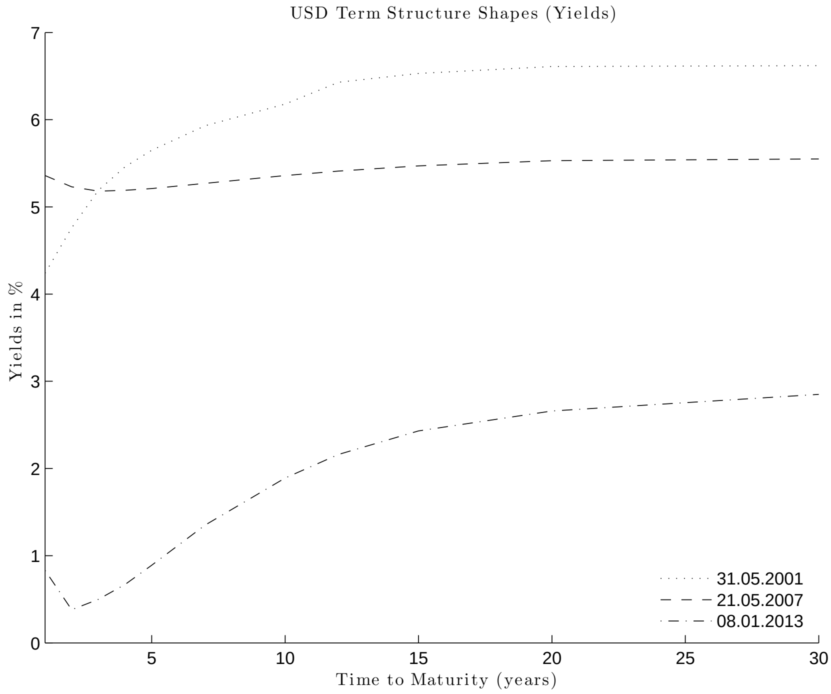

Figure 6: Three Types of Term Structures: Positive Slope, Flat and Combined.

Results of a comparison of three different approaches are summarized in Table (8). *Smoothing* denotes the procedure based on smoothing previously bootstrapped yields described in section 3. *Direct Estimation* denotes the procedure based on prices described in 3 and *G-N* for the procedure based on prices where Gauss-Newton is considered in six dimensions instead of splitting the problem and then perform direct search on HCPs.

The results reported show that the yield-based procedure provides the worst results in terms of fitted prices although it provides reasonable forward rates, see figure (8). Related to the Gauss-Newton procedure, we observe that this algorithm is very fast, with rate of convergence similar to the Newton's method. It gives the a fairly good approximation when the cost function is approximately zero but it is highly dependent on the starting point $\mathbf{x}^0$. Using the results from the previous section a run (unconstrained) of the algorithm provides good results when it converges. Unfortunately it converges in only 52.9 % of the cases. Therefore, although the results from Table (8) are the best in terms of accuracy and CPU time, they are only valid when the method converges, which is not the case in almost half of our sample (i.e more than 520 days).

Table 7: Convergence of the Smoothing Algorithm. RMSE and CPU time (in seconds). The algorithm was run varying the level of the HCP search for swap rates as of 03.06.2013.

Level	**RMSE**	**Seconds**	**# of Points**
4	1.1597	0.0317s	15
8	1.1392	0.1462s	505
12	1.1241	3.1185s	12245
16	1.1217	70.3840s	261537

Table 8: Accuracy and CPU time (average) for 1.000 estimations using three different methodologies: i) Smoothing bootstrapped rates ii) Direct estimation on prices: splitting the problem in two iii) Direct estimation on prices: solving a 6-dimensional least squares problem with Gauss-Newton.

	Smoothing	**Direct Estimation**	**G-N**
RMSE	1.5639	0.0142	0.0102
MAE	0.2421	0.0011	0.0011
Max	106.3494	100.2944	100.3191
Min	93.2342	99.8158	99.8943
CPU	0.1719	3.7596	0.0168

Table 9: Empirical proxies and their interpretation.

