Yield Curve Estimation at the National Bank of Poland

Spline Based Methods, Curve Smoothing and Market Dynamics*

Estymacja krzywej dochodowości w Narodowym Banku Polskim

Metody splajnowe, wygładzanie krzywych, dynamika rynku

*Marek Marciniak***

Abstract

The paper presents methods used at the National Bank of Poland for yield curve estimation. A comparative analysis of parsimonious and polynomial models is conducted and the models evaluated according to several criteria. The results indicate that B-spline models stabilized with a variable roughness penalty (VRP) have best overall performance.

A new approach to implementing the roughness penalty is proposed. Instead of the commonly used continuous stabilizer a modification of the difference penalty used by Eilers and Marx (1996) is applied. A modified discrete stabilizer is compatible with the continuous penalty function while facilitating analytical solutions and reducing time of computation.

Finally, yield curve estimates for Poland are presented and their dynamics analyzed. A simple time-series based test is applied to evaluate the influence of unexpected events on the bond market.

Keywords: yield curve estimation, B-splines, curve smoothing, market dynamics

JEL: C14, E43, G12

Streszczenie

Artykuł przedstawia metody stosowane w Narodowym Banku Polskim do estymacji krzywych dochodowości. Autor przeprowadza analizę porównawczą modeli oszczędnych i wielomianowych oraz ocenia modele na podstawie kilku kryteriów. Wyniki wskazują na najwyższą niezawodność modeli B-splajnowych stabilizowanych za pomocą zmiennej sankcji krzywizny (VRP).

Autor proponuje nowy sposób implementacji sankcji krzywizny. Zamiast powszechnie używanego stabilizatora ciągłego stosuje modyfikację stabilizatora różnicowego wykorzystanego przez Eilersa i Marxa (1996). Własności zmodyfikowanego stabilizatora odpowiadają ciągłym funkcjom sankcji, a jego zaletą jest ułatwienie rozwiązań analitycznych i skrócenie czasu obliczeń.

W ostatniej części artykułu autor prezentuje oszacowania krzywej dochodowości dla Polski oraz przeprowadza analizę ich dynamiki. Do oceny wpływu nieoczekiwanych zdarzeń na rynek obligacji wykorzystany jest prosty test oparty na szeregach czasowych.

Słowa kluczowe: estymacja krzywej dochodowości, B-splajny, wygładzanie krzywych, dynamika rynku

* The views and opinions expressed herein are those of the author and do not necessarily reflect the position of the National Bank of Poland.

** National Bank of Poland and Warsaw School of Economics, Collegium of Economic Analysis

Introduction

The purpose of this paper is to present methods used at the National Bank of Poland for yield curve estimation and discuss in more detail some related issues. The topics include both a general overview of the basic concepts as well as a detailed discussion of some specific aspects of yield curve extraction and analysis. The scope of the paper is therefore fairly wide and can be attractive as much for a practitioner familiar with the field, seeking concrete solutions to some practical problems, as for a person with only little prior knowledge of the topic.

The first part of the paper provides a brief discussion of the most relevant aspects of yield curve modelling. First, several basic models for extraction of zero-coupon and forward rates from coupon bearing instruments are presented. Two classes of models are considered: parsimonious and polynomial based models. The first group includes the approaches developed by Nelson and Siegel (1987) and Svensson (1994), the second – a standard piecewise polynomial model by McCulloch (1971 and 1975) and its B-spline based versions by Fisher, Nychka and Zervos (1994) with smoothing methods put forward by Waggoner (1997) and Anderson, Sleath (2001). Technical details and practical aspects of model implementation are considered. Moreover, based on yield curve estimates for the Polish bond market a comparative analysis of the Svensson model and B-spline VRP model is conducted. The criteria for model evaluation are formulated and used to choose the most reliable approach.

A conclusion is drawn that the B-spline VRP models show best overall performance. They have virtually unlimited calibration possibilities and thus enable adequate smoothness adjustment while offering comparable or better goodness of fit than parsimonious models. A potential weakness of the B-spline based smoothed curves is considerable time expenditure for computation. The problem intensifies for more accurate models covering a wide range of maturities. It is related mainly to the curve stabilizer whose continuous form requires numerical integration over a squared second derivative of the function used to approximate the yield curve. Moreover, the yield curve function itself has to be evaluated recursively, which further extends the time of calculation.

A way to overcome the problem of a computationally demanding curve smoothing scheme was proposed by Eilers and Marx (1996) who used a discrete difference penalty instead of the continuous stabilizer. The authors, however, applied only a single scalar parameter to control the smoothness of the entire curve thus giving up the flexibility of the variable roughness penalty (VRP) approach. A method of merging both approaches in one model has so far not been discussed in much detail in the literature. This paper seeks to fill this gap.

An extension to the model by Eilers and Marx is proposed, which enables different level of smoothing in each segment of the yield curve. The method preserves virtually all advantages of the continuous approach by Anderson and Sleath while signifcantly simplifying notation and calculation. The result is a simple, flexible yield curve extraction model with relatively low computational requirements.

Finally, the paper presents the VRP model based yield curve estimates for Poland. An analysis of their basic properties and dynamics is conducted and the results discussed. A simple time-series based test is proposed to evaluate the influence of unexpected events and news releases on the bond market. The test proves a useful tool for evaluating yield curve reaction when data contain a considerable noise component and no accurate measures of market expectations are available. The behaviour of interest rates may be difficult to explain when an undefined (but large) number of factors translate into chaotic yield curve swings with no reasonable economic interpretation. The test constitutes a kind of a filter discriminating between meaningless and meaningful yield curve movements. By comparing the level of interest rates before and after an event evaluated as (tested to be) statistically significant it is possible to approximate the extent to which the event was anticipated and a surprise component. The method may be especially useful when market expectations are not known and have to be estimated from prices of financial instruments.

The paper is organized into three main sections. Section 1 contains a review of the basic parsimonious and polynomial yield curve models. The B-spline basis is introduced and applied to piecewise polynomial models. A modification of the standard smoothing mechanism for B-spline VRP models is proposed and discussed.

Section 2 presents results of the statistical analysis and comparison of the Svensson model and B-spline VRP models.

In Section 3 the term structure and dynamics of the bond market in Poland are analysed. A test for significance of exogenous events on the bond market is proposed and evaluated.

Finally, results of the paper are summarized and conclusions drawn.

1. Yield curve estimation

The main problem in yield curve analysis lies in the fact that interest rates are often not directly observable. This is the case for zero-coupon and forward interest rates, which are most interesting with respect to their information content. Recall that the most accessible measure of interest rate – yield to maturity – is actually dependent on the shape of the entire yield curve. Therefore, the information it conveys may reflect the influence of different – even not neighbouring – yield curve segments.

Extraction of zero-coupon rates from money market or zero-coupon instruments is straightforward or not necessary as these instruments are frequently quoted by their zero-coupon rate. In the case of the interest rate swaps (IRS) the calculations are simple as well. With IRS quotes available for virtually all maturities and their coupon payments due at equal time intervals it is straightforward to use bootstrapping methods to calculate zero-coupon rates.

Calculations are more complicated when the area of interest is the bond market. Theoretically, a basic bootstrap method would also be possible if only bond issues maturing at equal time intervals with regular coupon payment schemes were available. In practice this requirement is often not met. As a result, zero-coupon rates and implied forward rates cannot be extracted directly from prices of coupon bearing securities - they have to be estimated. The methods of yield curve estimation from bond prices are the main focus of this paper. Yield curve extraction and smoothing on the basis of zero-coupon instruments will be discussed only briefly.

Estimation of the term structure of interest rates requires the following issues to be taken into account:

1) Which market (market segment) is a proper source of data?

2) Which type of interest rates (zero-coupon rates, instantaneous rates, forward rates, or discount factors) should be directly estimated?

3) Which yield curve model and which functional form of the curve should be applied?

4) Which method is the best for parameter estimation?

In this paper some of the issues mentioned above are discussed. In each case the relevancy of a problem is indicated and a description of implications of making specific choices provided.

1.1. Yield curve models and their functional forms

1.1.1. Criteria for choosing a model

Yield curve models are characterized by the following features:

– goodness of fit (flexibility),

– smoothness,

– stability of results (robustness to changes in the data),

– numerical stability and time of computation.

Goodness of fit

Goodness of fit plays a crucial role if interest rate estimates are used for pricing purposes. In order to ensure adequate precision of estimates it is necessary to use a sufficiently flexible curve model. A flexible curve – precisely reflecting the current market situation may be a useful tool for e.g. identifying optimal maturity segments for a bond issuer or investor. If in a certain market segment a "hump" or a local minimum has emerged and the corresponding yields deviate from their fair values then it may be possible to purchase or sell (issue) a bond under favourable conditions and make extra profits.

Smoothness of the curve

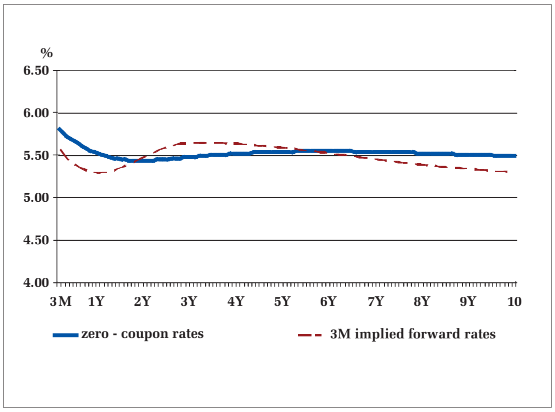
Smoothness of the yield curve may be a key feature if the purpose of analysis is not identification of mispriced securities but rather analysis of general properties and dynamics of the yield curve and extraction of implied expectations regarding inflation, interest rates or future state of the economy. Obviously, also for purposes of pricing securities an adequate degree of smoothness is necessary but to a lower degree.

A yield curve may be fitted to zero-coupon rates, forward rates, and instantaneous rates or to discount factors. If the curve reflects zero-coupon interest rates then its smoothness is crucial, especially if implied forward rates – expressing market expectations – will be extracted from it. This is due to the fact that any fluctuations of the zero-coupon curve increase in size when implied forward rates are calculated. The result may be a heavily fluctuating curve of forward rates (see Figure 1 for example).

Stability of the curve

The term stability is actually related to four different properties of a yield curve model. They are:

Figure 1. *Zero-coupon vs implied forward rates - curvature amplification*



Source: own calculations.

a) robustness to outliers,
b) robustness to minor changes in the data set,
c) (non)uniqueness of estimates,
d) numerical stability.

Robustness of yield curve estimates to mispriced securities in the data set is closely linked to its smoothness. An overly elastic curve will be in general too vulnerable to distortion caused by outliers. In practice it is very difficult to find the optimal relation between smoothness and flexibility. The problem becomes especially important in the case of non-liquid markets with only a small number of different bond series (issues). In such circumstances the most sensible solution is to employ appropriate filtering of data. This may be a non-trivial task ? in contrast to developed markets removing a mispriced security from the sample may be not possible if it has a high share in the market's total volume.

The second category of stability ? robustness to changes in the sample ? is closely related to the previous one, though it concerns somewhat different issues. In this case it means ensuring an adequate "stiffness" of the curve, so that removing, changing or adding instruments in the data set does not have a disproportionate effect on its shape. This property may be crucial if the model is supposed to be used for pricing new bond issues, or if it describes a small market, where a bond repayment or a new issue have a significant impact on the total size of the market. A frequently applied measure for stability of this kind is the sum of out of sample pricing errors. The minimization of its value is often used as a criterion for finding the optimal degree of smoothness of the curve (see Section 1.5 for more details).

The question of sensitivity to changes of the sample is especially important in the context of influence that movements in one segment of the yield curve have on another. For obvious reasons a change in the treasury bill rates should have only a limited impact on the long end of the yield curve. It is worth mentioning that the models analysed in this paper show significant differences in this respect.

The problem of non-uniqueness of estimates concerns models with non-linear error of fit function which have multiple local optima. In such cases (e.g. the Svensson model) repeating the calculations several times may yield randomly changing results, which reduces their credibility and hinders the analysis.

Low numerical stability is a serious problem characteristic for models with high correlation of the basis functions constituting a curve. A significant level of correlation causes very unfavourable properties of matrices used to represent a model. They manifest themselves in co-existence of extremely large and low eigenvalues, which potentially lead to very high rounding errors and errors resulting from exceeding the maximal value limit.

The trade-off between the flexibility (goodness of fit), smoothness and stability of the yield curve is the basic feature of any interest rate model. Excess flexibility leads to loss of smoothness (humps) and to a drastic fall in its stability. This phenomenon makes the appropriate calibration of a yield curve model with respect to all three criteria the central issue of yield curve modelling.

1.2. Parametric yield curve estimation

Parametric yield curve estimation methods were developed first. This was mainly due to the fact that it was possible to obtain results without a need for fast computers – analytically with the least squares method. Development of yield curve estimation methods began in the 70-ties, after publication of the seminal paper by J.H. McCulloch in 1971. McCulloch was one of the first authors to present a relatively comprehensive, general, theoretically justified and convenient approach to yield curve estimation.

