Funded by
the European Union

This project was funded by the **EU** through the **EU-EBRD Local Currency Programme**.

Yield Curve Methodology

Contents

1. Methodology
2. Yield curve model
3. Model estimation
4. The publication of yield curve

1 Methodology

The yield curve estimation for a particular date is based on a moving historical data window of 160 days. This approach allows the model to estimate the yield curve even if recent data for some tenors is missing. Operationally, the update will be performed once per week, and the estimated curve for a specific date will be published. However, in the background, we will estimate the curve for all days within the window in a single batch using the time-stacked technique, explained in more detail in Section 3. Having the curve on a daily basis is crucial for the accurate valuation of incoming transactions on their respective days.

The actual data used for the estimation are[1]:

- fixed-rate MDL-denominated sovereign instrument' transactions in the primary, secondary markets, quotes, and
- interbank transactions, and
- NBM certificate operations, and
- NBM repo operations, and
- the calibrated model parameters and the model structure.

The model is estimated as a zero curve, with corresponding par curve, discount factors, and forward curve being derived from it, all of the former enabling the accurate valuation of existing instruments.

[1]National Bank reserves the right to exclude from the calculation transactions deemed to be carried out at exaggerated rates/prices.

2 Yield curve model

The underlying model is the Exponential Polynomial Model[2] that describes the zero-coupon yield curve on each particular day t. It has time-variant parameters $\beta_{0,t}$, $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, and a constant parameter λ. The continuously compounded zero yield $y_{t,\tau}$ on this day t and selected tenor τ is modeled as:

$$y_{t,\tau} = \beta_{0,t} + \beta_{1,t}\frac{1-e^{-\lambda\tau}}{\lambda\tau} + \beta_{2,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + 2\beta_{3,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\left(1+\frac{\lambda\tau}{2}\right)\right), \quad (1)$$

Note that in our estimation routine, time is measured in days, and we use the Actual/365 day count convention when establishing cash flows for observed instruments and determining model parameters. Consequently, years are considered to have 365 days or 366 days in leap years. However, when yields are quoted in terms of percent per annum, we assume 365 days in a year.

However, estimating these parameters independently each day in a data- scarce environment would lead to highly unstable estimates. Therefore the estimation procedure assumes that $\beta_{i,t}1$ and $\beta_{i,t}2$ are somehow connected, that is, they do not change much from period to period. How $\beta_{i,t}$ changes over time t is formalized by a stochastic model described below.

At the same time, estimating a model over 160 days rolling window with four time-variant beta parameters would imply a huge set of parameters to be estimated, while we are actually interested in the last set of the parameters to identify today's curve only (later denoted as t^{max}). The next week will use a new 160 day window.

Therefore, the estimation routine applies the following constraints: We create a grid of days, denoted as $s_1, s_2, \ldots, s_N$, which are evenly distributed to cover the entire 160-day estimation window.

[2]Serge-Moulin, (2018). Modeling the interest rate curve: the Exponential Polynomial Model ("EPM"), the true extension of Nelson-Siegel, ALM Vision

Specifically, we ensure that $s_1 \leq t^{min}$ and $s_N \geq t^{max}$. In our default calibration, we use a 56-day frame, making each frame s_k - s_{k-1} equal to 56 days for all k. However, the general model allows flexibility in this regard. This results in $N = 5$ points, and hence five set of parameters instead of 160. When constructing the grid, we position s_N so that it aligns with the last settlement date of all available transactions in the data set (typically today plus one or two days), placing it in the middle of the last frame. Subsequently, we set the remaining $s_1, s_2, \ldots, s_{N-1}$ according to the selected frame to cover the entire estimation window.
All β_{i,s_k} parameters in the selected grid are modeled by the random walk process:

$$\beta_{i,s_k} = \beta_{i,s_{k-1}} + \varepsilon_{i,k}, \quad \text{for } i = 0, \ldots, 3,\ k = 2, \ldots, N, \tag{2}$$

where $\varepsilon_{i,k} \sim N(0, \sigma(\beta_i)^2)$ with $\sigma(\beta_i)$ being time-invariant standard errors of innovations.

