

 <p>MAGKS</p> <p>Aachen Siegen Marburg Gießen Kassel Göttingen</p>	<p>Joint Discussion Paper Series in Economics</p> <p>by the Universities of Aachen · Gießen · Göttingen Kassel · Marburg · Siegen</p> <p>ISSN 1867-3678</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

No. 39-2018

Vahidin Jeleskovic and Anastasios Demertzidis

**Comparing different methods for the estimation of
interbank intraday yield curves**

This paper can be downloaded from
<http://www.uni-marburg.de/fb02/makro/forschung/magkspapers>

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Comparing different methods for the estimation of interbank intraday yield curves

Vahidin Jeleskovic * and Anastasios Demertzidis **

ABSTRACT

In this paper, we compare three different models, namely the Nelson-Siegel model, the Svensson model and the Diebold- Li model, for the estimation of an intraday yield curve on the Italian interbank credit market e-MID. Using a sample which spans from October 2005 until March 2010, the first important finding is that all three models are highly suitable for the estimation of an intraday yield curve providing superior empirical results when compared with similar works on e-MID. The second important finding is that, based on different in sample statistics, the Svensson model dominates the other two models before, during and after the financial crisis from 2007. Moreover, the Nelson-Siegel model seems to dominate the Diebold- Li model although these differences in goodness-of-fit between these two models may not be statistically significant.

Keywords: Interbank credit market, e-MID, intraday yield curve, Nelson-Siegel model, Svensson model, Diebold- Li model.

JEL Codes: C12, C13, E43, G01

* Department of Economics, Quantitative Methods in Economics, University of Kassel, Nora-Platiel-Str. 4, D-34109 Kassel. Email: Jeleskovic@uni-kassel.de

** Department of Economics, Quantitative Methods in Economics, University of Kassel, Nora-Platiel-Str. 4, D-34109 Kassel. Email: Demertzidis@uni-kassel.de

1. Introduction

Is there an implicit intraday interest on interbank credits? This question has been assessed recently in different papers. Jurgilas and Žikeš (2013) and Merrouche and Schanz (2010) in the UK and Furfine (2001 and 2002) in the US. By using linear models, they found out that there is a downward trend in the intraday interest rate, meaning that the interest rates in the analyzed interbank credit markets are higher in the morning and lower in the afternoon. In all these studies, authors stress that these results are in line with the theoretical argumentation given by themselves. Abbassi et al. (2017), base their analysis on secured funding data and use a linear model as well. They find out that after the start of the financial crisis, the intraday term structure of interest rates may not be only monotone falling during a day.

Regarding the e-MID market (Mercato Interbancario dei Depositi), the only electronically organized interbank credit in the Euro area and in the US, different studies focus on the estimation of intraday term structure in different periods. Angelini (2000) was the first one to analyze the intraday behavior of interest rates on the e-MID market. Using a linear model for the intraday interest rates and based on hourly means of the intraday interest rates in the period from July 1993 to December 1996, he finds only very weak evidence for an existing downward intraday term structure.¹ This low evidence is shown in the estimated term structure where the difference of the interest rate in the morning and in the afternoon differs only to a very small degree. Based on his premise, the main force of the intraday interest rates are variations in the market liquidity.

Baglioni and Monticini (2008), apply also a linear model using hourly means to estimate the intraday term structure in the sample from January 2003 until December 2004. They find weak statistical evidence for a downward trend in the intraday structure which is also reflected in a relatively small difference between interest rate in the morning and in the afternoon. They state that the main drive behind these movements is the higher credit risk, in terms of the counterparty risk, in the morning rather than in the afternoon.

Using two data samples from 11th of July to 10th of September 2007, Baglioni and Monticini (2010) redo their analysis from the year 2008. In this second analysis, they find evidence for a downward trend in the intraday term structure which becomes steeper after the outbreak of the financial crisis in 2007. In addition, here they state that these facts can be observed due to higher credit risk in the morning than in the afternoon.

¹ When mentioning a linear model, we refer to linear regression model.

Baglioni and Monticini (2013) also estimate an intraday term structure on the e-MID market, using three different extended linear models, based on the difference of the average of the interest rates between 09:00 a.m. and 01:00 p.m., called the morning rate, and the average of interest rate between 02:00 p.m. and 06:00 p.m. called the afternoon rate. By using a sample ranging from January 2007 to April 2009 they again find evidence for a downward trend in the term structure of interests. Based on their models this downward trend becomes even steeper after the outbreak of the financial crisis in August 2007 and the steepest after the collapse of Lehman Brothers in September 2008. They also argue that intraday interest differs from the morning to the afternoon due to higher counterparty credit risk as well as due to market liquidity constraints. Furthermore, they state that the interest rates may be influenced by incoming news in this in this particular period.

Furthermore, Demertzidis and Jeleskovic (2016) introduced the concept of the spot intraday yield curves (SIYC-s) and differ from the previous studies in two major points, namely the use of tick- by- tick interest rate data and the use of a nonlinear model. For the time period from October 2005 to March 2013, they showed that the SIYC can be modeled and estimated by a standard nonlinear model which is used by many researchers and central banks (Diebold and Rudebusch, 2013), namely by the Nelson-Siegel model (hereafter NSM for Nelson-Siegel model). The authors achieve an R^2 of up to 0.424 on average, which is remarkably high since they use tick-by-tick data. The authors conclude that one should move from the assumption of linear models for the estimation of SIYC towards explicit modelling of the nonlinear dynamics. The second very interesting empirical result is that the goodness-of-fit become significantly higher after the outbreak of the financial crisis. Thus, one should expect higher nonlinear systematic dynamics of yield curves during turmoil on interbank credit markets. The authors attribute this fact to the more intensive process of incoming news within a day during the financial crisis.

The NSM has been modified and extended by many researchers. Among others, Bliss (1996) with his three-factor model interpretation, Björk and Christensen (1999), with their five factor NSM, Christensen et al. (2009) and Christensen et al. (2011) with their arbitrage free interpretation of the NSM and Chen and Niu (2014) with their adaptive dynamic NSM, modified and / or extended the model.

One important model modification improves the original NSM significantly from the theoretical as well as from a practical point of view is proposed by Svensson (1994) (hereafter SVM for Svensson model). The major highlight of the SVM is modeling a second hump in the

yield curve. This model is used for the estimation of the yield curve by many central banks, including the ones of Germany, Norway, Spain, Sweden and Switzerland (BIS, 2005). According to De Pooter (2007), this model should be used when estimations of a larger variety of yield curves or more complex dynamics of the yield curves is necessary. Hence, this model should be used in times of higher volatility, e.g. in times of a financial crisis.²

The SVM is also used by many researchers for the estimation of the yield curve for different markets. Among others, Schich (1997) for the German bond market, Clare and Lekkos (2000) for the bond yield curves in the US, Germany and the United Kingdom (UK) and Gürkaynak et al. (2007) for the US bond market, use the SVM for the estimation of the yield curve.³

Another popular modification of the NSM is the Diebold and Li (2006) model (hereafter DLM for Diebold- Li model), which also has been used widely in practice and theory. Mönch (2008) e.g. in his study confirms that the model from Diebold and Li provides a good statistical fit for a variety of yield curves.