Parameter	**Proxy**	**Interpretation**
x_1	30Y Swap	Long Term Rate
x_2	10Y - 3M	Slope
x_3	2·(10Y + 3M)	Curvature

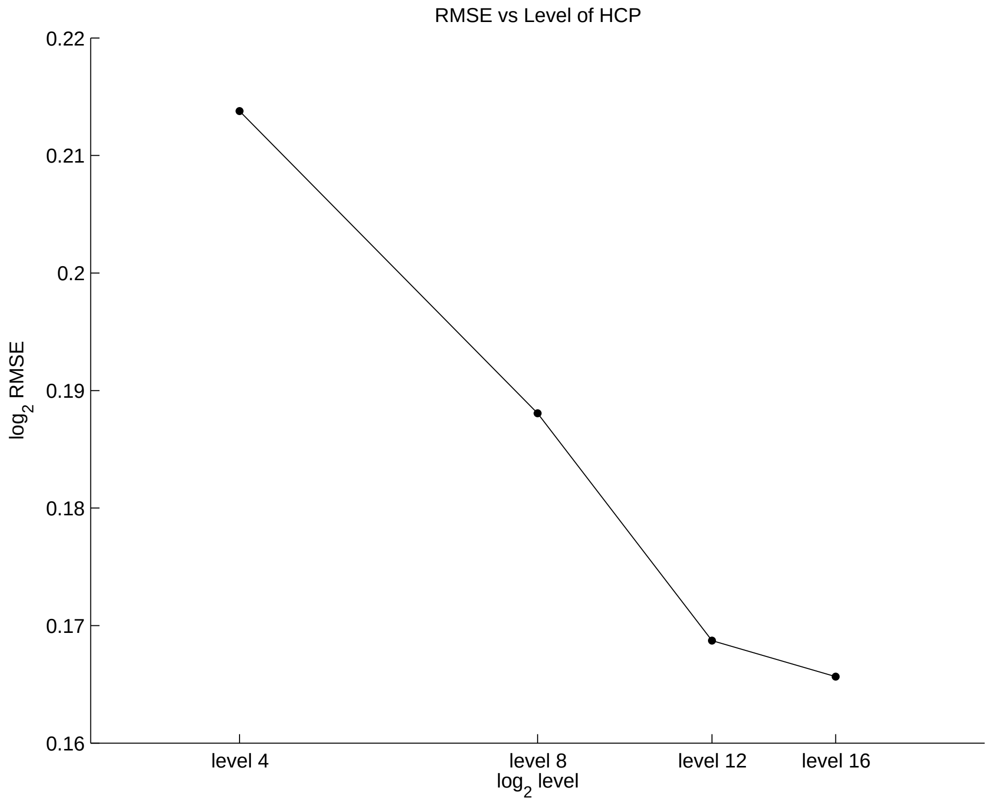

Figure 7: Convergence of the Smoothing Algorithm. The algorithm was run varying the level of the HCP search for swap rates as of 03.06.2013.