1.2.1. Polynomial and piecewise polynomial models

Yield curve estimation methods are in general based on several common assumptions. First, it is assumed that the observable bond prices can be represented as a sum of discounted future cash flows (coupon and principal payments). Second, it is assumed that cash flows due at the same time are discounted with the same rate regardless of the time to maturity or coupon of the underlying security. This implies that for all bonds from the same credit class there exists one common curve of discount rates. Based on these assumptions a bond price equation can be formulated as follows:

$$p_i + ai_i = c_i \sum_{l=1}^{n} \delta(t_l) + V_i \delta(t_n) \tag{1}$$

where p_i denotes a clean price of the i-th bond, ai – accrued interest, c_i – coupon paid on the i-th bond, V_i – nominal value, $\delta(t_l)$ – value of the discount factor for cash flows due at t_l.

McCulloch (1971) suggested approximating the term structure of interest rates directly by the discount function $\delta(t)$, with the latter defined as a linear combination of k linearly independent[1] basis functions:

$$\delta(t) = 1 + \sum_{j=1}^{K} a_j f_j(t) \tag{2}$$

[1] In this context the linear independence means only a lack of *perfect linear dependence*.

where $f_j(t)$ denotes values of the j-th basis function for t. In order to obtain economically interpretable results, the discount curve should be positive and monotonously decreasing within its domain. This is equivalent to the condition of positive implied forward rates.

In equation (1) the price of a bond is defined as a linear combination of discount factors (2), the latter being a linear function as well. With a combination of these equations it is possible to represent the bond price as a linear combination of the relevant parameters:

$$y_i = \sum_{j=1}^{K} a_j x_{ij} \qquad (3)$$

where:

$$y_i = p_i + ai_i - nc_i - V_i$$

$$x_{ij} = c_i \sum_{l=1}^{n} f_j(l) + V_i f_j(n) \qquad (4)$$

The above formulation of the estimation problem facilitates the application of the least squares method to calculate the parameters a_j. The choice made by McCulloch to directly estimate the discount curve was motivated mainly by this convenient feature. Zero-coupon and implied forward rates can then easily be derived from the estimated discount curve.

Single polynomial

The initial choice of a basis function in equation (2) was a polynomial of degree k, formally: for $j= 1,\ldots,k$ where ($k >= 3$). However, the simplicity of this solution is achieved at the cost of several unfavourable properties.

In the case of a single polynomial it is very difficult to obtain a desirable goodness of fit with an adequate level of smoothness and stability. The properties of the curve depend on a degree of the polynomial fitted. A higher degree increases the accuracy of fit at the cost of lower stability and smoothness, especially at the long end of the curve. As a result, for high degree polynomials the discount function has a tendency to large swings between observations for longer maturities. It is usually accompanied by a distortion of curve monotonicity. A direct consequence of discount function instability is a chaotic shape of the forward rate curve.

The instability at the long end of the discount curve stems from the mathematical properties of its basis functions. First, the standard polynomial basis exhibits very high co-linearity, which causes serious numerical difficulties. Second, the instability at the long end of the curve is additionally amplified by high values of the basis functions $f_j(t)$ and their derivatives for large t values, which may cause strong fluctuations of the curve for longer maturities. Third, very high basis function values force the corresponding parameters to take very low values (in the order of the fourth decimal place). The resulting rounding errors which arise during numerical calculations in combination with high basis function correlations may translate into a relatively low numerical stability and precision of the estimates.

Modelling discount functions using only a single polynomial has yet another significant drawback. The curve constructed in this way disregards the distribution of bonds with respect to their maturities. A consequence may be that when applying a simple OLS method to fit the curve, segments of the yield curve with high concentration of observations (usually the short end of the curve) will dominate the overall fit at the cost of medium and long maturities. An attempt to solve this problem through increasing the degree of the polynomial does not yield satisfactory results (for previously described reasons). One of feasible methods of reducing this problem is to adjust the distribution of bonds by applying different weightings to them. In order to ensure equal influence of each yield curve segment on the final fit, the weightings should be adversely proportional to the number of bonds in each segment.

Piecewise polynomials

A reasonable solution to some of the aforementioned problems is to approximate the discount curve using a piecewise polynomial function. Models based on piecewise-polynomial splines offer a convenient solution to many of the above mentioned problems. In contrast to single polynomial curves, splines facilitate reduction of the degree of polynomials. In effect, polynomial splines provide greater stability at the long end of the curve. The method was first proposed by McCulloch for second order splines (1971) and subsequently extended to cubic splines (1975).

In the McCulloch's method the maturity domain is divided into a number of segments with predefined location and fixed length. In each of them the curve is defined as a polynomial of order three. The points joining the adjacent curve segments are described as knots or knot points. In order to ensure that the spline curve is continuous and twice differentiable it is necessary to impose equality constraints on the function values and their derivatives at the knot points. A curve fulfilling these criteria can have the following form:

$$\delta(t) = 1 + a_1 t + a_2 t^2 + a_3 t^3 + \sum_{j=4}^{k} a_j (t - t_j)^3 \phi_j(t) \qquad (5)$$

where $\phi_j(t) = 1$ for , and $\phi_j(t) = 0$ otherwise[2].

[2] Based on Bekdache, Baum (1997).

The discount function remains a linear combination of the parameters a_i. Price of a bond as defined by (1) is linearly dependent on the discount factors. Combining both formulas enables representation of a bond price as a linear combination of the parameters and facilitates the application of OLS estimation.

An important issue is setting the number and positioning of knots. McCulloch suggests setting their number as an integer closest to a square root of the number of instruments used for estimation. To ensure equal influence of all instruments on the final fit of the curve, the author recommended locating the knots so that there are an equal number of bonds between each two adjacent knots.

Flexibility of the spline-based curves resulting from a full discretion in setting the number and location of knots allows achieving a reasonable fit of virtually any curve. Nevertheless, it has to be kept in mind that increasing the number of knots improves the fit, but at the cost of lower smoothness and stability. An overly elastic curve can lead to a loss of monotonicity of the discount function.

A problem which has not been resolved by substituting a single polynomial by a piecewise polynomial function is the *co-linearity of the basis functions*. Actually, it has been reduced to some extent, but remains at the level of approx. *50-90%*[3], which is very high. In combination with very low parameter values this may result in significant errors of estimates and a low numerical stability of the results.

1.2.2. Parsimonious models

The main weakness of the models discussed so far is their insufficient smoothness and stability. Implied forward rates calculated from polynomial-based discount functions have a tendency to highly oscillatory term structure. Moreover, they do not have any asymptotic convergence properties and – consequently – for maturities beyond the observed spectrum usually take totally unrealistic and non-interpretable (often negative) values. This is directly related to the asymptotic properties of polynomials, which always have infinite boundaries. An immediate consequence is a low usefulness of the standard polynomial models in empirical research, despite their flexibility and ease of calculation.

A solution to these problems is provided by parsimonious yield curve models. The most known model of this class was put forward by Nelson and Siegel (1987). The authors assumed that the instantaneous interest rates are generated by a stochastic process, which can be expressed by a differential equation of order two. A formula obtained by solving this equation describes the term structure of instantaneous interest rates and has the following form:

$$f(t) = \beta_0 + \beta_1 \exp\left(-\frac{t}{\tau_1}\right) + \beta_2 \left[\left(\frac{t}{\tau_1}\right)\exp\left(-\frac{t}{\tau_1}\right)\right] \qquad (6)$$

where $f(t)$ denotes the *instantaneous* implied forward rate for an infinitesimally short period starting at t.

Zero-coupon rates can be calculated by averaging the corresponding instantaneous rates:

$$R(t) = \frac{1}{t}\int_0^t f(x)dx \qquad (7)$$

which gives:

$$R(t) = \beta_0 + (\beta_1 + \beta_2)\left[1 - \exp\left(-\frac{t}{\tau_1}\right)\right] \Big/ \left(\frac{t}{\tau_1}\right) - \beta_2 \exp\left(-\frac{t}{\tau_1}\right)$$

The curves of forward and zero-coupon rates are functions of four parameters: $\beta_0, \beta_1, \beta_2, \tau_1$. They can take several basic shapes: monotonous increasing or decreasing, humped, or S-shaped.

By directly modelling the curve of forward rates and applying very smooth functions, Nelson and Siegel avoided the problem of forward curve instability, characteristic for polynomial spline models. Beside their stability and smoothness Nelson-Siegel curves possess another convenient feature. Namely, they have constant asymptotic limits. A direct consequence of this is that the first derivative converges to zero as $t \to \infty$. As a result, the curve gradually becomes flat for longest maturities. This is a desirable property as it reflects lack of sufficient information to differentiate between forward rates in different segments with very long maturities.

Another significant advantage of the Nelson-Siegel model is a straightforward interpretation of its parameters. A direct result of

$$\lim_{t\to\infty} R(t) = \beta_0 \qquad (8)$$

is that the value of β_0 should correspond to zero-coupon rates for ultra long maturities.

At the short end of the curve we have:

$$\lim_{t\to 0} R(t) = \beta_0 + \beta_1 \qquad (9)$$

which implies that the sum of parameter values β_0 and β_1 should be equal to the level of the shortest interest rates.

On the basis of the zero-coupon curve (7) it is easy to formulate the discount function:

$$\delta(t) = \exp(-R(t) \cdot t) \qquad (10)$$

Note that it is non-linear in parameters. Therefore it is not possible to express a price of a

[3] Depending on the set of basis functions and their domain – author's calculations.

coupon bearing bond as a linear combination of parameters. In consequence, estimation of a yield curve with this method requires numerical computation, which is a clear disadvantage in comparison to polynomial models.

A significant weakness of the Nelson-Siegel model, resulting from its low elasticity, is goodness of fit lower than in the case of polynomial models. When the curve is fitted to an irregular set of data points this can result in relatively large deviations of model values from actually observed rates.

A method to overcome this obstacle was proposed by Svensson (1994), who extended the Nelson-Siegel model by adding a new component to the equation (6), thus introducing a new hump in the curve:

$$f(t) = \beta_0 + \beta_1 \exp\left(-\frac{t}{\tau_1}\right) + \beta_2 \left[\left(\frac{t}{\tau_1}\right)\exp\left(-\frac{t}{\tau_1}\right)\right] + \\ + \beta_3 \left[\left(\frac{t}{\tau_2}\right)\exp\left(-\frac{t}{\tau_2}\right)\right] \qquad (11)$$

The corresponding curve of zero-coupon rates has the following form:

$$R(t) = \beta_0 + (\beta_1 + \beta_2)\left[1-\exp\left(-\frac{t}{\tau_1}\right)\right] \Big/ \left(\frac{t}{\tau_1}\right) - \beta_2 \exp\left(-\frac{t}{\tau_1}\right) + \\ + \beta_3 \left[\left[1-\exp\left(-\frac{t}{\tau_2}\right)\right] \Big/ \left(\frac{t}{\tau_2}\right) - \beta_2 \exp\left(-\frac{t}{\tau_2}\right)\right] \qquad (12)$$

Two additional parameters enabled a significant increase in the elasticity of the curve, without causing any significant loss of the smoothness and asymptotic properties of the model.

The extended Nelson-Siegel model by Svensson offering a satisfactory precision of fit and a smooth shape of implied forward curve became very popular in the middle 90-ties. A considerable number of central banks worldwide have since then been using it for estimation of the term structure of zero-coupon and forward rates[4]. Nevertheless, the model has a number of weaknesses, e.g. a limited ability to fit irregular yield curve shapes, a tendency to take extreme values at the short end, and a relatively strong co-dependence of estimates in different – even non-neighbouring – segments of the yield curve.

1.3. Choosing the objective function and functional form of the curve

1.3.1. Choosing a type of interest rates for direct estimation

In the case of parsimonious models the functional form of a yield curve is predefined. Consequently, the choice to estimate directly instantaneous rates, zero-coupon rates or discount factors is only a matter of computational convenience.

Piecewise polynomial curves are not attached by definition to any kind of interest rates and can be used to approximate any of them. However, as the fitted curve inherits all properties (smoothness, stability, flexibility, etc.) from the underlying piecewise polynomial the choice which interest rates are to be directly estimated becomes crucial. The basic criteria which must be considered are: quality of estimates and ease of computation.

In order to make an optimal choice it would be advisable to try each of the feasible options. This was done e.g. by Bolder and Gusba (2002) for the Canadian bond market[5]. Applying a number of criteria to evaluate the models authors conclude that the best results are achieved when the curves of (in the rank order) zero-coupon rates, discount factors and forward rates are estimated directly.

The results for the discount curve are much like for the zero-coupon rates. Differences emerge mainly at the short end of the curve. Direct approximation of forward rates yields unsatisfactory results. Moreover, the model implementation in this case is much more complicated and requires more time for computation.