The situation is sketched in figure 1. Formally, we set $\beta_{i,t}$ as:

$$\beta_{i,t} = \begin{cases} \beta_{i,s_k} & \text{if } s_k = t \\ \frac{s_k - t}{s_k - s_{k-1}} \beta_{i,s_{k-1}} + \frac{t - s_{k-1}}{s_k - s_{k-1}} \beta_{i,s_k} & \text{if } s_{k-1} < t < s_k \end{cases} \tag{3}$$

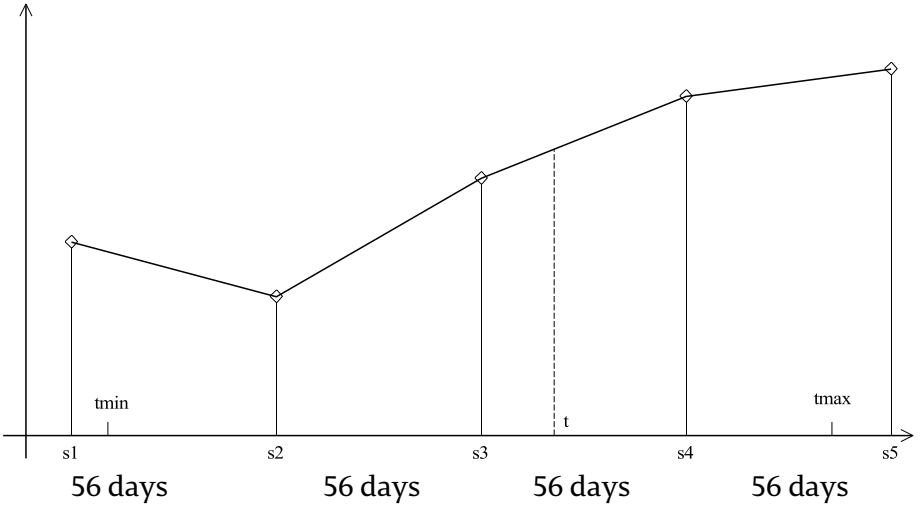

Figure 1: Interpolation of betas

There are two ways how the continuously compounded zero rates $y_{t,\tau}$ from equation (1) can be linked to the observed prices in historical transactions. In our approach, we link the zero rates $y_{t,\tau}$ with the observed transaction via discounted summation of instrument's cash-flows. An alternative approach sometimes used in practical models of the yield curves is to directly (after suitable data transformations) link the zero rates $y_{t,\tau}$ to yield-to-maturity quotes of the historical transactions. This, however, introduces internal inconsistency to the model which is more significant for markets where fixed income instruments bear high coupon rates.

Let $P_{i,t}$ be the dirty price of some paper[3] i in time t. The price can be calculated from the deterministic cash flows and unknown discount factors

$$P_{i,t} = \sum_j c_{i,j} d_{t,\tau_j}, \tag{4}$$

where j indexes all future cash-flows of the i paper, $c_{i,j}$ is the j-th cash-flow, d_{t,τ_j} is a discount factor between t and $t+\tau_j$ periods, when the cash-flow occurs. The discount factors are implied by the yield curve expressed in continuously compounded yields $y_{t,\tau}$ in percent per annum:

$$y_{t,\tau} = -\frac{100 \cdot M}{\tau} \log(d_{t,\tau}), \tag{5}$$

where M is a model parameter calibrated to 365, representing the number of days in a year.

Now we know how to calculate a theoretical dirty price $P_{i,t}$. However, we observe dirty prices with some error:

$$P_{i,t}^{obs} = P_{i,t} \exp(\eta_{i,t}/100), \tag{6}$$

where $\eta_{i,t}$ is a measurement error expressed in percent and is modeled as $\eta_{i,t} \sim N(0, \sigma(\eta_{i,t})^2)$. Parameter $\sigma(\eta_{i,t})$ will depend on many things, see section 3.

[3] We use a general; term "paper" to refer any fixed income debt instrument, money market instrument, deposit auctions, repo operations, etc.