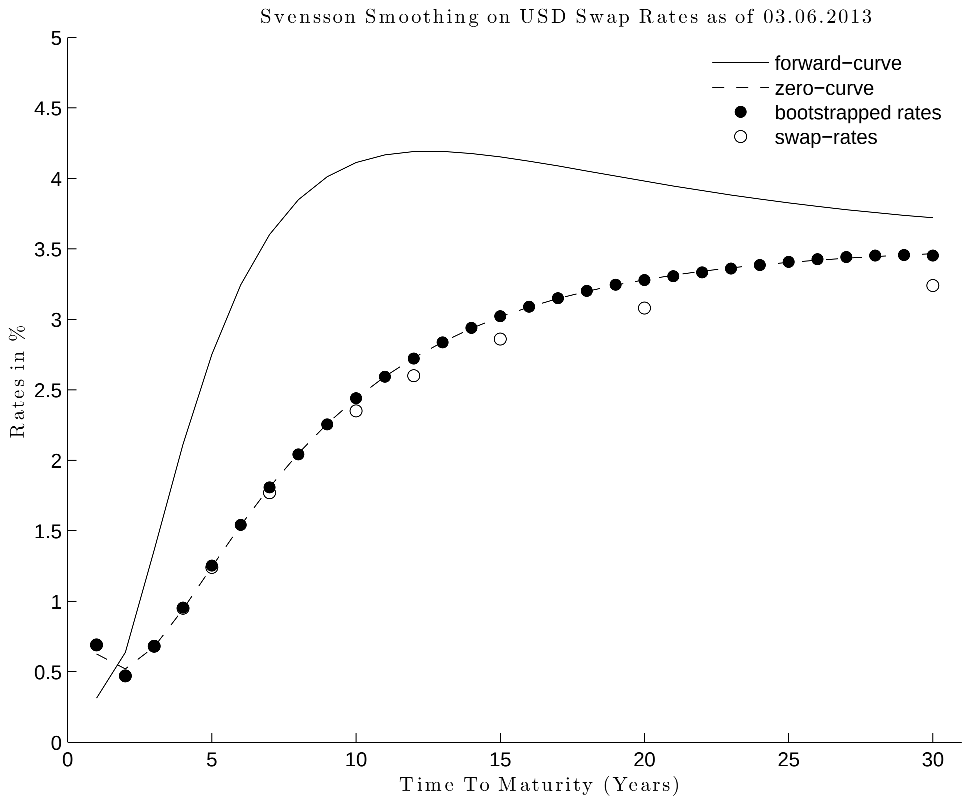

Figure 8: Svensson smoothing of previously bootstrapped (black circles) swap rates as of 03.06.2013. In a first step, Hermite interpolation method was needed in order to strip swap rates (white circles). Zero-rate curve (dotted line) and its respectively forward curve (continuous line) are also reported.

For the direct estimation procedure the six-dimensional non-linear problem was divided in two subproblems: i) A 4-dimensional problem solved by a Newton type method (with modified Hessian) and ii) A two-dimensional search made on half the hyperbolic cross points (level 6). This procedure provides the most reliable results, with high accuracy and a low computational cost, considering that it is a direct search method. In fact, the CPU time is much more lower than those reported for similar analysis (see Bliss, Ahi, Erdogan, and Şener (2012)).

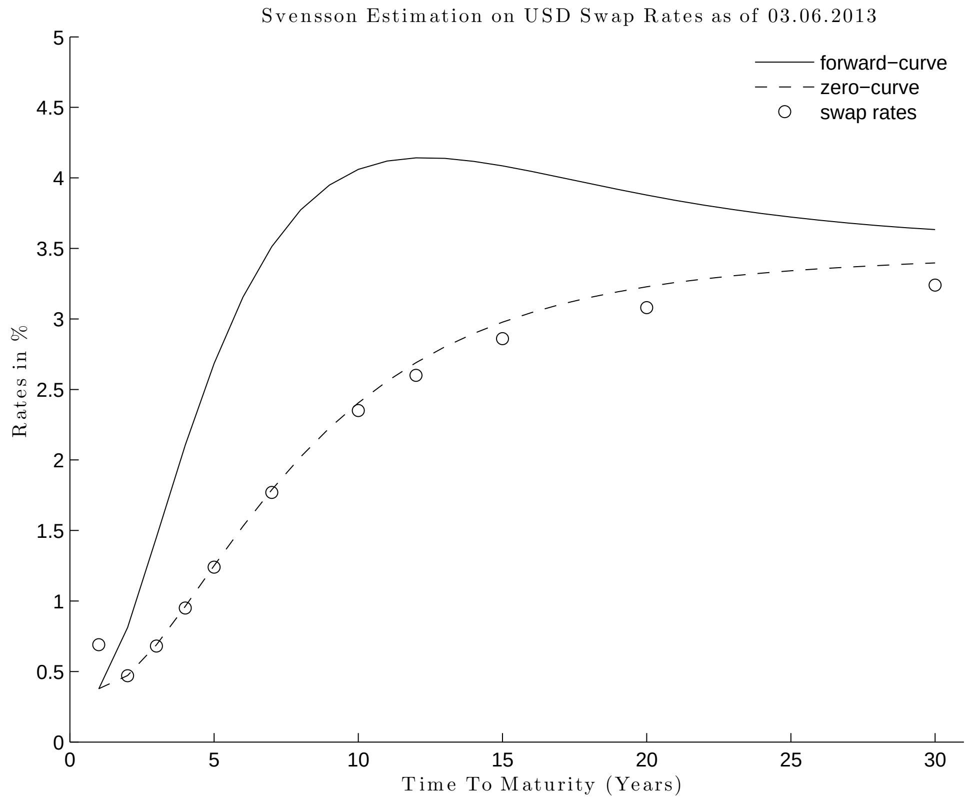

Figure 9: A calibration of the Svensson model directly to swap rates as of 03.06.2013 with the proposed algorithm.