Due to the best quality of results for the purposes of this paper direct approximation of zero-coupon rates was chosen.

1.3.2. The objective function

In the case of yield curve modelling the objective can take two forms: [1] minimize a sum of squared price errors or [2] squared yield errors. Despite the mutual uniqueness of the price-to-yield transformation estimates obtained for the above objectives differ significantly. This is due to the non-linear relation between price and yield, and the differences in yield elasticity of the price in different maturity segments. In practice modified duration is used as a measure of this elasticity. Yields of bonds with longer maturities are in general more sensitive to changes in prices. Therefore in the case of the objective function [1] with equally weighted price errors, long term bonds are actually given highest weightings. The result is less accuracy at the short end of the curve[6]. The problem can be solved in two ways: by applying higher weightings to short term instruments[7], or by choosing the objective function [2], i.e. by minimizing squared yield errors. Both approaches should in general yield similar results. However, the implementation of the second one entails more computational complexity and time expense for

[4] See e.g. Csajbok (1998).

[5] Other publications containing this kind of analysis include: Bliss (1996), Bekdache, Baum (1997).

[6] See e.g. Svensson (1995).

[7] See e.g. Csajbok (1998).

estimation. For this reason weightings proportional to the reciprocal of the duration are applied in this paper.

1.4. B-spline models

1.4.1. B-spline basis

The problem of co-linearity and the parallel problem of numerical instability characteristic for standard piecewise polynomials can be resolved by applying a B-spline basis. B-spline basis allows to represent a piecewise polynomial as a linear combination of relatively simple and only lowly correlated elements called basic splines (hereafter abbreviated to B-splines), which facilitate some elementary mathematical operations like differentiation and integration[8].

A single B-spline is a combination of a number of standard polynomials. A cubic B-spline (most often used) is based on at least five knot points which define four adjacent intervals for which only the B-spline is positive in value. In the remaining area of the domain its value equals 0. The B-spline basis is constructed from a series of B-splines based on a common set of knot points.

In order to present the technical details of the basis construction in a more comprehensible way let us suppose that we want to construct the basis for a cubic piecewise polynomial with knots located at $\{k_0,\ldots,k_N\}$. Let us call them primary knots (as opposed to the auxiliary knots), as they are located within the domain of the curve fitted. In each of the intervals between any two adjacent knots the value to the curve is defined as a linear combination of all B-splines, which can be constructed on the given knot set. From the fact that a cubic B-spline takes non-zero values for exactly four adjacent intervals it follows that for e.g. the first interval: $\in< k_0,k_1 >$ it is necessary to use knots starting at (taking positive values from) k_{-3},k_{-2},k_{-1},k_0 . So the first three B-splines begin with auxiliary knots, which are not contained in the domain of the curve. An analogous situation occurs in the last interval $x\in< k_{N-1},k_N >$. The B-splines which constitute the basis within this interval begin at $k_{N-4},k_{N-3},k_{N-2},k_{N-1}$, and end at $k_{N-4},k_{N-3},k_{N-2},k_{N-1}$ respectively. Also in this case three auxiliary knots are needed.

As shown above, the basis for a piecewise polynomial function with $N+1$ knots $\{k_0,\ldots,k_N\}$ is composed of $N+3$ B-splines starting at $k_{-3},\ldots,k_{N-1}$ and based on a set of main and auxiliary knots: $\{k_{-3},\ldots,k_{N+3}\}$. Any cubic piecewise polynomial S(x) can be represented as a linear combination of these $N+3$ B-splines:

$$S(x)=\sum_{i=-3}^{N-1} a_i B_i(x) \tag{13}$$

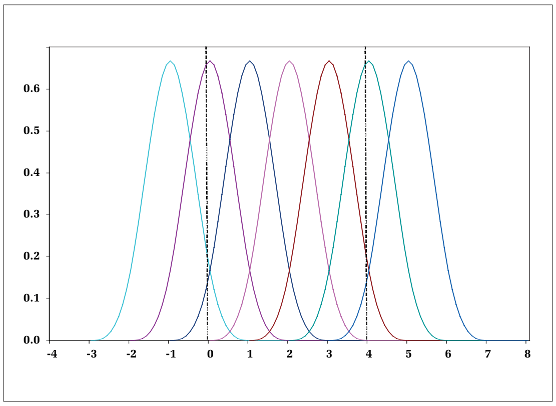
where $B_i(x)$ denotes a B-spline starting at $x = k$, and a_i – a corresponding parameter. For instance, the B-splines constituting a basis for a cubic polynomial spline on the interval <0, 4> with knots located at x={0, 1, 2, 3, 4} are presented in Figure 2.

Before we show how to determine the value of S(x) let us introduce some basic notation and terms. A B-spline of degree n based on a sequence of primary knots $\{k_0,\ldots,k_N\}$ is a combination of ordinary polynomials of degree $[n-1]$. It is therefore $[n-2]$ times differentiable on the interval including its auxiliary knots $\{k_{-3},\ldots,k_{N+3}\}$. A cubic B-spline has a degree equal 4. Let us introduce the following notation: i-th B-spline (taking positive values from the i-th knot) of degree n will be denoted as: $B_{i,n}(x)$.

Unfortunately the value of a B-spline at a given x cannot be represented by a single (non-compound) function. The evaluation is usually done recursively

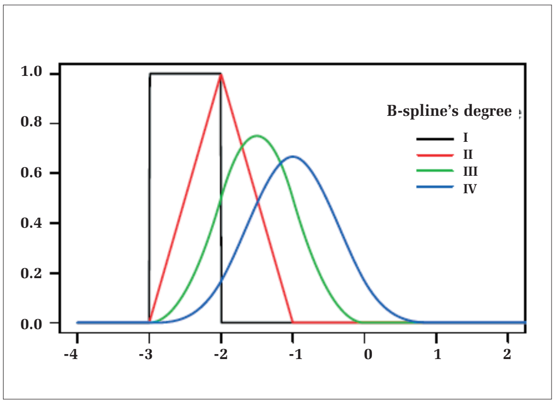
[8] A detailed discussion and derivation can be found in Lancaster, Salkauskas (1986).

Figure 2. *B-spline basis for the interval <0,4>; equidistant knots*



Source: own calculations.

Figure 3. *B-splines of different degrees*



Source: own calculations.

on a basis of B-splines of lower degrees using the following formula:

$$B_{i,n}(x) = \frac{x-k_i}{k_{i+n-1}-k_i} B_{i,n-1}(x) + \frac{k_{i+n}-x}{k_{i+n}-k_{i+1}} B_{i+1,n-1}(x) \qquad (14)$$

for $i = [-3, -2,\ldots, N-1]$, $n = [1, 2,\ldots]$, where k_i denotes position of the i-th knot. To use this calculation scheme it is necessary to determine the values of a B-spline of degree 1:

$$B_{i,n}(x) = \begin{cases} 0 : x \in (\infty, k_i) \\ 1 : x \in < k_i, k_{i+1}) \\ 0 : x \in < k_{i+1}, \infty) \end{cases} \qquad (15)$$

With application of the above formulas it is easy to evaluate any B-spline $B_{i,n}(x)$ and to find the value of a piecewise polynomial $S(x)$ for any $x \in < k_0, k_N >$. Shapes of B-splines of different degrees are presented in Figure 3.

1.4.2. Least Squares estimation

B-splines facilitate curve modelling with piecewise polynomials. If a curve is to be fitted to a set of *observable* points (e.g. money market rates or interest rate swap quotes) then the application of the least squares method is possible. In the remainder of this section practical issues related to B-spline modelling are discussed. In order to make the notation transparent and easy to follow the following convention is introduced: a) capital letters in bold type represent *matrices*, b) small letters in bold type represent *column vectors*, c) all letters in standard type (not bold) represent *scalar* values.

Let us suppose that we want to fit a piecewise polynomial curve $S(x)$ to a set of observable money market rates *r(x)* for $x \in \{x_1,\ldots,x_m\}$. The polynomial is based on a set of primary knots $\{k_0,\ldots,k_N\}$. Due to its convenient statistical properties the sum of squared errors of fit will be used as a measure of goodness of fit and denoted S^2. The purpose of estimation is to find a curve which minimizes the value of the following expression:

$$S^2(\boldsymbol{a}) = \sum_{i=1}^{m} \left(S(x_i) - r(x_i)\right)^2 \qquad (16)$$

where m denotes the number of data points $(x_i; r(x_i))$, to which a piecewise polynomial $S(x)$ is fitted.

The LS curve estimate is defined as a vector of parameters, for which the following set of conditions is fulfilled:

$$\frac{\partial S^2(\boldsymbol{a})}{\partial a_j} = 0 \quad \text{for j=-3,}\ldots, N-1. \qquad (17)$$

A solution of the above system of equations gives so called *normal equations*, which can be reformulated in a matrix form as follows:

$$\boldsymbol{B}^T\boldsymbol{B}\boldsymbol{a} - \boldsymbol{B}^T\boldsymbol{r} = 0 \qquad (18)$$

where:

$$\mathbf{a} = \begin{bmatrix} a_{-3} & a_{-2} & \ldots & a_{N-1} \end{bmatrix} \qquad (19)$$

$$\mathbf{r} = \begin{bmatrix} r(x_1) & r(x_2) & \ldots & r(x_m) \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} B_{-3}(x_1) & B_{-2}(x_1) & \cdots & B_{N-1}(x_1) \\ B_{-3}(x_2) & B_{-2}(x_2) & \cdots & B_{N-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_{-3}(x_m) & B_{-2}(x_m) & \cdots & B_{N-1}(x_m) \end{bmatrix}$$

After a proper transformation of equation (18) we obtain the standard form of the LS estimator:

$$\boldsymbol{a} = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{r} \qquad (20)$$

As already mentioned B-splines can be used to construct piecewise polynomial curves with very convenient mathematical properties. Nevertheless, curve fits obtained with this method are in principle identical to curves obtained on a basis of standard polynomials. The difference is purely technical – but crucial – and consists in the correlation of columns in the matrix $\boldsymbol{B}$. Low level of co-linearity between the columns enables fast and precise inversion of the cross-product matrix $\mathbf{B}^T\mathbf{B}$. The next difference lies in the simplicity of imposing restrictions (e.g. for smoothing or stabilizing) on piecewise polynomials. In the case of B-spline curves a number of relatively complex restrictions[9] can be imposed in a linear form. This feature of B-splines represents a significant advantage over the standard polynomial basis. This issue is be discussed in some more detail in the next subsection.

1.5. Non-parametric estimation methods

The B-spline model discussed in the previous subsection is to a large extent analogous to the model proposed by McCulloch based on standard polynomials Just like in the case of standard polynomials, goodness of fit and smoothness are dependent on the number and location of knots. As a result, B-spline models inherit some weaknesses of standard piecewise polynomials. In this subsection we discuss a method to solve the problem of yield curve's excess elasticity put forward by Fisher, Nychka and Zervos (1994).

1.5.1. Smoothed splines

In the spline models described so far the preferred degree of curve smoothness and goodness of fit was controlled by changing the number and location of knots. Despite some adjustability the method had significant constraints and bore serious problems of

[9] However in some cases through some simplifications.

technical nature. Fisher, Nychka and Zervos avoided these problems by fixing the knots and applying a widely known stabilizer[10] to ensure an adequate level of smoothness:

$$P(\mathbf{a}) = \int_a^b (S''(x))^2 dx \tag{21}$$

where a and b are locations of the first and last knot point respectively. $P(\mathbf{a})$ is a measure of the curvature of S(x). It takes values ranging from 0 for a straight line and gradually approaches infinity for curves with increasingly oscillatory behaviour.

To obtain a desired goodness of fit and an adequate level of smoothness the objective function (16) is modified by adding a scalar multiple of $P(\mathbf{a})$:

$$\mathcal{G}(\mathbf{a}) = S^2(\mathbf{a}) + \lambda \mathcal{P}(\mathbf{a}) \tag{22}$$

which gives:

$$\mathcal{G}(\mathbf{a}) = \sum_{i=1}^{m} \left(S(x_i) - r(x_i)\right)^2 + \lambda \int_a^b \left(S''(x)\right)^2 dx$$

where λ represents a scalar smoothness parameter.

The stabilizer $P(\mathbf{a})$ and the objective function $\mathcal{G}(\mathbf{a})$ can actually be applied both to the standard and B-spline piecewise polynomials. However, the convenient mathematical properties of B-splines make it possible to considerably simplify the notation and calculations, which explains their widespread use.