Besides the description of how measurement errors in equation (6) are determined, the model is now completely specified. Table 1 lists the parameters of the model we have so far.

Parameters	Estimated?	Comment
β_{i,s_k}	yes	For $i = 0, 1, 2, 3$ and $k = 1, \ldots, N$
λ	no	Position of the first hump
$\sigma(\beta_i)$	no	Innovation size to β_{i,s_k} process
$\sigma(\eta_{i,t})$	no	Measurement error size of paper i in t

Table 1: Model parameters and their brief description

The estimation of β_{i,s_k} and the modeling of the measurement errors $\sigma(\eta_{i,t})$ are described in Section 3.

3 Model estimation

In this section we describe how β_{i,s_k} are estimated. Let B be a vector of all estimated β_{i,s_k}; Θ be a vector of all other (calibrated) parameters; Y be a set of all observations of all available papers i for days $t^{min} \leq t \leq t^{max}$. Then likelihood of the observed data in this model will be $L(Y|B, \Theta)$. As Θ are fixed, we search for B to maximize the fit (density).

$$p(B|Y, \Theta) \propto p(B|\Theta)L(Y|B, \Theta) \tag{7}$$

That means we optimize B to attain:

$$\max_B p(B|\Theta)L(Y|B, \Theta) \tag{8}$$

The likelihood and density of B parameters are equivalent. The density is equal to the density of implied innovations in their respective independent distributions

$$p(B|\Theta) = \prod_{i=0}^{3}\prod_{k=2}^{N} \frac{1}{\sigma(\beta_i)\sqrt{2\pi}} e^{-\widehat{\varepsilon_{i,k}}^2/(2\sigma(\beta_i)^2)}, \tag{9}$$

where $\widehat{\varepsilon_{i,k}}$ are realizations of the innovations given B.

The log-likelihood is then given as:

$$\begin{aligned} \log(p(B|\Theta)) &= -\frac{3(N-1)}{2}\log(2\pi) - \sum_{i=0}^{3}\log(\sigma(\beta_i)) \\ &\quad - \sum_{i=0}^{3}\sum_{k=2}^{N} \frac{\widehat{\varepsilon_{i,k}}^2}{2\sigma(\beta_i)^2} \end{aligned} \tag{10}$$

The likelihood of observations is given by a distance of observed dirty prices from their theoretical ones scaled by the measurement error sizes, that is:

$$L(Y|B, \Theta) = \prod_{i,t}^{\text{all obs}} \frac{1}{\sigma(\eta_{i,t})\sqrt{2\pi}} e^{-\widehat{\eta_{i,k}}^2/(2\sigma(\eta_{i,t})^2)}, \tag{11}$$

where $\widehat{\eta_{i,k}}$ is the actual distance from the theoretical dirty price

$$\widehat{\eta_{i,t}} = 100 \log(P_{i,t}^{obs}/P_{i,t}). \tag{12}$$

Thus, the estimation algorithm can be summed up as follows:

1. Process input data Y into a suitable structure
2. Get calibrated parameters Θ
3. Get initial guess of betas B_0
4. Optimize B starting at B_0 to attain maximum of $\log(p(B|Y, \Theta))$
5. Use gradient and Hessian of $\log(p(B|Y, \Theta))$ to boost iterations
6. Obtain optimal $\widehat{B}$
7. This implies historical and current yield curves $y_{t,\tau}$ for all $s_1 \le t \le s_N$

It is important to note that the estimation algorithm is a stacked-time estimation, that is, the estimation is done for all $s_1, \ldots, s_N$ simultaneously.

Information weights

This section describes the model for determination and calibration of measurement errors in equation (6). The model for measurement errors provides a way to influence a model behavior in response to the observations of various types. In other words, the model for measurement errors provides means of how to weight various observations and therefore how much the observations would influence the formation of the estimated curve.