Among others, Tam and Yu (2008) for the US, the Japanese and the German bond market and Afonso and Martins (2012) for the United States and Germany use the DLM for the estimation of the yield curve. Furthermore, this model is also used from a practical point of view in different studies, e.g. to model and forecast the term structure of futures on oil contracts (Grønberg and Lunde, 2016).⁴ Besides the different studies of yield curve estimations, many analyses focus on the comparison of different yield curve estimation methods. These studies try to, empirically, find out which model suits the best under different conditions and different markets and countries.

Csajbok (1999) compares different estimation methods for the yield curve, including different spline-based methods as well as the NSM and the SVM, for the Hungarian bond market. One of his key findings is, that the SVM is superior to the NSM and different spline-based methods for the estimation in Hungary. This may be because according to Csajbok the SVM is able to capture a more complex variety of yield curves. Ganchev (2009) models and estimates the spot rates for the Bulgarian bond market. In his study he uses different estimation methods including

² Angelini et al. (2011) state, that interbank credit market rates become more volatile in times of a crisis.

³ Besides the original NSM, many researchers have modified the SVM as well. E.g. De Rezende and Ferreira (2008), propose a five-factor model, Christensen et al. (2009) present a dynamic version of the model and De Rezende (2011) presents a six-factor model. However, many of these models are mostly not used from a practical point of view.

⁴ The DLM has been extended/ or modified. Laurini and Hotta (2010) extend the model through a Bayesian estimation method using the Markov Chain Monte Carlo Simulation. Bernadell et al. (2005) present a regime-switching extension of the DLM by linking expectations of different macroeconomic variables to the estimated yield curve.

also the NSM and the SVM. One major finding is, that the NSM has a poorer performance than the SVM. Aljinović et al. (2012) focus in their study on the comparison between the NSM and the SVM for the estimation of the yield curve on the Croatian financial market. They find out that the Svensson model is superior to the Nelson-Siegel model. Moreover, Ioannides (2003) uses different spline-based models, the NSM and the SVM in order to estimate the yield curve in the UK. By estimating the yield curve with different methods, he shows, that the SVM and the NSM outperform the other used spline-based methods. By comparing the SVM with the NSM model, he points out that the SVM is more suitable than the NSM for the yield curve estimation in the UK.

To the best of our knowledge, no study or analysis has focused on the comparison of different nonlinear models and methods for the estimation of yield curve for an interbank credit market, neither on an intraday day basis, nor for higher maturities.

Due to the importance and empirical validity of the previously described three models, the goal of the paper is manifold: first, we aim to find out, whether the NSM, the SVM and the DLM are able to model the SIYC. The second purpose is to discover which model is the most suitable for estimating the SIYC. Using a sample from October 2005 to March 2010 we also put focus on the different states of the interbank credit markets by dividing our sample into different sub-periods according to different relevant events during the financial crisis starting in 2007. Hence, the importance and the consequences of the financial crisis are explicitly considered.

Following e.g. Angelini (2000) and Baglioni and Monticini (2008, 2010), who use one-hour intervals for the estimation of an intraday term structure on e-MID by applying the linear regression with hourly dummies, we also construct the SIYC over intraday time intervals. However, we do not use one-hour intervals but 30-minute intervals, meaning 30-minute averages for the interest rates.⁵

The paper is organized as follows: After the introduction, we present in section two our data sample and the main descriptive statistics. Section three describes the applied models. In section four we present the empirical results. Here, we first examine whether each model is capable of modeling the SYIC and in the second part we perform the model comparison. In the last part of the section we discuss our empirical results. Section five concludes.

⁵ The use of different intervals instead of tick-by-tick data is used in different studies focusing on limit order books see e.g. Kempf and Mayston (2005) and Hautsch and Jeleskovic (2008) for financial markets. Moreover, Engler and Jeleskovic (2016) apply the Multivariate Multiplicative Error Model to analyze the order book data on e-MID using 5-minute intervals.

2. e-MID and descriptive statistics

The trading activity on the market begins each day at 08:00 a.m. and ends at 06:00 p.m. During this time credits with a minimum credit value of 50.000 euro can be traded. The maturity of credits ranges from overnight credits (ON) up to one year.⁶

During the transaction process the duration, the interest rate, the specific time and the amount of each credit are known. Furthermore, also the Quoter (bank which puts the order for the transaction in the limit order book) and the Aggressor bank (bank which selects and accepts the specific credit transaction) are known due to a specific code which consists of two letters, referring to the country of origin and four digits which refer to the specific bank.

The exact time of repayment may not be known exactly, but the maximum maturity of the ON credits is predefined by the system itself. If an Italian bank is involved in the credit transaction, either as a Quoter or as an Aggressor, the latest repayment time point of the ON credit is at 09:00 a.m. the next day. If no Italian bank is involved the latest repayment time is at 12:00 (noon) the next day.

For our analysis we use a data sample starting on 03.10.2005 up to 31.03.2010. This is a large sample and includes times before, during and after the financial crisis of 2007 and contains 377745 overnight transactions.

As pointed out in many studies (see e.g. Baglioni and Monticini, 2008 and Baglioni and Monticini, 2010) in the time band between 08:00 a.m. and 09:00 a.m., the trading activity in the e-MID market is very low in terms of volume and number of transactions. Thus, it can be characterized as not sufficient in this particular daily time period. This fact can also be observed in our data sample. Only 5829 overnight transactions occur during the time between 08:00 a.m. and 09:00 a.m., which are approximately five transactions per day, in the whole sample period.

Furthermore, as stated by Gürkaynak et al. (2007) the estimation of the yield curve behaves oddly based on securities with a very short maturity. According to their analysis this fact can be observed due to the relative low liquidity of securities with low maturity. As pointed out by Angelini (2000) this fact can also be observed in the e-MID. We can observe this trend also in our whole data sample. During the time band between 05:00 p.m. and 06:00 p.m. , only 5975 transactions occur during the whole sample period, meaning that only approximately five overnight transactions per day take place in the market during this daily time period.

⁶ The ON segment represents more than 90% of the credit transaction in terms of volume and number of trades.

By considering these two facts, meaning a small number of transactions and low volume between the time bands 08:00 a.m. and 09:00 a.m. and 05:00 p.m. and 06:00 p.m. we focus our estimations for the SYIC during the time between 09:00 a.m. . and 05:00 p.m., which is in line with the previous studies as mentioned above. Thus, in our analysis 365941 out of 377745 overnight transactions in the sample period are considered, stating for 96.88 % of all the overnight transactions in the sample period in the e-MID.

Out of these overnight transactions, in 345105 transactions at least one Italian bank was involved, either as a credit lender or as a borrower within the transaction. This represents 94.31% of all ON transactions. In the remaining 20836 overnight transactions no Italian bank was involved. These credits were completed between foreign banks, accounting for 5.69 % of all ON transactions.