Eilers and Marx (1996) observed that with equidistant knots the value of $P(\mathbf{a})$ can be approximated with the following expression:

$$\mathcal{P}(\mathbf{a}) \approx \sum_{j=-1}^{N-1} \left(\Delta^2 a_j\right)^2 \cdot const \tag{23}$$

Using the above expression the minimum value of the stabilizer $P(\mathbf{a})$ can be found as follows:

$$\frac{\partial \mathcal{P}(\mathbf{a})}{\partial a_j} \approx \frac{\partial}{\partial a_j}\left(\sum_{j=-1}^{N-1}\left(\Delta^2 a_j\right)^2\right) = 2\sum_{j=-1}^{N-1}\left(\Delta^2\right)^2 a_j = 2\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{a} \tag{24}$$

where $\boldsymbol{D}$ is a matrix form of a second order difference operator. For a piecewise polynomial based on [N+1] primary knots: $\{\kappa_0, \dots, \kappa_N\}$ and defined by [N+3] parameters: $\{a_{-3}, \dots, a_{N-1}\}$ the matrix $\boldsymbol{D}$ has dimension [N+1]x[N+3] and the following form:

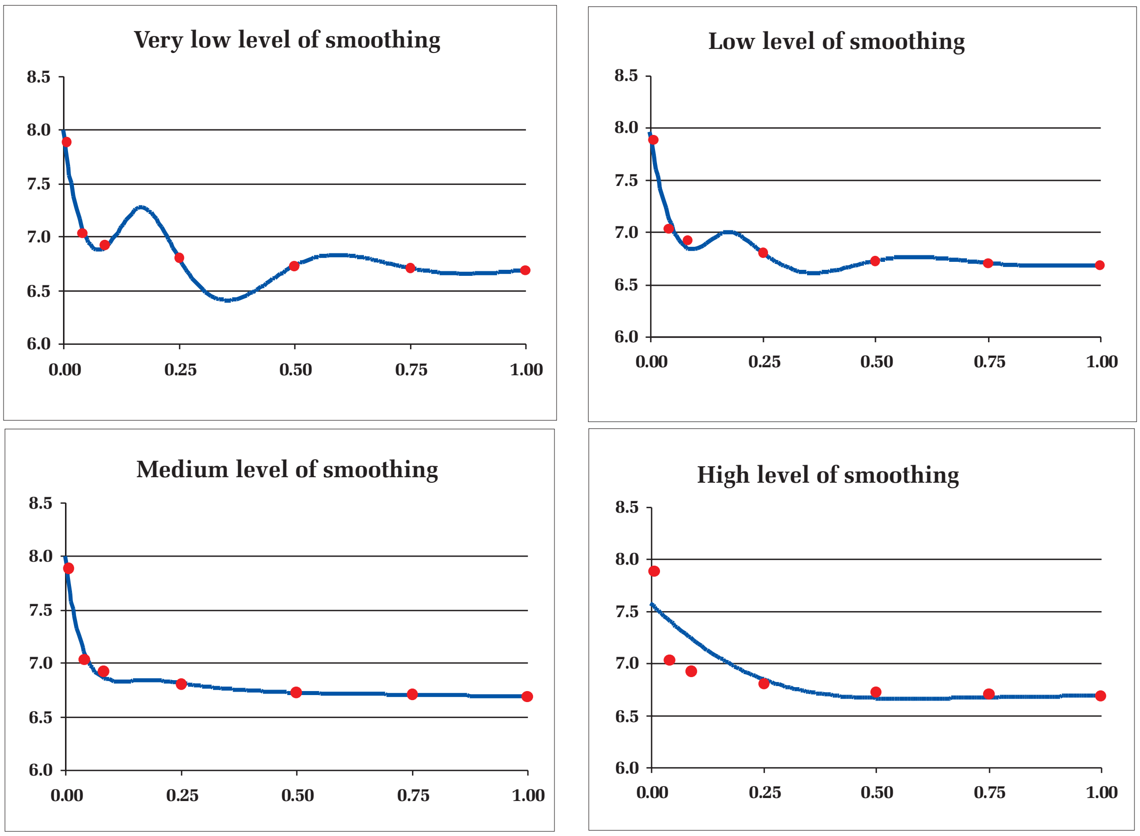
$$\mathbf{D} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \tag{25}$$

Just like in the case of the objective function (16) the vector of parameters for which the value of (22) is

---

[10] See e.g. Haerdle, Linton (1994).

Figure 4. *Money market rates in Poland - curve estimates for different levels of smoothing*



Source: own calculations.

minimized – the curve estimator – can be found by solving the following system of differential equations:

$$\frac{\partial \mathcal{G}(\mathbf{a})}{\partial a_j} = 0 \tag{26}$$

or equivalently:

$$\frac{\partial \mathcal{S}^2(\mathbf{a})}{\partial a_j} + \lambda \frac{\partial \mathcal{P}(\mathbf{a})}{\partial a_j} = 0 \text{ , for } j=-3,\ldots, N-1.$$

A combination of the results in (18) and (24) gives the following set of first-order conditions which constitute a solution to the problem defined by (26)[11]:

$$\boldsymbol{B}^T \boldsymbol{B a} - \boldsymbol{B}^T \boldsymbol{r} + \lambda \boldsymbol{D}^T \boldsymbol{D a} = 0 \tag{27}$$

After solving the above equation for the parameters a we obtain the OLS estimator of the curve parameter vector:

$$\boldsymbol{a} = (\boldsymbol{B}^T \boldsymbol{B} + \lambda \boldsymbol{D}^T \boldsymbol{D})^{-1} \boldsymbol{B}^T \boldsymbol{r} \tag{28}$$

The most significant component of the above formula is the smoothness parameter λ which controls the shape of the curve. By changing its value (level of smoothing) it is possible to influence the *effective number of parameters*[12]. Reducing their effective number results in a smoother curve at the expense of worse accuracy of fit. On the other hand, decreasing the value of the smoothness parameter, i.e. increasing the effective number of parameters yields more flexible and fluctuating curve with a better goodness of fit. In the Figure 4 curves estimated with this method are presented along with the influence of the changing value of the smoothness parameter.

Application of the above described smoothing splines model requires to develop a method of setting the appropriate value of λ. Fisher et al. (1994) proposed to use the so called *Generalized Cross Validation* (GCV) approach to find the optimal value of λ for each single day. Although the method is reported to give very good results, its implementation may cause serious problems of technical nature. Using numerical methods to determine the optimal value of λ (by e.g. the GCV criterion) requires a repeated estimation of the curve for a set of λ values. If the curve is being fitted to an observable set of data points then a single iteration of the estimation process (for a given value of λ) proceeds analytically on the basis of the formula (29). In this case the application of GCV is relatively simple. If however, the curve is being fitted to unobservable interest rates (e.g. bond market zero-coupon rates) then the application of linear estimators is not possible and numerical methods must be used to find the fit within each iteration. In such a situation the time expense increases dramatically and can reach several minutes for a single curve[13]. If the value of λ is to be set for each curve separately then the total estimation time would rise dramatically, which could make the analysis of long time series very time demanding and inconvenient. For this reason in practice a constant smoothness parameter is applied, and its value set arbitrarily or through a one-off use of GCV or other method[14].

1.5.2. Variable roughness penalty models

The piecewise polynomial models with a single parameter controlling smoothness of the entire curve have serious limitations. In general they stem from the fact that these models disregard the differences in volatility and smoothness of different yield curve segments. By definition, they choose an "average" smoothness for the whole curve, which frequently causes an inadequate level of accuracy (large pricing errors) for short maturity instruments, and strongly fluctuating interest rates at the long end of the curve[15].

A solution to this problem was put forward by Waggoner (1997), who made the FNZ method more flexible by applying a Variable Roughness Penalty (VRP) and modifying the objective function as follows:

$$\mathcal{G}(\mathbf{a}) = \sum_{i=1}^{m} \left(S(x_i) - r(x_i) \right)^2 + \int_a^b \lambda(x) \left(S''(x) \right)^2 dx \tag{29}$$

where a i b denote respectively the beginning and the end of the fitted function's domain.

Introducing a maturity-dependent power of smoothing made it possible to achieve both: a desirable goodness of fit and an appropriate level of smoothness for each yield curve segment. In the Waggoner's model $\lambda(x)$ is a piecewise constant function comprised of 3 segments. The division to segments was made arbitrarily – according to the natural market division – into the segments of short-, medium-, and long-term US treasury securities. The fixed division into the intervals: <0; 1), <1; 10), <10; 30> is undoubtedly a significant constraint on the model's adjustability.

Anderson and Sleath relax this constraint and apply a completely continuous smoothing function controlled by three parameters, given by:

---

[11] For a more detailed discussion see e.g. Bolder, Gusba (2002).

[12] Details on determining the effective number of parameters can be found in Fisher, et al (1994).

[13] Example of estimation times for a number of curve approximation methods can be found in the paper by Bolder and Gusba (2002), p. 66. The estimation time of a single curve depends mainly on the number of knots, choice of starting parameter values and the desired accuracy of fit.

[14] See e.g. Waggoner (1997) or Anderson, Sleath (2001).

[15] See Waggoner (1997 p.1); Bliss (1996).

$$\log\lambda(t)=\beta_0-(\beta_0-\beta_1)\exp\left(\frac{-t}{\beta_2}\right) \qquad (30)$$

The proposed method of smoothing facilitates precise calibration of the model but requires application of numerical procedures to find the optimal set of smoothing parameters, evaluate the roughness penalty and finally – to find the fit. Apart from increasing the complexity of calculations this leads to a potentially longer time of estimation.

1.5.3. Extension of the FNZ model to a piecewise continuous penalty function

In order to ensure high flexibility of a spline based yield curve model without necessitating numerical integration, it is possible to combine both of the above described approaches.

In this section an extension of the FNZ model is proposed to a piecewise continuous penalty function, which enables analytical evaluation of the smoothness component

$$\mathcal{P}\big(\lambda(x),\mathbf{a}\big)=\int_a^b\lambda(x)\big(S''(x)\big)^2\,dx \qquad (31)$$

The method proposed is based on the findings by Eilers and Marx (1996), who applied a discrete difference stabilizer instead of the above presented penalty on the integral of the squared second derivative. Let us generalize their model – with a single roughness parameter controlling the smoothness of the entire curve – to a case of a piecewise continuous stabilizer.

For $S(x)=\Sigma_{i=-3}^{N}a_iB_i(x)$ defined as a polynomial cubic B-spline with equidistant knots located at $\{k_0,k_1,\ldots,k_N\}$ the following relation holds:

$$\mathcal{P}=\frac{1}{h^2}\int_{k_0}^{\kappa_N}\lambda(x)\Big(\sum_i a_iB''_{i,4}(x)\Big)^2dx \qquad (32)$$

where h denotes a constant inter-knot distance. Due to the properties of B-splines, which take non-zero values only in a relatively narrow interval, the exact bounds of the indexes in the sums in the formulas for their derivatives will be omitted.

With the formulas for B-spline derivatives given by de Boor (1978), it is possible to transform the above expression in terms of second degree B-splines:

$$\mathcal{P}=\frac{1}{h^2}\int_{k_0}^{\kappa_N}\lambda(x)\Big(\sum_i\Delta^2a_iB_{i,2}(x)\Big)^2dx \qquad (33)$$

Given the properties of 2-nd degree B-splines, further transformation[16] leads to the following formulation:

$$\begin{aligned}\mathcal{P}&=\sum_i(\Delta^2a_i)^2\frac{1}{h^2}\int_{k_0}^{\kappa_N}\lambda(x)B_{i,2}^2(x)dx+\\&+2\sum_i\Delta^2a_i\Delta^2a_{i-1}\frac{1}{h^2}\int_{k_0}^{\kappa_N}\lambda(x)B_{i,2}(x)B_{i-1,2}(x)dx\end{aligned} \qquad (34)$$

The first term on the right side is equivalent to the difference penalty applied by Eilers and Marx (1996). Let us recall that it has the following form:

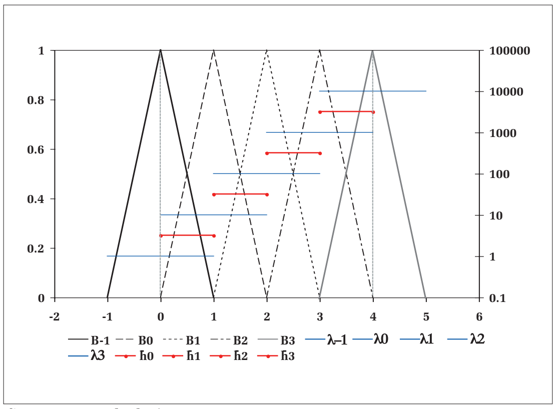
$$P=\sum_i\lambda(\Delta^2a_i)^2 \qquad (35)$$

The authors observed that there is a very strong connection between both penalties. It results from the equivalence between P and the first component of $\mathcal{P}$: Note that if we attach a different level of smoothness λ to each B-spline B_i then the first term in (35) can easily be transformed to the difference penalty P:

$$\frac{1}{h^2}\sum_i(\Delta^2a_i)^2\int_{k_0}^{\kappa_N}\lambda_iB_{i,2}^2(x)dx=\sum_i\ddot{h}_i(\Delta^2a_i)^2\cong P \qquad (36)$$

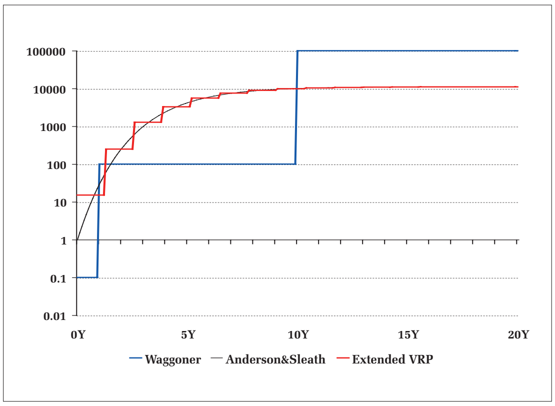
[16] For a step-by-step derivation see Eilers and Marx (1996).