We distinguish the following dimensions along which we can specify the information weight: residual maturity, market type, off-the run or on-the-run bonds, and volume. Equation (13) reflects these dimensions breaking the measurement error $\sigma(\eta_{i,t})$ for paper i in period t into the factors.

$$\sigma(\eta_{i,t}) = \sigma \cdot \phi_{i,t}^{mkt} \cdot \phi_{i,t}^{rm} \cdot \phi_{i,t}^{off} \cdot \phi_{i,t}^{spr} \cdot \phi_{i,t}^{vol}, \tag{13}$$

where:

- σ represents the weight between model and data,
- $\phi_{i,t}^{mkt}$ represents the weight depending on the market segment,

- $\phi_{i,t}^{rm}$ represents the weight according to residual maturity,
- $\phi_{i,t}^{off}$ represents the weight on/off-the-run bonds,
- $\phi_{t,i}^{spr}$ represents the weight according to the spread, and
- $\phi_{i,t}^{vol}$ according to the volume.

We describe the factors one-by-one. The factors take exponential form.[4]

Factor σ represents the weight between the model and data. This allows us to influence whether the estimated yield curve follows data more closely, or is more likely to smooth out the data and follow more closely the model.

σ is modeled as

$$\sigma = e^{\alpha}, \tag{14}$$

where α is a chosen constant. The choice of the exponential form is explained below. If α is increased, then we put more weight on the model and the yield curve is more likely to be smoothed out over the time. If α is decreased, then we put more weight on the data and the yield curve will tend to follow the data more closely, especially in periods of fast changes.

[4]The choice of exponential form is motivated by an ease of interpretation of the parameters. In order to illustrate this, let $\alpha_1$and $\alpha_2$be two exponents of the factors, therefore $\sigma(\eta_{i,t}) = \sigma \exp(\alpha_1 + \alpha_2)$. Note that the likelihood equation (11) involves actual error scaled by the standard error which can be approximated:

$$\frac{\widehat{\eta_{i,t}}}{\sigma(\eta_{i,t})} = \frac{\widehat{\eta_{i,t}}}{\sigma \exp(\alpha_1 + \alpha_2)} \approx \frac{\widehat{\eta_{i,t}}(1 - \alpha_1 - \alpha_2)}{\sigma}$$

It means that α parameters can be interpreted as effective decrease of actual error $\widehat{\eta_{i,t}}$ in fractions and that these fractions are additive.

Factor $\phi_{i,t}^{mkt}$ represents the weight of information coming from various market. The idea is to give higher weight to transactions from the money market than other markets, or more weight on primary than secondary markets. The market factor is

$$\phi_{i,t}^{mkt} = \begin{cases} e^{\alpha_{pmbill}} & \text{for the primary bill market} \\ e^{\alpha_{pmbond}} & \text{for the primary bond market} \\ e^{\alpha_{sm}} & \text{for the secondary market} \\ e^{\alpha_{quot}} & \text{for the secondary market quotes} \\ e^{\alpha_{intbnk}} & \text{for the interbank market} \\ e^{\alpha_{repo}} & \text{for the NBM repo operations} \\ e^{\alpha_{cd}} & \text{for the NBM CD auctions} \\ e^{\alpha_{judg}} & \text{for the judgmental observations} \end{cases}, \qquad (15)$$

where α_{cd} is the chosen penalty for the CD rates and similarly other alphas. The judgmental observations market is currently not used, but in general might be used for any judgmental observations that require the highest market weight.

Factor $\phi_{i,t}^{rm}$ represents the weight according to the residual maturity. The fundamental goal is to impose same measurement error if expressed in yield per annum. In addition, it is possible but only optional to impose a smaller weight on papers with long residual maturities. Let m be a residual maturity expressed in years of paper i in period t. The factor can be completely switched off, or take the form:

$$\phi_{i,t}^{rm} = \begin{cases} \left(\frac{1}{4}e^{-4m} + m\right) e^{m \cdot \alpha_{rm}} & \text{if switched on} \\ 1 & \text{if switched off} \end{cases}, \qquad (16)$$

where $\alpha_{rm} \geq 0$ is additional penalty for each year of the residual maturity.