Following different studies, (e.g. Gabbi et al., 2012 and Demertzidis and Jeleskovic, 2016) we separate our data sample into four periods. This is done, due to the fact, that our interest goes further than the simple analysis of the suitability of the different models. We are interested in finding out whether the models are capable of estimating the SIYC in different sub-periods and which model performs the best in the in different periods, and different states of the market, before, during and after the financial crisis. Hence, splitting up our sample in this way enables this kind of analysis.

The first period, which we call the pre-crisis period, starts on 03.10.2005 until 08.08.2007 - one day before the onset of the global financial crisis. The second period ranges from 09.08.2007, the onset of the crisis, up to the 14.09.2008, one day before the collapse of the bank Lehman Brothers. Hence, we define it as the first crisis period. The third period ranges from 15.09.2008 until 12.05.2009, one day before the last reduction of the key interest rate by the European Central Bank (ECB). We call this period the second period of the crisis. The last period ranges from 13.05.2009 until the end of the sample on the 31.03.2010. This period can be called the after-crisis period.⁷ The different periods for our estimations are summarized in table 1.

⁷ Brunetti et al. (2015) refer to the period from April 2009 to March 2010 as the after-crisis period.

Table 1: Presentation of the sub-periods

Period 1	03.10.2005-08.08.2007	Period before the crisis
Period 2	09.08.2007-14.09.2008	Outbreak of the crisis until the collapse of Lehman Brothers
Period 3	15.09.2008-12.05.2009	Lehman Brothers collapse until reduction of key interest rate
Period 4	13.05.2009- 31.03.2010	Key interest rate reduction until the end of the observation period

The main descriptive statistics for the credit transactions considered in our data sample are summarized in the tables 2- 5.

Table 2: Descriptive statistics: days and observations ⁸

	Whole Sample	Period 1	Period 2	Period 3	Period 4
Number of days	1641 (1149)	675 (473)	403 (281)	240 (166)	323 (229)
Transactions	365941	182876	97281	41858	43926
Mean of transactions per day	318.49	386.63	346.19	252.16	191.82

Based on table 2 we can see that, the mean number of transactions in the whole sample is 318.49 trades per day. What is more interesting is that the number of trades is the highest before the crisis (period 1) and starts do drop slowly with the onset of the financial crisis in the second period. This trend becomes even more acute in period 3, the second crisis period, where the mean number of transactions drops dramatically, resulting in an even sharper drop in the number of transactions in period 4 in our data sample.

⁸ In parentheses: effective trading days, excluding weekends and holidays.

Table 3: Descriptive statistics: volume (in Million Euros)

	Whole Sample	Period 1	Period 2	Period 3	Period 4
Daily average	13116.69	19779.45	13000.34	6977.88	3947.49
Mean per Transaction	41.18	51.16	37.55	27.67	20.58

Regarding the descriptive statistics in terms of volume, we can state that, the trading volume, as daily average volume and mean per transaction, follow the same trend as the number of trades in table 3. We see that the volume is the highest before the crisis, drops in periods 2 and more in period 3. The lowest volume per day and per transaction is found in period 4.

Table 4: Descriptive statistics: interest rates

	Whole Sample	Period 1	Period 2	Period 3	Period 4
Mean	2.605	3.050	4.036	2.029	0.355
Std. Dev.	0.032	0.014	0.0389	0.081	0.027

Considering the descriptive statistics of the interest rate, which are calculated over half hour time intervals, we can see that the mean of the interest rate is quite high in period 1 and the highest in period 2. After the culmination of the financial crisis the interest rate dropped in period 3 and even more in period 4. Regarding the standard deviation, one can see that the smallest grade of variation of the interest rates is observed in the first period in our data sample whereas the highest one is in the second period. After the second period the standard deviation is successively declining in the periods 3 and 4. These results for the standard deviation rely on the fact that before the outbreak of the financial crisis the dynamic of interest rates is quite flat.⁹ On the other hand, this implicates that the strongest variation in the dynamic of interest rates can be assumed in the period 2.

⁹ This fact was also observed by Baglioni and Monticini (2013)

As already discussed by Demertzidis and Jeleskovic (2016) the market functions properly before the crisis and in the first period of the crisis. They further state that the market does not function properly in periods 3 and 4, meaning that the effective allocation of credits is no longer possible. This effect is also supported in our data sample in terms of volume and number of trades.

3. Methodology of the SYIC estimation

As already mentioned above and in contrast to previous studies, we use 30- minute intervals for the estimations of the SICY-s. There are at least two reasons to use half hour intraday intervals. First, the construction of the SIYC becomes more precise, and thus, the estimation of the SIYC as well. Second, from the practical point of view the traders on the e-MID may be more interested in the nowcasting of the interest rate in shorter time intervals due to the fact that they trade more frequently within the intraday time domain. Hence, in our opinion the use of half-hour intervals is an appropriate solution for the tradeoff between avoiding the noise in the tick-by-tick data and the practical advantage of not-using intervals which are too long.

Therefore, in our data sample we use 14 mean interest rates per day, meaning 14 intervals, starting from 09:00 a.m. - 09:30 a.m. which represents the first interval at the time stamp 09:30 a.m. , until the time band from 04:30 p.m.- 05:00 p.m. which represents the last intraday interval for 05:00 p.m.

To estimate the empirical SIYC, it is necessary to define the maturity of each credit transaction in our data sample. We calculate the maturity of each credit interval as the difference between the time stamp of the particular half hour interval within a day and 06:00 p.m., when the market closes on each day. Thus, because we use the pre-described time bands between 09:30 a.m.- 05:00 p.m. , the maximum maturity is 8.5 hours (09:30 a.m.- 06:00 p.m.) and the minimum maturity is one hour (05:00 p.m.- 06:00 p.m.).¹⁰

We can state that these time intervals of 30 minutes generate a high number of observations needed for the empirical analysis of the SIYC on the e-MID market. We point out that we also estimated the yield curve using time intervals of one, five and fifteen minutes and within the interval of one hour. However, the results do not differ qualitatively and are even slightly worse in terms of quantitative results.¹¹

¹⁰ The first interval lies between 09:00 a.m. and 09:30 a.m.

¹¹ The problem occurs by the use of shorter time intervals that in a certain number of intervals there are no credit transactions. This may cause some artifacts and impact negatively the numerical optimization. However, these results can be provided on request.

3.1 The Nelson- Siegel Model

Nelson and Siegel (1987) propose the following equation for the estimation of the spot rate R of different maturities (m):

$$R(m) = \beta_0 + \beta_1 \frac{1-e^{-\frac{m}{\tau}}}{\frac{m}{\tau}} + \beta_2 \left(\frac{1-e^{-\frac{m}{\tau}}}{\frac{m}{\tau}} - e^{-\frac{m}{\tau}} \right) \quad (1)$$

Where β_0 , β_1 , β_2 specify the parameters to be estimated and τ denotes the time constant associated with the equation.

β_0 , is a constant. For a maturity which is approaching infinity, the spot rates converge to this value. The second term $\beta_1 \frac{1-e^{-\frac{m}{\tau}}}{\frac{m}{\tau}}$, refers to the slope of the specific yield curve and the third term of the model $\beta_2 \left(\frac{1-e^{-\frac{m}{\tau}}}{\frac{m}{\tau}} - e^{-\frac{m}{\tau}} \right)$, is important for the modeling of a hump or a U- shape in the yield curve. In our case R is the mean of interest rates within a half hour interval and m is the maturity defined as above.