Figure 5. *Second degree B-splines and the standard (overlapping) vs separated smoothness parameters*



Source: own calculations.

Figure 6. *Proposed discrete extension to the VRP vs penalty functions used by Waggoner and Anderson & Sleath*



Source: own calculations.

where $\ddot{h}_i$ denotes (appropriately re-scaled) i-th (corresponding to the i-th spline) element of the vector $\ddot{h}$ of roughness parameters.

The main advantage of using P instead of $\mathcal{P}$ is a significantly reduced complexity of notation and computation while the properties of the penalty are mostly preserved. Eilers and Marx point out that the difference penalty P is a good discrete approximation of $\mathcal{P}$:

The expression (37) can easily be transformed to a matrix form, however its weakness is the lack of a direct interpretation of the roughness vector $\ddot{h}$. This is a direct consequence of properties of the second degree B-splines, which cover 2 adjacent inter-knot intervals and partially overlap with the neighbouring splines. For this reason it is not possible to relate a given value of the vector $\ddot{\mathbf{h}}$ to a single segment of the fitted curve, and to interpret $\ddot{\mathbf{h}}$ as a function. However, this can be achieved after a proper transformation of $\ddot{\mathbf{h}}$.

Figure 5 contains a schematic representation of the relation between ordinates of the roughness penalty vector and the corresponding B-splines. Interpretation of this relation is complicated by the fact, that there are two different penalty function values attributed to each interval and corresponding to two distinct B-splines. Nevertheless, it is straightforward to show that each two penalties which correspond to adjacent B-splines can easily be replaced by a single value. It can be shown that its value is a weighted average of the overlapping penalties in each inter-knot interval. An example of the separated smoothness penalties is represented by the red segments in Figure 5.

Note that if before integration the curve domain is divided into intervals defined by its knot points then P can be represented as follows:

$$P = \frac{1}{h^2}\sum_j \left\{ (\Delta^2 a_{j-1})^2 \int_{k_j}^{k_{j+1}} \lambda_{j-1} B^2_{j-1,2}(x)dx + (\Delta^2 a_j)^2 \int_{k_j}^{k_{j+1}} \lambda_j B^2_{j,2}(x)dx \right\} =$$
$$= \frac{1}{h^2}\sum_j \left\{ \ddot{h}_{j-1}(\Delta^2 a_{j-1})^2 + \ddot{h}_j(\Delta^2 a_j)^2 \right\} \tag{37}$$

In the above formula in each interval the function is integrated twice, with two different smoothing parameters applied. The procedure can be simplified by applying a common smoothness parameter $\tilde{h}_j$:

$$\ddot{h}_{j-1}(\Delta^2 a_{j-1})^2 + \ddot{h}_j(\Delta^2 a_j)^2 = \tilde{h}_j \left\{ (\Delta^2 a_{j-1})^2 + (\Delta^2 a_j)^2 \right\} \tag{38}$$

which gives:

$$\tilde{h}_j = A'_{j-1}\ddot{h}_{j-1} + A''_{j-1}\ddot{h}_j$$

where $A'_{j-1} = \frac{(\Delta^2 a_{j-1})^2}{(\Delta^2 a_{j-1})^2 + (\Delta^2 a_j)^2}$ $\quad A''_{j-1} = \frac{(\Delta^2 a_j)^2}{(\Delta^2 a_{j-1})^2 + (\Delta^2 a_j)^2}$

It is now straightforward to see that after transformation the smoothness parameter vector $\ddot{h}$ can be interpreted as a piecewise continuous smoothing function:

$$\tilde{h}(x) = \tilde{h}_j \quad \text{for} \quad x \in \langle k_j; k_{j+1}) \text{, where } j = \{0,1,\ldots, N-1\}. \tag{39}$$

The vector $\tilde{\mathbf{h}} = \left[\tilde{h}_0, \tilde{h}_1, \ldots, \tilde{h}_{N-1}\right]^T$ can be represented as follows:

$$\tilde{\mathbf{h}} = \mathbf{A}\ddot{\mathbf{h}}, \quad \text{where} \tag{40}$$

$$\mathbf{A} = \begin{bmatrix} A'_{-1} & A''_{-1} & 0 & \cdots & 0 & 0 \\ 0 & A'_0 & A''_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A'_{N-2} & A''_{N-2} \end{bmatrix} \tag{41}$$

As indicated above, the modified difference stabilizer (37) represents a relatively accurate and convenient approximation of a piecewise continuous smoothing function. In practice it proves to be also a satisfying approximation of completely continuous, smooth functions.

The number of different values taken by the proposed smoothing function equals the number of inter-knot intervals. In practice their number is usually set to 10–20. For instance, Anderson and Sleath (2001) use on average 12 knots, whereas Bolder and Gusba (2002) about 20. A piecewise constant function taking about 10-20 different values enables a satisfying approximation of the majority of smooth functions. This implies that the method proposed in this paper offers much better flexibility than e.g. the three-parameter stepwise smoothing curve applied by Waggoner (1997). Moreover, due to the possibility of evaluating $\mathcal{P}(\lambda(x), \mathbf{a})$ analytically, the increase in elasticity virtually does not cause an increase in complexity of calculations and – consequently – longer estimation time. Figure 6 presents a possible approximation of the continuous smoothing function used by Anderson and Sleath by the means of a piecewise constant function. It is evident that for 15 knots the approximation will be satisfactory for the majority of practical applications. The function $\tilde{h}(x)$ described by equation (40) can be therefore treated to some extent as a quasi-continuous function.

The main advantage of the proposed difference stabilizer P is the convenience of application. With the matrix notation and analytical evaluation, its implementation and usage is much simpler and faster than in the case of continuous smoothing functions.

In order to facilitate the transformation of the expression (37) to a matrix form let us reformulate P as follows:

$$P = \sum_i \left(\sqrt{\ddot{h}_i} \Delta^2 a_i \right)^2 \qquad (42)$$

It is now easy to see that

$$\boldsymbol{P} = \left(\Lambda^{1/2} \boldsymbol{D a} \right)^T \left(\Lambda^{1/2} \boldsymbol{D a} \right) = \boldsymbol{a}^T D^T \Lambda \boldsymbol{D a} \qquad (43)$$

where Λ is a diagonal matrix constructed on the basis of $[N+1]$ ordinates of vector $\ddot{\mathbf{h}}$. The result is a computationally convenient and easily differentiable quadratic form.

Now the question of the parameterization of the piecewise constant function has to be resolved. The function values are defined by the parameter vector $\ddot{\mathbf{h}}$. With no additional restrictions imposed the number of parameters is equal to the number of knots. However, for parameter numbers encountered in the literature ranging from 10 to 20, this would translate into several additional parameters. The over-parameterization would cause considerable difficulties and result in much longer time of estimation. Furthermore this would make the results less reliable and reasonable.

One of possible solutions to this problem is to express the smoothness parameters as a function with a parsimonious parameterization. Let us denote this – in an ideal case monotonous and smooth – function as $l(x)$, then the parameter vector $\ddot{\mathbf{h}}$ defining the smoothness function can be represented as follows:

$$\ddot{h}_i = l(k_i), \text{ for } i = 0, 1, \ldots, N, \qquad (44)$$

where k_i denotes a location of the i-th main knot.

The above parameterization of the smoothing function based on the function (31) has been used for the purposes of this paper.

1.6. Zero coupon rate models for the bond market

So far the discussion concentrated mainly on presenting analytical solutions to the problem of curve fitting to observable data (e.g. money market rates). This section presents a general framework for estimation of bond market zero coupon rates both with the Svensson model and cubic spline models.

As a starting point let us represent the bond price equation in terms of cash flows:

$$p_i = \sum_{j=1}^{n} c_{i,j} d_{tj} \qquad (45)$$

where p_i denotes the settlement (dirty) price of the i-th bond, $c_{i,j}$ – the j-th cash flow on the i-th bond (including principal repayment), d_{tj} – a discount factor for cash flows due at the time t_j.

Taking into consideration the remarks we made on the choice of the type of interest rates to be directly estimated (see Section 1.3), we decide to fit the curve to zero coupon rates as the most reasonable solution.

Due to its computational simplicity and notational convenience the convention of continuously compounded interest rates is used. A unique relation between the discretely and continuously compounded rates ensures that using the latter does not cause any bias of the results. In the final phase of the calculations continuous rates will be translated into discrete rates.

Let us begin by representing the bond price equation (46) in a matrix form:

$$\hat{p}_i(\theta) = \boldsymbol{c}_i^T \boldsymbol{d}_i(\theta) \qquad (46)$$

where $\hat{p}_i$ denotes a fitted price of the i-th bond, θ – the curve parameters, $\mathbf{c_i}$ – vector of cash flows, $\mathbf{d_i}$ – vector of discount factors corresponding to cash flows on the i-th bond.

Vectors $\mathbf{c_i}$ and $\mathbf{d_i}$ have the following form:

$$\boldsymbol{c}_i = \left[c_{i,1}, c_{i,2}, \ldots, c_{i,ni} \right]^T \qquad (47)$$

$$\boldsymbol{d}_i = \left[\delta_{i,1}(\theta), \delta_{i,2}(\theta), \ldots, \delta_{i,ni}(\theta) \right]^T$$

where $c_{i,j}$ denotes the j-th cash flow on the i-th bond for $j = \{1, 2, \ldots, ni\}$, ni – the number of cash flows on the i-th bond[17], $\delta_{i,j}(\theta)$ denotes the value of the discount function (conditional on θ) for j-th cash flow on the i-th bond.

With the above notation at our disposal the vector of fitted bond prices can be formulated as follows:

$$\hat{\boldsymbol{p}}(\theta) = \left[\hat{p}(\theta)_1, \hat{p}(\theta)_2, \ldots, \hat{p}(\theta)_m \right]^T = diag\left[\boldsymbol{C}^T \boldsymbol{D}(\theta) \right] \qquad (48)$$

where m denotes the number of bonds in the data set, $diag[.]$ – vector of the diagonal elements of a square matrix, and finally

$$\mathbf{C}_{[n.\max \times m]} = \left[\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_m} \right] \qquad (49)$$

$$\boldsymbol{D}(\theta)_{[n.\max \times m]} = \left[\boldsymbol{d}_1(\theta), \boldsymbol{d}_2(\theta), \ldots, \boldsymbol{d}_m(\theta) \right]$$

In both of the above formulas the length of the vectors $\mathbf{c_i}$ and $\mathbf{d_i}(\theta)$ has been increased to be equal to the length $n.max$ of the vector corresponding to a bond with the largest number of cash flows. The vectors were extended by concatenating a number of zeros equating the vector's length to $n.max$. With this modification it was possible to apply the compact and very convenient formulation (49) for the bond price.

The main component of the objective function – squared price error – can be given the following form:

$$\min_{\theta} \left[\left(\boldsymbol{p} - \hat{\boldsymbol{p}}(\theta) \right)^T \boldsymbol{W} \left(\boldsymbol{p} - \hat{\boldsymbol{p}}(\theta) \right) \right] \qquad (50)$$

where $\mathbf{p}$ denotes a vector of the observable bond prices, $\mathbf{W}$ – a diagonal matrix of weights

[17] The last coupon and the principal repayment are discounted together.

corresponding to each single bond. The weightings are used in order to increase the influence of the most liquid and benchmark bonds on the parameter estimates, and to capture differences in the duration of the instruments. The result should be a yield curve reflecting each bond market segment with equal precision and giving a reasonable overall picture of the term structure of interest rates (see Section 1.3).

For a full specification of the model two elements are still missing: a functional form of the discount curve and the objective function. In the next two subsections we discuss the specification details of the two models under consideration – the Svensson model and a VRP cubic spline model.

1.6.1. Discount function and objective function in the VRP model

Let us now present the functional form of the discount curve (48) for the B-spline VRP model. The curve of zero coupon rates will be modelled in the convention of continuous interest rates. The maturity range covered and the location of knots are determined by the maturity structure of instruments in the respective bond market. For the purposes of this paper a yield curve model covering the maturity range from 0 to 12 years was implemented. For the ease of calculation equidistant knots were chosen and located at maturities of every second year, i.e. {0, 2, 4, 6, 8, 10, 12}.