Note that $\phi_{i,t}^{rm}$ scales measurement error $\sigma(\eta_{i,t})$ in equation (6) which is expressed as percent of the dirty price. The linear term m in equation (16) scales measurement error into the yield per annum. For example 1% of price difference in 1Y residual maturity is equivalent to m x 1 percent difference for residual maturity m in terms of equivalent error in the yield per annum.
However, keeping only m as the linear term would imply that the

measurement error of the price would converge to zero with residual maturity m going to zero. In practice, the measurement errors do not converge to zero for a fact as simple as a limited number of significant digits used for quoting prices. Another reason is that there is a thin market for instruments with very short maturities as there can be some nontrivial transaction costs close to maturity date.

Therefore we add the term $1/4e^{4m}$ in order to guarantee that $\phi_{i,t}^{rm}$ is 0.25 if residual maturity approaches to zero. Finally, the multiplicative term $e^{m \cdot \alpha_{rm}}$ represents additional penalty for each year of the residual maturity, but if α_{rm} is negative, then the model imposes more weight on longer maturity instruments.

Factor $\phi_{i,t}^{off}$ represents the weight of on-the-run bonds versus off-the-run bonds.

The idea is to penalize observations of off-the-run bonds. Let Δm_{min} be a chosen number of years. Let m denote a residual maturity of paper i observed in time t. The bond i is considered off-the-run in period t if there is a newer paper (not necessarily t-bond) with a residual maturity not bigger than $m + \Delta m_{min}$. Then the off-the-run bond factor is:

$$\phi_{i,t}^{off} = \begin{cases} e^{\alpha_{off}} & \text{if } i \text{ is off the run in } t \\ 1 & \text{if otherwise} \end{cases}, \qquad (17)$$

where $\alpha_{off} \geq 0$ is the chosen penalty for off-the-run bonds.

Factor $\phi_{t,i}^{spr}$ weights according to the observed spread. The idea is to give a smaller weight to transactions with higher spreads.

Let S^j be a spread (in percentage points of the annual yield) typical for transactions on the market j. Let $S_{t,i}$ be the observed spread of the observed data (auction results, or quotes, etc.). Then the transaction spread factor is

$$\phi_{t,i}^{spr} = e^{\alpha_{spr,j}(S_{t,i} - S^j)/100}, \qquad (18)$$

where $\alpha_{spr,j} \geq 0$ is the chosen elasticity for the market type j.

Factor $\phi_{i,t}^{vol}$ represents the weight according to the observed transaction

volumes. Here, the idea is to give a smaller weight to transactions with lower volumes. Let V^j be a typical volume for transactions on the market type j. Let $V_{i,t}$ be the observed transaction volume of paper i in period t. Then the transaction volume factor is:

$$\phi_{i,t}^{vol} = \begin{cases} \left(V^{pmbill}/V_{i,t}\right)^{\alpha_{vol,pmbill}} & \text{for the primary bill market} \\ \left(V^{pmbond}/V_{i,t}\right)^{\alpha_{vol,pmbond}} & \text{for the primary bond market} \\ \left(V^{sm}/V_{i,t}\right)^{\alpha_{vol,sm}} & \text{for the secondary market} \\ \left(V^{intbnk}/V_{i,t}\right)^{\alpha_{vol,intbnk}} & \text{for the interbank market} \\ \left(V^{cd}/V_{i,t}\right)^{\alpha_{vol,cd}} & \text{for the NBM CD auctions} \\ \left(V^{repo}/V_{i,t}\right)^{\alpha_{vol,repo}} & \text{for the NBM repo operations} \end{cases}, \tag{19}$$

where $\alpha_{vol,j} \geq 0$ is the chosen elasticity for the market type j.

4 The publication of yield curve

The yield curve is estimated based on the results of the most recent MDL-denominated state securities transactions in the primary, secondary markets and quotes published by primary dealers, interbank transactions, NBM certificate operations and NBM repo operations, every Wednesday. The yield curve is published weekly, on Wednesday, until 14:00 on the official website of the National Bank of Moldova.

If the estimation and publication of the yield curve fall on a holiday, these processes are carried out in the next working day.

The yield curve is estimated and published by the Financial Markets Department.