The estimation of the NSM relies on the same procedure as in Demertzidis and Jeleskovic (2016). We estimate each parameter of the NSM by fitting $R(m)$ based on formula (1). During this process we apply a numerical optimization where we apply an objective function over τ , whereas each parameter is estimated simultaneously in each optimization step using the ordinary least squares (OLS) method. During our analysis we use the `fminbnd` function for our optimization process, with default settings. The optimization bounds for τ lie between 0 and 10000 during our estimations.

3.2 The Svensson Model

In order to increase the goodness-of-fit and the flexibility of the yield curve Svensson (1994) extended the NSM by adding a fourth term. By adding this fourth term, it is possible to model a second hump, or a second U- Shape, in the yield curve (Svensson 1994). He validates his findings by estimating the yield curve of Swedish government bonds in the time between May 1992 and June 1994.

For the estimation of the spot rate R , with a yield to maturity denoted m , Svensson uses the equation:

$$R(m) = \beta_0 + \beta_1 \frac{1-e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1-e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} - e^{-\frac{m}{\tau_1}} \right) + \beta_3 \left(\frac{1-e^{-\frac{m}{\tau_2}}}{\frac{m}{\tau_2}} - e^{-\frac{m}{\tau_2}} \right) \quad (2)$$

Where b is: $\beta_0, \beta_1, \beta_2, \beta_3$ are the parameters of the estimated yield curves and the parameters τ_1 and τ_2 are the time constants of the model.

In this equation the term $\beta_3 \left(\frac{1-e^{-\frac{m}{\tau_2}}}{\frac{m}{\tau_2}} - e^{-\frac{m}{\tau_2}} \right)$ defines the second hump, or the second U shape in the yield curve and the parameter τ_2 the position of this positive or negative hump. All the other parameters, including their asymptotic properties, can be defined like the model proposed by Nelson and Siegel (Svensson, 1994).

In his model, Svensson uses the Maximum likelihood method in order to estimate the parameters. According to Svensson the estimated prices can be fitted to the actual (observed) prices also with the general method of movements and the nonlinear least squares method (Svensson, 1994).

In our case, we use the nonlinear least squares method where we apply the Matlab and the optimization toolbox. However, as Gilli et al. (2010) report there may be a significant problem with the objective function when optimizing the Svensson model. As the authors report, the optimization problem might be non-convex and there may be different local minima. To avoid these problems, we use at first the genetic algorithm. Having optimized the parameters in the Svensson model via the genetic algorithm, we take the optimal parameters to use them as starting values for the numerical optimization in the second step. We are convinced that this procedure will lessen the problem of starting values and local minima.

3.3 The Diebold- Li Model

Diebold and Li (2006) modified the original NSM and use at first a two-step estimation method for the parameter estimation. In their work they fitted the yield curve using a three-factor model based on the NSM. Equation (3) presents the three-factor model from Diebold and Li (2006).

$$y(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_{3t} \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right) \quad (3)$$

Diebold and Li interpret the parameters β_{1t}, β_{2t} and β_{3t} as latent dynamic factors which vary over time and thus, they are state-dependent. The loading on β_{1t} equals one, which can be viewed as the long-term factor. The term $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$ is the loading of the parameter β_{2t} , which starts at the value of one and guarantees a quick and monotonical decay towards 0. so, it can be interpreted as the short-term factor. The factor loading of β_{3t} , is $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}$. The value starts at 0,

increases in the beginning and then decays to zero, so it can be viewed as the medium-term factor (Diebold and Li, 2006).

Another important insight of the this extension is that the parameters β_1 , β_2 and β_3 can be interpreted in another way than in the original NSM. Diebold and Li interpret these parameters as the level, slope and curvature of the yield curve respectively (Diebold and Li, 2006).

The last parameter λ , which is $1/\tau$ in the original NSM, explains the exponential decay rate. When λ takes small values, it results in a slow decay, so the model can fit the yield curve better at long maturities. If λ takes large values the decay is faster, resulting in a better fit at short maturities. Besides the decay rate the parameter λ defines where the loading of β_3 achieves his maximum. In their work this parameter stays constant at the value of 0,0609 for every given t (Diebold and Li, 2006).

We do the estimation process based on Diebold et al. (2006). The Kalman filter method for the yield curve, due to the fact, that obtain better results with this method than with the original proposed two step method for the estimation.¹²

For the estimation of the SIYC using the DLM we use of the SSM econometrics toolbox in Matlab. Here for the state vector x_t and the observation vector y_t the parametric form us given by the following linear state- space functions:

$$x_t = A_t x_{t-1} + B_t u_t \quad (4)$$

$$y_t = C_t x_t + D_t \varepsilon_t \quad (5)$$

Here, u_t and ε_t are unit-variance white noise vector processes which are uncorrelated. In this representation the first equation is called the state equation and the second one is the observation equation. The parameters of the model, A_t , B_t , C_t and D_t are referred to as the state transition, state disturbance loading, measure sensitivity and observation innovation matrices, respectively.

The DLM is formulated in such a way that level, slope and curvature follow a VAR (1) or autoregressive process of first order and as such the model forms a state space system. As already mentioned, we use the interpretation of Diebold, Rudebusch and Aruoba, stating transition equation, which govern the dynamics of the state vector and it is written as:

¹² The results of the SIYC estimation of the DLM using the two-step method can be submitted upon request.

$$\begin{pmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{11} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{pmatrix} + \begin{pmatrix} \eta_t(L) \\ \eta_t(S) \\ \eta_t(C) \end{pmatrix} \quad (6)$$

Whereas the corresponding observation equation is written as:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_1) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & -e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & -e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & -e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} + \begin{pmatrix} e_t(\tau_1) \\ e_t(\tau_2) \\ \vdots \\ e_t(\tau_N) \end{pmatrix} \quad (7)$$

In the vector notation the DLM can be rewritten as the following state space system for the 3-D vector of mean-adjusted factors f_t and the observed yields y_t :

$$(f_t - \mu) = A(f_{t-1} - \mu) + \eta_t \quad (8)$$

$$y_t = \Lambda f_t + e_t \quad (9)$$

With the orthogonal, Gaussian white noise processes η_t and e_t are defined as following:

$$\begin{pmatrix} \eta_t \\ e_t \end{pmatrix} \sim WN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right) \quad (10)$$

In this setting, it is assumed that the stochastic terms for the state factor disturbances η_t are correlated leading to a non-diagonal covariance matrix Q which is non-diagonal. On the other hand, the diagonality of the covariance matrix H is assumed so that the deviations of the observed yields among all maturities are uncorrelated.

The latent states are to be defined as the mean-adjusted factors:

$$x_t = f_t - \mu \quad (11)$$

And the deflated or, intercept-adjusted yields as:

$$y'_t = y_t - \Lambda\mu \quad (12)$$

And then substitute into the equations above.