The zero-coupon curve is defined as a linear combination of 9 B-splines, which altogether are based on 7 primary knots and 6 auxiliary knots: $k_i \in \{-6,-4,\ldots,16,18\}$:

$$r(t) = \boldsymbol{b}(t)\theta \quad \text{, for} \qquad (51)$$
$$\boldsymbol{b}(t) = \left[b_{-3}(t), b_{-2}(t), \ldots, b_5(t)\right]$$

where $\boldsymbol{\theta}$ is a column vector of parameter values with length equal to the number of B-splines, $b_t(t)$ – denotes the value of the i-th (starting at the k-th knot) B-spline.

With the notation introduced as above the value of a discount factor defined by (48) can be represented as follows:

$$\delta_{i,j}(\theta) = \exp\left(b(\tau_{i,j})\theta \cdot \tau_{i,j}\right) \qquad (52)$$

where $\tau_{i,j}$ denotes time to the j-th payment on the i-th bond.

Estimation of the parameter vector $\boldsymbol{\theta}$ amounts to finding a minimum of the objective function whose general form is represented by (30). With the expression (51) it can be rewritten in a matrix form:

$$\min_{\theta}\left[(\boldsymbol{p} - \hat{\boldsymbol{p}}(\theta))^T \boldsymbol{W}(\boldsymbol{p} - \hat{\boldsymbol{p}}(\theta)) + \theta^T \boldsymbol{D}^T \Lambda \boldsymbol{D}\theta\right] \qquad (53)$$

A variety of numerical procedures can be used to find the optimal value of $\boldsymbol{\theta}$. For the purposes of this paper Newton-type algorithms were used.

1.6.2. Form of the discount and objective function in the Svensson model

In the case of the Svensson model also a curve of zero-coupon continuously compounded interest rates was chosen to be directly estimated. Given the properties of the function there is no need to explicitly restrict the area of domain - unlike in the case of the VRP model.

A value of a single discount factor defined by (48) can be represented as follows:

$$\delta_{i,j}(\theta) = \exp\left(\vartheta(\tau_{i,j};\theta) \cdot \tau_{i,j}\right) \qquad (54)$$

where $\vartheta(\tau_{i,j};\theta)$ denotes the value of a zero-coupon curve for a time to maturity corresponding to the j-th cashflow on the i-th bond. The curve is defined by a set of six parameters $\boldsymbol{\theta}$ and its functional form is given by (12).

The main advantage of the parsimonious Svensson model is a relatively high level of smoothness both of the forward and zero-coupon rate curve. Therefore it is not necessary to impose additional stability constraints on the curve, and the objective function (51) can be used.

2. Svensson model vs VRP model – analysis of estimates for the Polish bond market

In this section we analyse the properties of the parsimonious model by Svensson and the B-spline model with a variable roughness penalty (VRP). The analysis is conducted using data from the Polish bond market from 03.2004 to 03.2006[18].

The yield curve model encompasses bonds with actual time maturity ranging from 1 to 12 years at the time of estimation. Due to low liquidity and potentially ineffective market pricing, bonds with less than one year remaining to maturity were not included in the data set. The short end of the curve was estimated using the interbank deposit market fixings - WIBOR rates. A single yield curve was approximated using a combination of 8 WIBOR rates and 13-15 treasury bond prices. Weightings applied to the instruments are dependent on their duration, and in case of bonds also liquidity.

2.1. Setting the smoothness penalty for the VRP model

In the previous section a general specification of the models has been presented. A crucial issue which has

[18] All calculations were made with **R** – the open source system for statistical computing. See http://www.r-project.org for details.

not been discussed yet is a method of choosing the roughness parameters for a VRP model.

An adequate level of roughness penalty should ensure a desirable smoothness of the curve, and improve stability of estimates. With the penalty parameters adjusted properly a stabilizer applied to the objective function serves as a kind of a "skeleton" supporting the curve when the data set used for estimation changes or – in particular – is reduced. For this reason, the optimal function of smoothness levels $\lambda(t)$ is usually chosen as the one which minimizes the sum of *out-of-sample errors*[19].

Bliss (1996) and Waggoner (1997) apply the out-of-sample method to calculate the optimal roughness penalty for the US bond market. They divide the data set into two equally large subsets, one of which is used to estimate the yield curve. Yield curve estimates are then applied to price bonds from the other subset and calculate the out-of-sample errors of fit.

Given the low number of instruments in the Polish bond market applying the same method would be not feasible and methodologically incorrect. For this reason a modification of this approach has been used in this paper. The method is called *leave-one-out cross validation*[20] and consists in multiple estimation of a single curve each time with one bond omitted from the sample. The price estimates for omitted bonds are compared with their actual market pricing. A sum of squared pricing errors for all bonds calculated this way is a measure of the out-of-sample pricing error. It is minimized by adjusting the parameters of the smoothing function. Anderson and Sleath use function with the following form:

$$\log \lambda(t) = \beta_0 - (\beta_0 - \beta_1)\exp\left(\frac{-t}{\beta_2}\right) \qquad (55)$$

The same function has been applied in this paper. The parameters were calculated on the basis of curve estimates for 25 different days representing the middle of each month included in the data set, i.e. 03.2004–03.2006. Using a day from the middle of each month for calculation guarantees that all shapes that the yield curve has taken within the sample period are included in the analysis. Actually, a rough estimate of smoothness parameters could be obtained through out-of-sample error minimization for a sample of just several dates. Optimization based on a larger number of dates allows to fine tune the rough estimate. However, the reaction of estimates to additional dates in the analysed sample becomes insignificant after a certain number of dates included is reached. In the case of the VRP model implemented for Poland the smoothness parameter estimates were rather stable when the number of dates in the sample exceeded 20. As a result, the smoothness parameters calculated this way should be a good approximation of the full sample estimates.

The reason for the relatively low number of curves used to determine the smoothing parameters is mainly the time expenditure for computation. Estimation of a single curve takes about 30 seconds. If the minimum is reached after 20 iterations on average, then the time needed to compute $\lambda(t)$ is 0.5x20x14=140 minutes. Using a larger number of days for calculation would result in a significantly higher time expenditure with potentially only slight improvement in precision. It is not uncommon to conduct out-of-sample optimization on a relatively small subset of available data. This approach was used by e.g. Bolder and Gusba (2002).

2.2. Goodness of fit

This subsection evaluates the models by comparing their goodness of fit. In order to ensure a complete comparability of results the same data set is used for estimation of both models.

Two measures of goodness of fit are used. The first is the *root mean squared error* (RMSE) defined as:

$$RMSE = \sqrt{\sum_{i=1}^{m} \frac{(\hat{p}_i - p_i)^2}{m}} \qquad (56)$$

where m denotes the number of instruments used for estimation. The second measure is the *mean absolute error* (MAE) defined as:

$$MAE = \sum_{i=1}^{m} \frac{|\hat{p}_i - p_i|}{m} \qquad (57)$$

In RMSE more weight is assigned to extraordinarily high error values. Large differences between RMSE and MAE indicate a large number of large errors of fit.

[19] This optimality criterion for the roughness penalty parameters was used by e.g. Waggoner (1997), Anderson and Sleath (2001), Bolder and Gusba (2002).

[20] It is used also by e.g. Anderson and Sleath (2001) for the British bond market.

Table 1. *Goodness of price fit (PLN/100 PLN of nominal value)*

Model	Price RMSE				Price MAE			
	mean	median	std. dev.	3Q-1Q	mean	median	std. dev.	3Q-1Q
Svensson	0.076	0.071	0.026	0.027	0.060	0.056	0.018	0.023
VRP	0.070	0.068	0.026	0.021	0.054	0.052	0.017	0.016

Source: own calculations.

Table 2. *Estimates of mean error (bias) and RMSE/MAE statistics in different yield curve segments*

Model		Time to maturity (years)				
		<1; 2)	<2; 3)	<3; 5)	<6; 8)	<8; 10)
Svensson	bias	-0.006	0.013	-0.007	-0.009	-0.017
	RMSE	0.074	0.093	0.079	0.077	0.118
	MAE	0.058	0.074	0.059	0.061	0.075
VRP	bias	-0.001	0.009	-0.003	0.008	0.012
	RMSE	0.066	0.073	0.068	0.056	0.072
	MAE	0.050	0.057	0.051	0.042	0.043

Source: own calculations.

Table 1 presents the RMSE and MAE statistics for both yield curve models. In each case the statistics were calculated on the basis of ca. 500 separate estimations. The median and interquartile range (difference between the quartile III and I) were provided to facilitate the analysis of the distributions, in particular to take into account the influence of outliers, which can potentially distort the values of mean and standard deviation statistics.

Although the results are actually comparable, the VRP model performs slightly better – on average by 0.006. Distributions of errors from both models are skewed to the right – which is quite typical for distributions of squared variables. Significant differences between RMSE and MAE statistics indicate that there are a potentially large number of exceptionally high pricing errors. Table 2 contains estimates of the pricing bias (mean value of errors) and pricing error statistics for different yield curve segments.

For both models estimates of the mean error (bias) are not significantly different from zero (t-values lower than 1) in each maturity segment. However, in the sample the mean error was always higher for the Svensson model. The distribution of errors across maturities is roughly uniform although the Svensson model had a few significant pricing errors in the longest maturity segment as indicated by relatively high RMSE and moderate MAE.

The above results are compatible with the properties of both models (as described in section 1). The VRP model – with theoretically unlimited elasticity was expected to have a higher precision of fit. Relatively small differences between pricing errors of both models result from a small number of securities in the Polish bond market. The elasticity of the VRP model is not fully utilized when there are only one or two bonds between neighbouring knots. Estimates for other markets presented in the literature confirm this explanation. Pricing errors for the Svensson model are in general significantly higher than for the VRP model. In-the-sample MAE statistics obtained by Bliss (1996) for the US Treasury bond market were equal 0.18 for the Svensson model, and 0.10 for the FNZ B-spline model. MAE statistics for the Canadian bond market obtained by Bolder and Gusba (2002) were equal 0.73 for the Svensson model and 0.21 for the B-spline zero-coupon model. The above presented pricing errors for the US and Canada are significantly higher than the errors for Poland. This is again related mainly to a relatively small number of bonds in the Polish bond market, but it also confirms the efficiency of pricing at the short end of the curve (1–3 years) where the number of bonds is larger.

2.3. Smoothness and stability of estimates

Beside the goodness of fit, smoothness of a yield curve and stability of parameter estimates are the main determinants of the model's usefulness. Both models have similar performance in this respect. The curves of zero-coupon and forward rates are smooth and mostly identical, although minor differences occur in some cases. Figure A in Appendix presents samples of zero-coupon curve estimates for both models. Figure B presents the corresponding curves of forward rates.

Smoothness (and elasticity) of a yield curve is closely related to its stability, defined as robustness to changes in the data set. Table 3 contains results of out-of-sample analysis conducted for both models with the leave-one-out cross validation method.

Table 3. *Stability of estimates – out-of-sample pricing errors*

Model	Price RMSE				Price MAE			
	mean	median	std. dev.	3Q-1Q	mean	median	std.dev	3Q-1Q
Svensson	0.093	0.081	0.030	0.031	0.073	0.069	0.022	0.025
VRP	0.092	0.079	0.031	0.032	0.074	0.067	0.023	0.025

Source: own calculations.

Table 4. *Out-of-sample pricing errors in different yield curve segments for the Svensson and VRP model (PLN/100 PLN of nominal value)*

Model		Time to maturity (years)				
		<1; 2)	<2; 3)	<3; 5)	<6; 8)	<8; 10)
Svensson	RMSE	0.077	0.086	0.096	0.107	0.160
	MAE	0.058	0.071	0.073	0.085	0.130
VRP	RMSE	0.076	0.082	0.096	0.108	0.150
	MAE	0.059	0.067	0.074	0.085	0.118

Source: own calculations.

Before discussing the results, let us recall the basic features of both models which are critical for the results of the out-of-sample analysis. First, the Svensson model has a parsimonious functional form ensuring an adequate level of smoothness and stability. It is therefore expected that the model performs good in the stability analysis. The B-spline models are potentially perfectly elastic and – as a consequence – potentially very unstable. However, with the VRP methodology adopted here the model is optimized for stability (minimized out-of-sample errors). For this reason its results should be good as well while still offering comparable (or better) goodness of fit.

Although the results from both models are very similar for the Polish bond market, in most cases the VRP model shows slightly better performance. The out-of-sample error statistics in Table 3 are consistent with the results presented in the literature for other bond markets. Anderson and Sleath's (2001) MAE estimates for UK are 0.09 for the Svensson model and 0.088 for the VRP B-spline model.

Table 4 reveals interesting facts about the distribution of errors across maturities. There is a clear positive correlation between the value of RMSE and MAE statistics and time to maturity. This is related to a relatively low number of bonds in the medium and long maturity segments in the Polish bond market – since 2004 the gap between maturities of longest bonds ranged between 2 and 3 years. When a bond was removed from the sample in the out-of-sample analysis, the gap could increase to even 5 years. As a result, out-of-sample yield curve estimates in the medium and long segment deviated significantly from full sample estimates for Poland.