Therefore, the results enables to presents the direct-estimation procedure as a very reliable method to estimate Svensson's parameters (see figure (9) for an example of the forward curve). Finally, we want to study the stability of the parameter estimates. In practice, the parameters must be estimated with previous information. Therefore, we report the results for one-day-ahead analysis, only for the case of direct estimation. Namely, we have calculated the fitted prices for each day with the parameters estimated using the data from the day before. The results are presented in Table (10). As observed from the table, the level of accuracy remains very high.

Table 10: Fitting prices one-day-ahead using the proposed direct estimation method.

RMSE	**MAE**	**Max**	**Min**
0.0498	0.0042	100.6149	99.4710

We compare the inter-temporal stability to detect implausible results of the optimal parameter set. In Figure 10 we can appreciate that the estimate of the long rate closely tracks the 30Y swap rate.

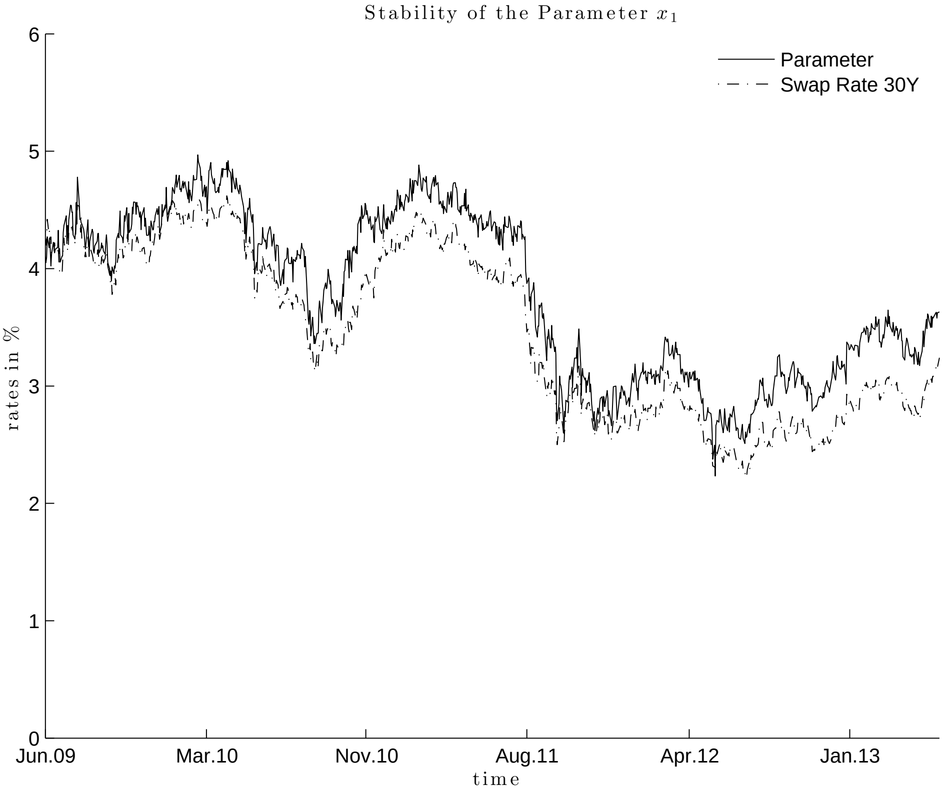

Figure 10: Stability of x_1. Represents the long term rate of the economy. Time series of x_1 (continuous) in % vs its proxy: 30Y swap rate (discontinuous).

5 Conclusions

The computational problem that arises in order to calibrate the parameters in Svensson model has been widely studied.

Our work is mostly related to Manousopoulos and Michalopoulos (2009), Gilli, Große, and Schumann (2010), Ferstl and Hayden (2010) and Bliss, Ahi, Erdogan, and Şener (2012). Our main objective has been to provide an estimation process that do not rely in any standard package, to avoid convergence problems. When the method is based on yields, the algorithm is very fast and the order of the errors is similar to those reported in Ferstl and Hayden (2010) and Bliss, Ahi, Erdogan, and Şener (2012).