Thus, the DLM state-space system may be rewritten as:

$$x_t = Ax_{t-1} + \eta_t \quad (13)$$

$$y_{t'} = y_t - \Lambda_{\mu} = \Lambda x_t + e_t \quad (14)$$

$$\eta_t = Bu_t \quad (15)$$

$$e_t = D\varepsilon_t \quad (16)$$

$$Q = BB' \quad (17)$$

$$H = DD' \quad (18)$$

As already mentioned, the yields-only model forms a state-space system, with a VAR(1) transition equation where the dynamics of the vector of latent state vector variables are summarized, and a linear measurement equation relating to the observable yields to the state vector. For the estimation purposes, we use the SSM toolbox using the smoother algorithms and the default specifications given by this toolbox. Due to the often referred to problem of the sensitivity of Kalman filter estimator on starting values, we span a grid of starting value for parameter λ in the range between 0.00001 and 0.5 and took the best estimates.

4. Results

In order to verify, whether these models are suitable for estimating the SYIC and to compare their empirical performance in each sub-period, we evaluate the models using three different measures, namely the standard R^2 since it also has been used by Nelson and Siegel (1984), the Root Mean Squared error (RMSE), which has been used by Svensson (1994) to evaluate his model and the Mean Absolut Error (MAE).

An important fact of the RMSE is that it is based on the squared errors and thus sensitive to outlier in the error distribution. Hence, relatively higher weights are put on the tails of the error distribution using RMSE as a measure of goodness-of-fit. In an analogous way, this holds true also for R^2 .¹³ On the other hand, the Mean Absolute Error (MAE) may be quite robust to the outliers, and thus, has some advantages compared to the other two. Therefore, we can consider the MAE as a robust measure of the goodness-of-fit. Hence the both measures of the goodness-of-fit, namely the MAE and the RMSE, may behave differently when one uses them for the purposes of the measures of the model fit. Relying on this fact, we will use all three measures when we are analyzing the three models applied in this paper.

¹³ However, the R^2 takes additionally the variation of the dependent variable into account.

In the first part of this chapter, we evaluate each model separately by the three measures over the four periods to identify if there are significant changes. In the second part, we compare the three models with each other over periods, again based on those three measures, and in the third part we discuss our findings.

4.1 Empirical results for the comparison between the periods

4.1.1 Evaluation based on the R^2

We calculate the R^2 for each day using the following formula:

$$R^2 = 1 - \frac{\sum(r_i - \hat{r}_i)^2}{\sum(r_i - \bar{r}_i)^2} \quad (19)$$

where r_i stands for the mean interest rate of the i^{th} 30- minute interval within a day, \hat{r}_i stands for its estimate interest rate and \bar{r}_i stands for the mean of the all sixteen r_i -s on the particular estimation day. Thus, we have as many estimates of R^2 -s as the number of days we consider in a sample. After that, we can analyze the statistical properties of calculated R^2 -s. In the same way, we proceed with other two measures of fit.

Table (5) presents the results for the R^2 in the different periods for the estimations of the intraday yield curve in the e-MID market using the NSM.

Table 5: R^2 of the NSM

	Period 1	Period 2	Period 3	Period 4
R^2	0.6732***	0.7398***	0.6612***	0.6259***
Std. Dev.	0.2190	0.2023	0.1977	0.2050
t-statistic	66.847	61.294	49.092	46.201

*** Denotes significance on the 1% level

Hence, these results present the mean and the standard deviations of the R^2 -s, and the t-statistics for the means in each period, respectively. First given the high mean of R^2 in each period, we can state that the NSM is capable of modelling the SYIC in the e-MID market. Second, we can also state that, the best performance for the modeling of the SYIC can be found in period 2 with an R^2 of 0.7398, thus, this is in the offset of financial crisis. This is a first support of results given by Demertzidis and Jeleskovic (2016) that the best performance of the NSM may be achieved in period 2. Although distortion on the interbank credit markets where noticeable, the market was still functioning properly, as already mentioned above. The second-best

performance is found in period 1 which is the pre-crisis period where an R^2 of 0.6732 is achieved. The performance drops slightly in period 3, where we achieve a mean R^2 of 0.6612 which is quite similar to the period 1. Furthermore, the smallest mean R^2 of 0.6259 is achieved in period 4. Although the achieved means of R^2 are remarkably high, we use the standard t-test in order to find out whether these means are significantly different from zero:¹⁴

$$t = \sqrt{n} \frac{\bar{x} - \mu_0}{s} \quad (20)$$

where n stands for the number of observations, \bar{x} stands for the mean of the respective goodness-of-fit statistic, in this particular case of R^2 , s stands for the standard deviation of that specific goodness-of-fit statistic, and μ_0 is zero, since we test against zero.

Based on the t-test we can state, that the R^2 are significantly different from zero even at the 1% level in each sub- period.

To find out whether the means of the considered statistics for goodness-of-fit from the different periods of the same model¹⁵ are significantly different, we use the two-sample two-tailed t-test between the periods:

$$t = \frac{\sqrt{nm} (\bar{x}_1 - \bar{x}_2)}{\sqrt{n+m} s} \quad (21)$$

$$s^2 = \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2} \quad (22)$$

where n and m stand for the number of observations of the same statistics from two periods, respectively, which we want to compare statistically with each other. The two statistics are in our case the means of the particular measure of fit, here \bar{x}_1 and \bar{x}_2 , and can stem from two different periods for the same model. s_1^2 and s_2^2 are the estimated variances of \bar{x}_1 and \bar{x}_2 . The results of the two-sample t-test between each period for the NSM are presented in table 6.

Table 6: Two-sample t-test of R^2 for the SIYC-s estimated by the NSM

	Period 2	Period 3	Period 4
Period 1	-4.151***	0.622	2.736***
Period 2		4.001***	6.283***
Period 3			1.714*

***, **, * Denote significant different means at the 1%, 5% and 10% level respectively.

¹⁴ We use this test also for the analysis of the MAE and the RMSE.

¹⁵ Or of two models from the same period.

The difference between period 2 and all other periods is significant even at the 1% level. Hence the NSM achieves significantly the best performance in period 2. In addition, the difference between period 1 and period 4 is highly significant, whereas the difference between period 1 and period 3 is statistically not significant. Between period 3 and 4 we can state significantly different means of R^2 only at the 10 % level. These results for the NSM are in line with the results provided by Demertzidis and Jeleskovic (2016). Therefore, the same economic discussion given by Demertzidis and Jeleskovic (2016) regarding their results for the NSM also holds in the case of results in this paper for the NSM when comparing different periods. Thus, the main conclusion is that the best performance of NSM is achieved in the period of the onset of the financial crises with a proper functioning interbank credit market.¹⁶

The means of R^2 for the SVM for the different sub-periods are presented in table 7.