2.4. Conclusions

The above presented empirical results were obtained for the Polish bond market characterized by a relatively low number of different bond issues. As a result, a thorough analysis of VRP model's capabilities was not possible, which resulted in fairly similar results for both models. Nevertheless, VRP outperformed other models by a small margin in most cases.

The advantages of the B-spline VRP models over the parsimonious Svensson model are widely known in the literature. While the B-spline models are considered fairly robust analytical tools[21] the Svensson approach has several widely known weaknesses which were not addressed directly in this paper. The most important of them are:

– low elasticity at the short end of the yield curve (see e.g. Gurazdowski 2003),

– high degree of instability and a tendency to "explode" at the short end (see e.g. Stamirowski 1999),

– non-uniqueness of estimates and their dependence on the starting point for estimation (see e.g. Csajbok 1998),

– high correlation of estimators, and low robustness of estimates to changes of instrument sets or their prices in distinct - even distant (not adjacent) - segments of the curve (see e.g. Anderson, Sleath 2001).

As the main advantages of the Svensson model – low complexity and computational requirements – have become less important[22] a number of institutions switched to piecewise polynomial VRP models. The VRP models are currently used by a large number of major central banks worldwide, e.g. Federal Reserve Banks in the US, the Bank of Japan, the Bank of England (BIS 2005).

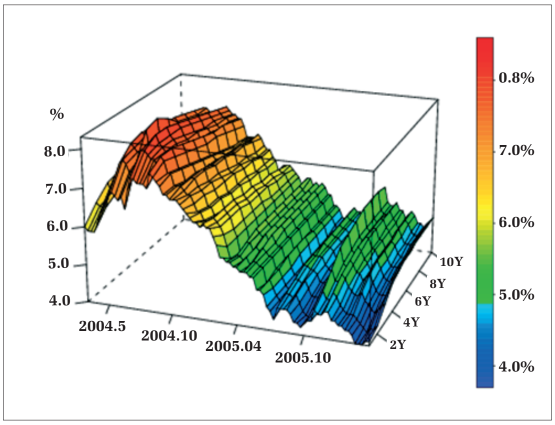
3. Interest rates in Poland - term structure and dynamics

In this section we use the VRP estimates to analyse bond market dynamics in Poland. In particular, we apply a simple time-series based method to evaluate the influence of economic events (e.g. monetary policy decisions) on the shape and level of the yield curve. Figure 7 presents the VRP model based zero-coupon rate estimates for the Polish bond market.

[21] See e.g. Bliss (1996), Waggoner (1997), Anderson, Sleath (2001), Bolder, Gusba (2002).

[22] Mainly due to a fast development of computer technology, which allowed a drastic reduction in the time of computation.

Figure 7. *Term structure of zero-coupon rates in Poland*



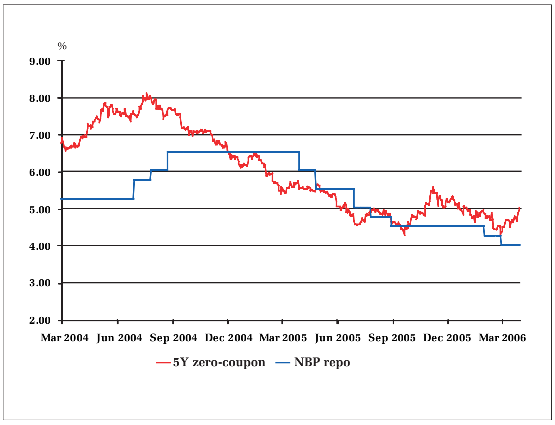
Source: Reuters; author's calculations.

3.1. Bond market dynamics in Poland

A comprehensive analysis of bond market dynamics should be based on a data panel consisting of time series of zero-coupon rates covering a full range of maturities. Zero-coupon rates obtained for Poland with the VRP model (see Figure 7) are appropriate for this purpose.

Within the framework of "traditional" time series analysis, complete identification of the processes determining the yield curve structure and dynamics would require a multi-dimensional and potentially integrated structural VAR model with a multivariate GARCH noise process. Such comprehensive time-series modelling is beyond the scope of this paper. The analysis conducted in this section is limited to univariate analysis of zero-coupon rates in a single yield curve segment. This is sufficient to fulfil the main objective of this section, which is to capture the basic properties of bond market dynamics in Poland and present a method of testing influence of economic events on interest rates.

Figure 8. *NBP base rate and 5-year zero-coupon interest rates*



Source: Reuters.

The results presented below were obtained for 5-year zero-coupon rates. The 5-year bond market segment has satisfactory liquidity and its dynamics should capture both the market's short term forecasts of central bank rates as well as medium term expectations regarding inflation, monetary policy and other factors like GDP growth and credit risk. The analysis covers the time range 03.2004–03.2006 and includes virtually one full cycle of monetary policy tightening and easing (see Figure 8). This should guarantee that all basic yield curve shapes are included in the sample thus enabling a proper and unbiased examination of the data generating process.

The time series of 5-year zero-coupon rates were tested for integration with the ADF and KPSS unit root tests. In either case strong evidence was found in favour of the unit root hypothesis, which implied integration of order 1. Further analysis of the differentiated time series revealed heteroscedasticity (variable variance of the noise process), high kurtosis (7,52 vs. 3 in a Gaussian distribution), and volatility clustering. The above features – characteristic for GARCH processes – were taken into account in the final specification of the model. Eventually the model was structured as a zero-mean AR(1)-GARCH(1,1) process with the following specification and parameter estimates:

$$\begin{aligned} \Delta r_t &= 0.8\Delta r_{t-1} + U_t\sigma_t \\ \sigma_t^2 &= 1.6\cdot 10^{-4} + 0.08(\Delta r_{t-1})^2 + 0.89\sigma_{t-1}^2 \end{aligned} \qquad (58)$$

where r_t denotes a 5-year zero-coupon rate at the time t, σ^2 – variance of the white noise component, and $U_t \sim N(0,1)$. The estimates were found statistically significant at the standard level 5%.

3.2. Influence of changes in central bank rates on bond yields

With the above presented model it is possible to forecast volatility in the bond market (at least for a short time horizon). Volatility forecasts can be used to formulate interval forecasts of the level of interest rates. The concept of interval forecasts consists in determining a fluctuation range (confidence interval) which – with a given probability – will not be exceeded within a given time in the future. A forecast with a 90% confidence interval assumes, that in only 10 out of 100 cases the realization of the analysed process will exceed the forecast interval bounds. If we assume, that these 10% correspond to market's reactions to unanticipated and significant events then within the same framework a test for the significance of these events can be constructed. Bounds of the interval forecast can be treated as critical values of the test and used for evaluating the realized reaction of the variable tested. If the bounds are exceeded then with a probability of making a mistake equal 10% a null hypothesis should be rejected that the market's reaction to a given event (and thus its influence) was statistically insignificant (realization of the noise disturbance).

In order to present a possible area of application of this test let us analyse the influence of changes in the National Bank of Poland base interest rate on the domestic bond market. The analysis covers several dates from the period March 2004 – March 2006.

The test is conducted on the basis of zero-coupon rates with maturities from 1 to 12 years estimated with the VRP model. Testing the significance of the yield curve's reaction to an event requires a prior estimation of a series of critical values corresponding to different market segments. A set of critical values allows constructing a line of critical values which can be used for graphical evaluation of the results. In Figure C zero-coupon rate estimates for the Polish bond market are presented for dates immediately before and after NBP's monetary policy decisions. The reactions to decisions in June, July 2004, and June, July 2005 were analysed. The data sample contains decisions which had, and which did not have a significant impact on the market. This should facilitate a proper evaluation of the method proposed.

It's worth noting that the width of confidence bands varies across dates and maturities. It is determined by two major factors: the forecast of volatility in a given maturity segment and a number of days between the dates compared. The width of the confidence band is proportional to volatility on the one hand, and to a square root of time between the dates compared on the other.

Charts presented in Figure C facilitate visual evaluation of the market's reaction to NBP's interest rate decisions. The test is conducted by comparing a zero-coupon curve a day after the NBP MPC's meeting with its 90% confidence interval forecast made before the decision. If the confidence bands are exceeded then the market's reaction was statistically significant and – by implication – the decision not fully expected.

The analysis of yield curve reactions to NBP decisions reveals some interesting facts about the market's expectations regarding future interest rates in Poland. In particular, the results indicate that the Reuters economist poll[23] has only limited value as a source of information about the market's interest rate expectations.

Before discussing the main findings let us briefly recall several NBP's monetary policy decisions between March 2004 and March 2006. The main focus is on the yield curve reaction and accuracy of the Reuters poll median forecast.

The decision made on 30.06.2004 was different from the median forecast. In spite of this the market's reaction was moderate – bond yields rose at the short end and remained stable at the long end – which implies that the survey median did not precisely reflect the market's expectations. The situation in July was similar. Although the decision on 28.07.2004 to raise the base rate by 25 bp. was different from the median forecast, the market's reaction was statistically not significant. Again, the median forecast did not accurately reflect the actual expectations. The rate cut in June 2005 was deeper than the survey median expectations. As a result, the entire yield curve shifted downwards – a usual reaction. The situation in July was also standard. The NBP's decision was in line with the survey median and was followed by virtually no reaction in the bond market.

From the above case-study analysis it is evident that it is difficult to assess if the market was surprised by a monetary policy decision (or any other event) by comparing survey median forecasts with the actual central bank's decision. To properly evaluate the market's reaction to e.g. central bank's decisions several issues have to be kept in mind. First, expectations of market participants – responsible for any price action – may differ from expectations of economists taking part in the survey. Second, expectations expressed in the survey may change before the actual outcome (e.g. interest rate decision) is known. Third, significant information is lost when the distribution of expectations is characterized by the median only. The problem is especially important in case of strongly skewed distributions, when up to 50 per cent of market participants (or analysts) may be surprised (and trigger a significant price adjustment) even if the median forecast was correct[24]. Fourth, some information about individual expectations is lost already at the stage of gathering individual forecasts. Note that the forecast provided by an economist in a survey is actually a rough approximation of the probability-weighted mean of different possible scenarios. The answer to the survey question is the same when a probability equal 55% or 95% is attached to a given scenario. This should not be a serious problem for surveys with large number of participants when individual approximation errors tend to compensate. In case of smaller surveys however, this may occasionally result in an apparent

[23] Reuters survey conducted each month among market analysts and concerning the expectations regarding the next NBP's interest rate decision.

[24] It is important to note that given the heterogeneity of expectations all market participants whose forecasts were different from the actual outcome are – by definition – surprised. Nevertheless, the market – as a whole – may be not surprised (no significant price adjustment occurs) if the distribution of individual expectations is symmetric.

consensus of expectations while actually the probability attached to the median scenario by the majority of participants may be close to 50%. Fifth, even if the interest rate decision is in line with expectations, the market may be surprised by the central bank's communiqué after the monetary policy meeting and react with a price adjustment.

The problems mentioned above have to be kept in mind when survey based forecasts are evaluated. However, they do not play a significant role when the market's reaction (surprise) is measured with the proposed time-series based method. Regardless of the origin or reason of the market's reaction, the participants are assumed to be surprised if the price adjustment after an event is statistically significant in a given yield curve segment. In all the cases of NBP decisions described above the proposed test gave reasonable results, in each case consistent with the corresponding ex post comments by market analysts.[25] Therefore, despite its limitations which include the inability to identify and separate the reasons for price adjustments, the test may be a useful analytical tool.

4. Summary

The purpose of this paper was to present the most commonly used parametric and spline-based methods of yield curve estimation. The basic concepts related to yield curve modelling were presented and some up-to-date techniques discussed in more detail. The class of B-spline models smoothed with a variable roughness penalty (VRP) was found most reliable. The B-spline VRP models – with virtually unlimited calibration capabilities – offer yield curve estimates with a high degree of precision, along with an adequate level of smoothness and stability. Due to their high accuracy and low co-dependence of results for different curve segments, VRP model based zero-coupon rate estimates can be used for precise modelling of yield curve structure and dynamics. An example of their application is the method of evaluating the influence of exogenous factors on financial variables (e.g. interest rates). The main advantage of the method proposed is its capability to distinguish between economically meaningful and random price movements with a certain degree of statistical objectivity. Due to a high noise component in realisations of virtually all financial market processes, it is usually very difficult to correctly verify hypotheses concerning the influence of unexpected events on the market. Empirical analysis suggests that the proposed test may be a useful approach to solving problems of this kind.

In the technical part of the paper a modification of the standard smoothing mechanism for B-spline models has been proposed. The discrete stabilizer developed is an extension of the widely known difference penalty. The method preserves the main advantages of continuous stabilizers while facilitating analytical solutions and significantly reducing the time of computation.