However, the reported results show that the direct optimization method based on observed prices, without using any previous interpolation, provides the best results in terms of accuracy. Moreover, the short CPU time and the stability of parameters makes the direct-search based on hyperbolic cross points a very suitable procedure to fit the Svensson model.

The steps can be computed very easily and there are not problems due to slow or even lack of convergence in the optimization algorithm. Moreover, only four initial parameters are needed, which can be considered as the empirical proxies reported in table (9).

As commented, related works are Bliss, Ahi, Erdogan, and Şener (2012), Bolder and Stréliski (1999) and Ferstl and Hayden (2010). In terms of CPU, Bliss, Ahi, Erdogan, and Şener (2012) needs to interpolate to obtain the initial points, though the CPU time is bigger than the one reported. Bolder and Stréliski (1999) provides very long CPU times. Ferstl and Hayden (2010) does not report CPU time.

Comparing to these previous works, our direct-search procedure provides very short CPU times, it can be directly applied without any previous interpolation, the convergence is always guaranteed and it avoids the problem of convergence to a local minimum. Finally, the stability of parameters makes this procedure very useful for practitioners.

References

L. Andersen. Discount curve construction with tension splines. *Review of Derivatives Research*, 10(3):227–267, 2007.

L. B. Andersen and V. V. Piterbarg. *Interest Rate Modeling, Volume I: Foundations and Vanilla Models*. Atlantic Financial Press, 2010.

T. Björk and B. Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348, 1999.

R. Bliss. Testing term structure estimation methods. *Advances in Futures and Options Research*, 9:197–232, 1997.

R. Bliss, E. Ahi, G. Erdogan, and E. Şener. Robust estimation of the term structure. *Available at SSRN 2086840*, 2012.

D. Bolder and D. Stréliski. Yield curve modelling at the bank of canada. *Available at SSRN 1082845*, 1999.

D. J. Bolder and S. Gusba. Exponentials, polynomials, and fourier series: More yield curve modelling at the bank of canada. *Bank of Canada*, Working Paper 2002-29, 2002.

A. J. Cairns. *Interest rate models: an introduction*, volume 10. Princeton University Press, 2004.

J. Caks. The coupon effect on yield to maturity. *The Journal of Finance*, 32(1):103–115, 1977.

C. De Boor. *A practical guide to splines*. Springer (New York), 2000.

F. X. Diebold and C. Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130:337–364, 2006.

E. Fama and R. Bliss. The information in long-maturity forward rates. *The American Economic Review*, pages 680–692, 1987.

R. Ferstl and J. Hayden. Zero-coupon yield curve estimation with the package termstrc. *Journal of Statistical Software*, 36(1):1–34, 2010.

P. E. Gill, W. Murray, and M. H. Wright. *Practical optimization*. Academic press, 1981.

M. Gilli, S. Große, and E. Schumann. Calibrating the nelson–siegel–svensson model. *COMISEF Working Paper Series*, 2010.

P. S. Hagan and G. West. Interpolation methods for curve construction. *Applied Mathematical Finance*, 13(2):89–129, 2006.

P. Lancaster and K. Salkauskas. *Curve and surface fitting. An introduction.* London: Academic Press, 1986, 1986.

C. L. Lawson and R. J. Hanson. *Solving least squares problems*, volume 161. SIAM, 1974.

F. LeFloc'h. Stable interpolation for the yield curve. *Applied Mathematical Finance (July 11, 2012)*, 2012.

P. Manousopoulos and M. Michalopoulos. Comparison of non-linear optimization algorithms for yield curve estimation. *European Journal of Operational Research*, 192(2):594–602, 2009.

C. Nelson and A. Siegel. Parsimonious modeling of yield curves. *Journal of business*, pages 473–489, 1987.

E. Novak and K. Ritter. A new method for global optimization using hyperbolic cross points. *Working Paper*, 2006.

H. Späth. *One dimensional spline interpolation algorithms.* AK Peters, Ltd., 1995.

L. Svensson. Estimating and interpreting forward interest rates: Sweden 1992-1994, 1994.

R. Werner and I. Ferenczi. Calibration of the svensson model to simulated yield curves. In *MathFinance Workshop, Frankfurt*, 2006.

J. Wiseman. The exponential yield curve model. *European Fixed Income Research, Technical Specification*, 1994.