Table 7: R^2 of the SVM

	Period 1	Period 2	Period 3	Period 4
R^2	0.7863***	0.8243***	0.7517***	0.7176***
Std. Dev	0.1740	0.1682	0.1723	0.1892
t-statistic	97.566	82.121	56.202	57.400

*** Denotes significance on the 1% level

Based on the means of R^2 we can state that the SVM is also suitable for an empirical estimation of the SYIC in the e-MID market. These means of R^2 follow the same tendency as in the previously presented results for R^2 of the NSM. The best performance of the SVM can be found in period 2, with a mean R^2 of 0.8243. The second best can be found in period 1 of 0.7863 whereas in period 3 the mean drops to 0.7517 and even more in period 4 where we achieve a mean R^2 of 0.7176. Moreover, these means are significantly different from zero at the 1% level as well, as the t-test indicates.

Regarding the significance among the different periods, we can state that the results of the two-sample t-test of the SVM, presented in table 8, are also qualitatively the same as in the case of the NSM. The only difference here is that the difference between the mean of period 1 and period 3, is significant at the 10 % level. This fact does not change qualitatively the implication of the results which can be taken over as for the NSM.

¹⁶ For further information see Demertzidis and Jeleskovic (2016). This conclusion holds also for the SVM and the DLM, as will be presented below.

Table 8: Two-sample t-test of R^2 for the SIYC-s estimated by the SVM

	Period 2	Period 3	Period 4
Period 1	-3.375***	1.847*	4.367***
Period 2		4.369***	6.732***
Period 3			1.832*

***, **, * Denote significant different means at the 1%, 5% and 10% level respectively.

The mean of R^2 for the DLM for the different periods can be found in table 9 where additionally the estimated mean of the λ parameter is presented.

Table 9: R^2 of the DLM

	Period 1	Period 2	Period 3	Period 4
R^2	0.6025***	0.6823***	0.6444***	0.6020***
Std. Dev	0.2491	0.2353	0.2057	0.2224
t-statistic	52.589	48.605	40.364	40.964
$\bar{\lambda}$	0.0007	0.1681	0.0407	0.0400

*** Denotes significance on the 1% level

First, we can state that also the DLM, like the NSM and the SVM, is capable of estimating the SIYC for the e-MID market. Also, like the NSM and the SVM the best performance can be found in period 2 where we achieve a mean of R^2 of 0.6823. In contrast to the previous models, the second-best performance cannot be found in period 1 but in period 3 with an R^2 of 0.6444. Furthermore, the results of the periods 1 and 4 are quite similar and lower than in the previously described periods, where we achieve an R^2 of 0.6025 and 0.6020 respectively. Like the previous two models, the R^2 of the DLM in each sub-period are statistically different from zero at the 1% level. Hence, the DLM is suitable for the modeling of the SIYC as well. These similarities in the findings are also supported by the two-sample t-test for the DLM presented in table 10.

Table 10: Two-sample t-test of R^2 for the SIYC-s estimated by the DLM

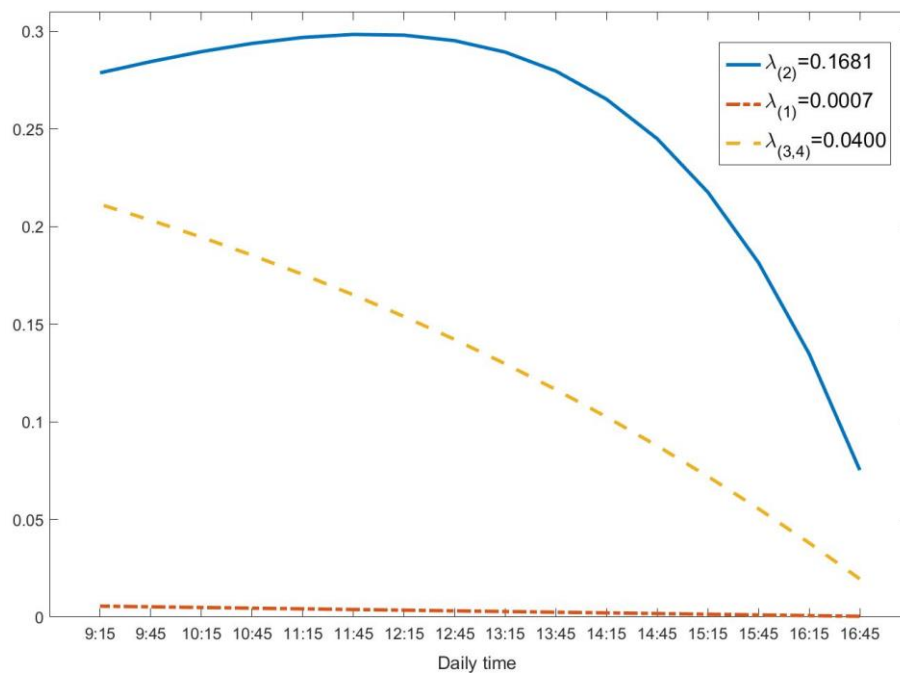
	Period 2	Period 3	Period 4
Period 1	4.344***	-1.947*	0.21
Period 2		1.724*	3.927***
Period 3			1.927*

***, **, * Denotes significant different means at the 1%, 5% and 10 % level respectively.

Based on this test, we can state that the differences between period 1 and 2 and for the periods 2 and 4 are significant also at the 1% level. The differences between period 1 and 3, period 2 and 3 and period 3 and 4 are significant at the 10% level whereas the difference between period 1 and 4 are not significant even at the 10% level. Thus, the best model performance can also be found here in period 2 and the worst one in period 4. Furthermore, the most important economic discussion regarding the goodness-of-fit in period 2 given by Demertzidis and Jeleskovic (2016) holds also for DLM. The only difference to the qualitative results achieved through NSM and SVM lies in the comparison between the periods 1 and 3 where, however, this difference between these two periods is significant only on the 10% level. Hence, we deem this evidence to be a matter for our future interest.

At this point, we tackle the same discussion about the curvature as in Diebold and Li (2006) which correspond to the following figure 1.

Figure 1: Curvature regarding estimated $\lambda_{(1)}$, $\lambda_{(2)}$, $\lambda_{(3)}$, and $\lambda_{(4)}$ for the periods 1, 2, 3, and 4, respectively.



As it can be seen in figure 1, the middle of each half hour interval is used to construct this graph which is correct when we assume the uniform distribution within the intervals. Due to the very small differences between estimated $\lambda_{(3)}$ and $\lambda_{(4)}$, and thus both corresponding curves cannot be graphically distinguished, we present only one curve for both periods which corresponds to $\lambda_{(3,4)}$. Note that in our case the first intraday interval from 9:00 a.m. to 9:30 p.m. has the longest

maturity and the last one from 04:30 p.m. to 05:00 p.m. has the shortest maturity. Hence, these functions are turned around compared with the curvature presented by Diebold and Li (2006).

We can recognize that the loadings on the curvature in period 1 is very small and quite flat whereas it has a negative slope and is monotonically decreasing. It is monotonically decreasing in periods 3 and 4 as well. However, the difference here is that these loadings are remarkably high, and the nonlinearity is to some extent obvious. So, all three loadings in periods 1, 3, and 4 support the hypotheses of monotonically decreasing interest rates during a day advocated by e.g. Baglioni and Monticini (2008) and explained by the intraday risk premium. However, the interesting result can be seen in period 2. A curvature with the maximum peak around noon is estimated. This is clear evidence of a highly nonlinear shape due to the curvature factor within the SIYC in period 2.