When interpreting and evaluating results presented in this paper it has to be kept in mind that they were based on interest rate estimates from only one market. The Polish market of treasury bonds is characterized by moderate liquidity and periodically reduced efficiency. Given additionally a relatively low number of bond series outstanding and gaps in the time to maturity spectrum of the existing bonds, this constituted a serious obstacle to a fully reliable statistical analysis. Nevertheless, the empirical results obtained for Poland are consistent with findings by other authors obtained for other markets, which supports the reliability of the results presented in this paper.

[25] According to comments published on Reuters newswires.

References

Anderson N., Breedon F., Deacon, M., Derry, A., Murphy, G. (1996), *Estimating and Interpreting the Yield Curve,* John Wiley & Sons, Baffins Lane, Chichester.

Anderson N., Sleath, J. (2001), *New estimates of the UK real and nominal interest rates*, "Working Paper", No. 126, Bank of England, London.

Bekdache B., Baum, C.F. (1997), *The Ex Ante Predictive Accuracy of Alternative Models of the Term Structure of Interest Rates*, "Working Paper", No. 372, Boston College, Boston.

BIS (2005), *Zero-coupon yield curves: technical documentation*, "BIS Papers", No. 25, BIS, Basle.

Bliss R.R. (1996), *Testing Term Structure Estimation Methods*, "Working Paper", No. 96-12a, Federal Reserve Bank of Atlanta, Atlanta.

Bolder D.J., Gusba S. (2002), *Expotentials, Polynomials, and Fourier Series: More Yield Curve Modelling at the Bank of Canada,* “Working Paper”, No. 2002-29, Bank of Canada, Toronto.

de Boor Carl (1978), *A Practical Guide to Splines*, Springer, Berlin.

Csajbók A. (1998), *Zero-Coupon Yield Curve Estimation from a Central Bank Perspective*, “Working Paper”, No. 1998/2, National Bank of Hungary, Budapest.

Eilers P.H.C., Marx B.D. (1996), *Flexible Smoothing with B-Splines and Penalties*, “Statistical Science”, Vol. 11, No. 2, p. 89-121.

Fisher M., Nychka D., Zervos D. (1994), *Fitting the term structure of interest rates with smoothing splines, Board of Governors of the Federal Reserve System*, Washington, D.C.

Gurazdowski E. (2003), *Wykorzystanie modelu zmiennej sztywności krzywej stóp terminowych do przybliżania krzywej rynku pieniężnego*, „Bank i Kredyt”, nr 2, s. 87-91.

Haerdle W., Linton O. (1994), *Applied Nonparametric Methods*, in: R.F. Engle, D.L. McFadden (eds.), *Handbook of Econometrics*, Volume IV, Elsevier Science, North Holland.

Lancaster P., Salkauskas K. (1986), *Curve and Surface Fitting. An Introduction*, Academic Press Ltd, London.

McCulloch J.H. (1971), *Measuring the Term Structure of Interest Rates*, “Journal of Business”, Vol. 44, p. 19-31.

McCulloch J.H. (1975), *The Tax-Adjusted Yield Curve*, “The Journal of Finance”, Vol. 30, No. 3, p. 811-830.

Nelson C.N., Siegel A.F. (1987), *Parsimonious Modeling of Yield Curves*, “Journal of Business”, Vol. 60(4), p. 473-489.

Stamirowski M. (1999), *Empirical application of the "Nelson and Siegel" parsimonious zero-coupon yield curve model*, „Materiały i Studia”, No.16, NBP, Warsaw.

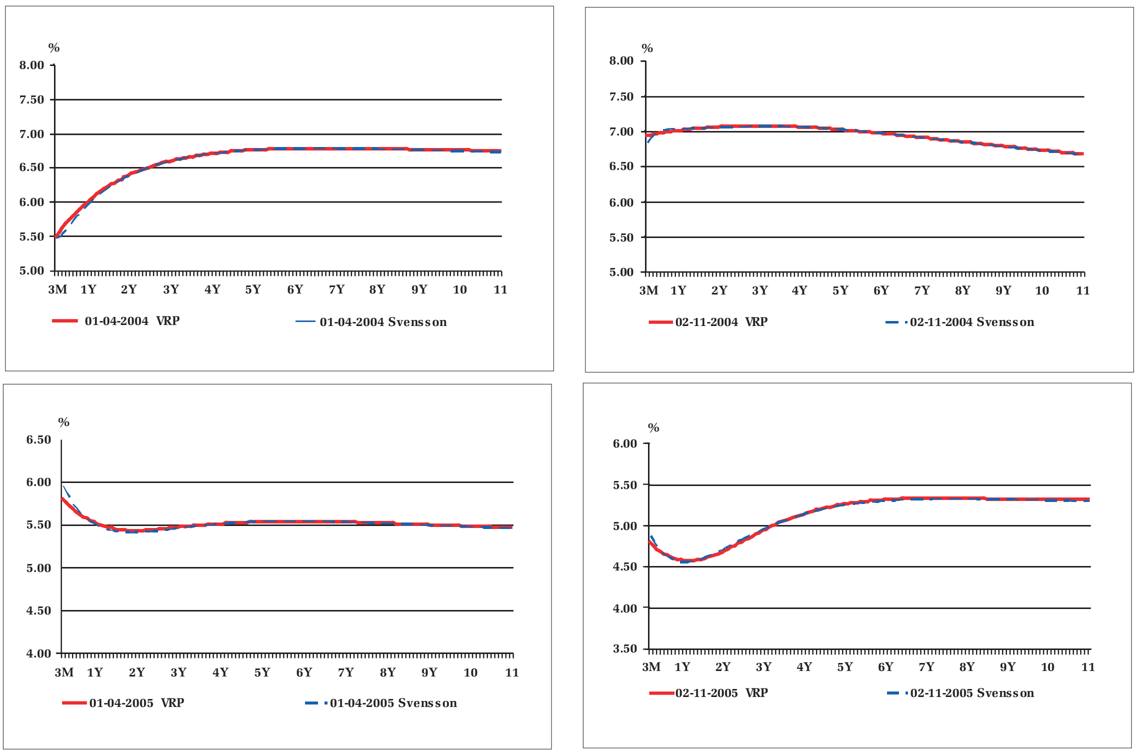
Svensson L.E.O. (1994), *Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994*, “Working Paper”, No. 4871, NBER, Cambridge, MA.

Svensson L.E.O. (1995), *Estimating Forward Interest Rates with the Extended Nelson & Siegel Method*, “Sveriges Riksbank Quarterly Review”, No. 3, p. 13-26.

Waggoner D.F. (1997), *Spline Methods for Extracting Interest Rate Curves from Coupon Bond Prices*, “Working Paper”, No. 97-10, Federal Reserve Bank of Atlanta, Atlanta.

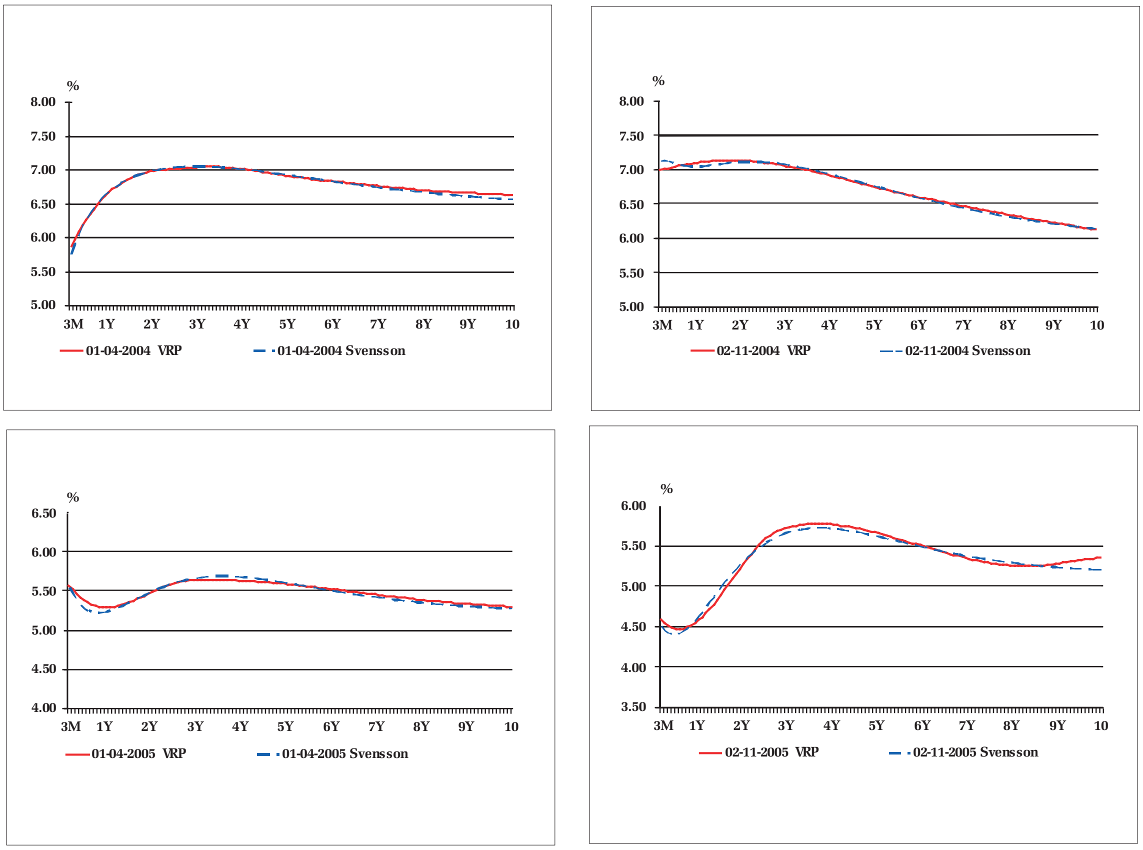
Appendix

Figure A *VRP vs Svensson model – zero coupon rate estimates*



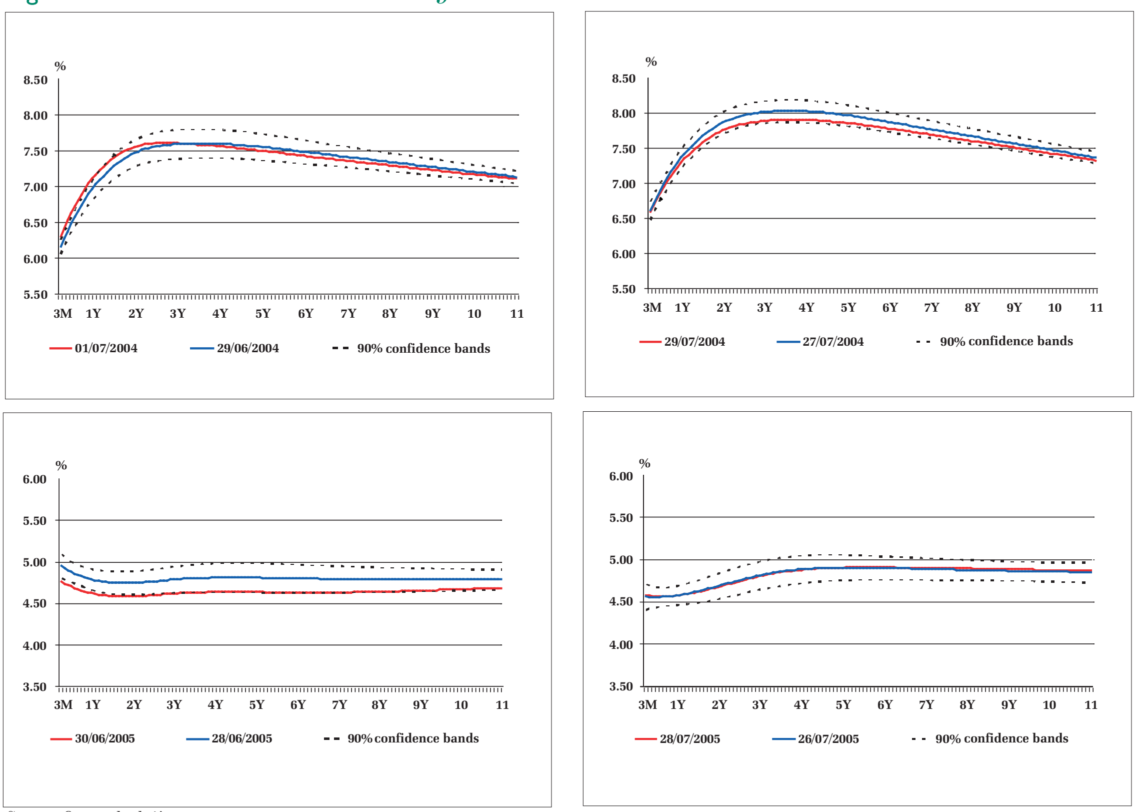
Source: Own calculations.

Figure B VRP vs Svensson model – 3-month forward rate estimates



Source: Own calculations.

Figure C Market's reaction to changes in NBP rates



Source: Own calculations.