Furthermore, regarding our findings from the tables 5,7 and 9 we can state that, based on the R^2 all three models are capable of modelling the SYIC in the e-MID market. To the best of our knowledge, such high R^2 have not been achieved in similar studies by analyzing the intraday interest rates on an interbank credit market.

In his empirical study, Angelini (2000) states that quite low R^2 of 0.02 for the modeling of the intraday term structure can be achieved, accenting this weak evidence for an intraday downtrend. As he uses a “pre-crisis period” we can state that the standard nonlinear models for the estimation of the SIYC surpasses linear models like of Angelini (2000).

Our results further indicate that our empirical findings are better than those obtained by Baglioni and Monticini (2008) where their model achieves an R^2 of 0.09 and who also estimate there the term structure in a pre-crisis period.

Baglioni and Monticini (2010) state that they achieve an R^2 of 0.34 before the outbreak of the crisis on 9th August 2007 and of 0.21 after that. Hence, from their point of view, there may be a higher difference between the morning and afternoon interest rates but on the other hand, it seems, due to lower R^2 that the assumption of the simple downtrend in the intraday term structure becomes less reasonable after the outbreak of the financial crisis. Moreover, our results indicate that the best goodness-of-fit can be achieved immediately after the outbreak of the financial crisis starting in 2007 using nonlinear models for the estimation of the SIYC.

In terms of goodness-of-fit the closest results to ours are those obtained by Baglioni and Monticini (2013) who are able to achieve estimated R^2 of 0.367, 0.402 and 0.424, using three

different linear models. Still, their results are still not nearly as good as the ones presented in this paper.

4.1.2 Evaluation based on the MAE

We calculate the MAE, analogy to the above analysis of R^2 -s, for each day based on the following formula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{r}_i - r_i| \quad (23)$$

where \hat{r}_i is the estimate for r_i .

The results for the means of MAE of the NSM are summarized in table 11.

Table 11: MAE of the NSM

	Period 1	Period 2	Period 3	Period 4
NSM	0.0047	0.0111	0.0313	0.0114
St. Dev.	0.0100	0.0137	0.0177	0.0087

For the NSM we can state, that based on the MAE, the best model performance can be found in period 1 followed by the periods 2, 4 and 3 respectively, whereas the difference between periods 2 and 4 can be considered relatively small. Based on these statistics we can use the two-sample t-test given by formula (21) to analyze the performance of the NSM based on the MAE between the different periods.

The results of the two-sample t-test between the periods are summarized in table 12.

Table 12: Two-sample t-test of MAE for the SIYC-s estimated by the NSM

	Period 2	Period 3	Period 4
Period 1	-7.380***	-23.507***	-8.609***
Period 2		-13.423***	0.264
Period 3			14.648***

***Denotes significant different means at the 1%.

Regarding table 12 we can state, that the MAE between all periods are statistically different at the 1% level, except the difference between period 2 and 4. Hence, the difference in MAE between these two particle periods cannot be considered as significant. This implies that the

MAE for the NSM is significantly the best in the period 1, before the crisis. The worst performance can be found in period 3, within the crisis, when the market is not functioning well.

The results of the means of MAE for the SVM are summarized in table 13.

Table 13: MAE of the SVM

	Period 1	Period 2	Period 3	Period 4
MAE	0.0033	0.0079	0.0253	0.0096
St. Dev.	0.007	0.0097	0.0143	0.0079

Based on table 13 we can state that the findings follow the exact same tendency as the NSM. Based on the MAE the best model performance can be found again in period 1, followed by periods 2, 4 and 3, whereas the difference between the periods 2 and 4 seems relatively small.

The two-sample t-test between the sub-periods confirm these findings. The results of the test are summarized in table 14.

Table 14: Two-sample t-test of MAE for the SIYC-s estimated by the SVM

	Period 2	Period 3	Period 4
Period 1	-7.501***	-25.635***	-10.586***
Period 2		-15.224***	-2.091**
Period 3			13.875***

***, ** Denotes significant different means at the 1% and 5% level respectively.

Based on this test, we can state that all differences are statistically different at the 1% level, except between periods 2 and 4 where we have statistically different means at the 5% level.

The results of the means MAE for the DLM are summarized in table 15.

Table 15: MAE of the DLM

	Period 1	Period 2	Period 3	Period 4
MAE	0.0054	0.0126	0.0318	0.0117
St. Dev.	0.0112	0.0153	0.0177	0.0091

Based on table 15 we can state that the DLM achieves the best performance based on the MAE in period 1 and the worst in period 3, as in the case of the NSM and the SVM. The difference here is, that using this model, the model performance in period 4 is better than in period 2. However, this difference in the performance of the DLM is not statistically different as the two-sample t-test between the periods in table 16 indicates.

Table 16: Two-sample t-test of MAE for the SIYC-s estimated by the DLM

	Period 2	Period 3	Period 4
Period 1	-7.382***	-22.204***	-7.449***
Period 2		-12.129***	0.720
Period 3			14.707***

*** Denotes significant different means at the 1% level.

Based on table 16 we can state, that the MAE between the periods are statistically different at the 1% level, except between period 2 and 4 where we cannot confirm statistically different MAE-s with NSM and DLM, but at 5% with SVM.

The results based on the MAE are quite different to the results given by the analysis of the R^2 which also shows up regarding RMSE in the next section. We will discuss this fact in section 4.3.

4.1.3 Evaluation based on the RMSE

In this section we present results for the RMSE. The RMSE can be calculated using the formula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{r}_i - r_i)^2} \quad (24)$$

Table 17 presents the mean RMSE in the different periods for the estimations of the SIYC in the e-MID market for the NSM .

Table 17: RMSE of the NSM

	Period 1	Period 2	Period 3	Period 4
RMSE	0.0060	0.0141	0.0397	0.0148
St. Dev.	0.0131	0.0172	0.0228	0.0122

Based on table 17 we can state that best model performance of the NSM can be found in period 1, followed by period 2, period 4 and period 3, respectively.

From the point of view of the statistical inference, these findings are mostly verified by the two-sample t-test between the different periods, which are summarized in table 18.

Table 18: Two-sample t-test of RMSE for the SIYC-s estimated by the NSM

	Period 2	Period 3	Period 4
Period 1	-7.261***	-23.032***	-8.456***
Period 2		-13.404***	-0.480
Period 3			14.003***

***Denotes significant different means at the 1%.

Here we can state that the results of the RMSE are statistically different even at the 1% between all periods besides period 2 and 4 where there is no statistically significant difference.

The means of the RMSE for the SVM are presented in table 19.

Table 19: RMSE of the SVM

	Period 1	Period 2	Period 3	Period 4
MAE	0.0044	0.0103	0.0331	0.0127
St. Dev.	0.0097	0.0126	0.0194	0.0113

Based on table 19 we can state that the fit of the SVM for the SIYC-s is best also in period 1 followed by period 2 and period 4 and 3 respectively. Hence these results are qualitatively in line with the results of the NSM.

The results of the two-sample t-test between the periods for the SVM can be found in table 20.

Table 20: Two-sample t-test of RMSE for the SIYC-s estimated by the SVM

	Period 2	Period 3	Period 4
Period 1	-7.207***	-24.581***	-10.038***
Period 2		-14.995***	-2.200**
Period 3			13.138***

***, ** Denotes significant different means at the 1% and 5% level respectively.

We can see that the results are statistically different at the 1% level among all periods except between the periods 2 and 4 where we can assume the significant differences at the 5% level.

The results of the mean of the RMSE for DLM are summarized in table 21.

Table 21: RMSE of the DLM

	Period 1	Period 2	Period 3	Period 4
RMSE	0.0076	0.0162	0.0412	0.0156
St. Dev.	0.0168	0.02	0.0236	0.0134

The results of the DLM indicate that the best model performance based on the RMSE can be also found in period 1 and the worst in period 3, as in the case of the NSM and the SVM. However, unlike the other two models the second-best performance is found in period 4. However, the difference between the periods 2 and 4 is not significant, which can be seen in table 22 where the results of the two-sample t-test between the periods are presented. In other periods these differences are statistically highly significant.

Table 22: Two-sample t-test of RMSE for the SIYC-s estimated by the DLM

	Period 2	Period 3	Period 4
Period 1	-6.301***	-19.729***	-6.224***
Period 2		-11.901***	0.424
Period 3			13.634***

***, ** Denotes significant different means at the 1% and 5% level respectively.

4.2 Empirical model comparison

In the previous part of this chapter we stated that all three models are capable of modeling the SYIC. Here we will compare the three different models based on the three measures of model performance with each attempting to answer the question of which model may be the best one for the modeling of the SIYC on the interbank credit market.

4.2.1 Model comparison based on the R^2

Based on the tables 5, 7 and 9 we can see that the SVM outperforms the NSM and the DLM in each period. Regarding the comparison between the NSM and DLM we can state that, the NSM

surpasses the DLM in each period, however, the results in periods 2 and 3 may not be different to a large extent.

To test if these differences in means are also statistically verified, we perform a two-sample t-test, based on formula (20), by testing the means of R^2 between these models. The results of these two-sample tests are summarized in Table 23 for each period.

Table 23: Two-sample t-test between the models for R^2

	Period 1	Period 2	Period 3	Period 4
NSM/SVM	-8.351***	-5.384***	-4.445***	-4.975***
SVM/DLM	12.746***	8.224***	5.151***	5.989***
NSM/DLM	4.635***	3.101***	0.758	1.193

*** Denotes significant different means at the 1% level.

Based on table 23. we can state that our previously described findings regarding models' performances and their comparison are also statistically confirmed. The difference in means between the SVM and the NSM and between the SVM and the DLM is significant even at the 1% level. Regarding the comparison between the NSM and the DLM we can state that the differences in periods 1 and 2 are also significant at the 1% level. As already stated in the comparison between the NSM and DLM, the means within period 3 and period 4 do not differ to any great degree from each other. This obviously does not lead to the rejection of the assumption for the equality of those two means.

4.2.2 Model comparison based on the MAE

By considering the tables 11, 13 and 15 for the comparison of the models we can state that the SVM dominates the NSM and the DLM in each sub-period as in the case of the R^2 . Regarding the comparison between the NSM and the DLM we can state that NSM surpasses the DLM in all sub-periods, though these differences are not as high as in the case of the SVM. To verify statistically our findings based on the MAE we also perform a two-sample t-test between the models. These findings are summarized in table 24.

Table 24: Two-sample t-test between the models for MAE

	Period 1	Period 2	Period 3	Period 4
NSM/SVM	2.353**	3.146***	3.362***	2.250**
SVM/DLM	-3.338***	-4.273***	-3.701***	-2.658***
NSM/DLM	-1.053	-1.197	-0.3	-0.437

***, ** Denotes significant different means at the 1% and 5% level respectively.

Based on table 24 we can state that the differences regarding the MAE between the NSM and the SVM are statistically different even at the 1% level in the periods 2 and 3, whereas in the period 1 and 4 the differences are significant at the 5% level. Regarding the comparison of the SVM and the DLM we can see that the differences in each period are different at the 1% level. Hence, the dominance of the SVM in comparison to the other two models can be statistically verified. The comparison of the NSM and the DLM shows, that the differences in each sub-period are not statistically different even at the 10%.

4.2.3 Model comparison based on the RMSE

Regarding the model comparison based on the RMSE given the results in tables 17, 19 and 21 we can state that the SVM surpasses the NSM and the DLM in terms of a lower RMSE in each period. By comparing the results of the NSM and the DLM we can state that the findings are quite similar, especially in the periods 2,3 and 4. These findings are also confirmed by the two-sample t-test shown in table 25.

Table 25: Two-sample t-test between the models for RMSE

	Period 1	Period 2	Period 3	Period 4
NSM/SVM	2.153**	2.973***	2.835***	1.893*
SVM/DLM	-3.597***	-4.164***	-3.392***	-2.466**
NSM/DLM	-1.627	-1.323	-0.574	-0.650

***, **, * Denotes significant different means at the 1%, 5% and 10% level respectively.

By considering these test results we can state that the differences in the RMSE between the SVM and the NSM are highly significant at the 1% in the periods 2 and 3, whereas they are significant at the 5% and 10% in the periods 1 and 4, respectively. By comparing the SVM and the DLM we can state that the differences in the means of the RMSE are significantly different

at the 1% in each period except period 4, where it is significantly different at the 5% level. Regarding the comparison between the NSM and DLM we can state that the differences based on the RMSE are not statistically significant.

4.3 Discussion of empirical results

The analysis regarding the goodness-of-fit which is measured by R^2 , MAE and RMSE for each model and over different periods reveals some interesting results. At first, all three models provide highly significant goodness-of-fit in each period so that one should consider these models when modeling SIYC on interbank credit markets. By taking a deeper look into the single periods the findings are also quite interesting. Again, these periods are defined as before, directly after the outbreak, during and after the financial crisis of 2007, which means periods 1, 2, 3, and 4 respectively. Having first fitted the SIYC-s to e-MID data and considering the MAE and the RMSE over those sub-periods, the qualitatively same results for NSM and SVM occur. That means that the best performance from both models was achieved in period 1 and the worst one in period 3. Moreover, the performance of these models seems to be better in period 2 than in period 4. The DLM is in line with results from MAE and RMSE for periods 1 and 3 where these results are the best and the worst ones, respectively. However, it is vice versa regarding periods 2 and 4. We point out that the results from the MAE and RMSE for the NSM and DLM between periods 2 and 4 are statistically not different. Hence, we can state that the results from DLM are not in conflict with them from NSM and SVM. Based on the facts that these four periods represent four different states of the market, one can conclude that only the SVM is able to recognize those different market states, and thus, has a further advantage over the NSM and DLM. Thus, when there is a need to recognize different market states on the interbank credit market, rather the SVM should be applied for these purposes.¹⁷

The results look differently regarding R^2 when comparing the performance of the models over different sub-periods. Overall, the best goodness-of-fit could be achieved in period 2 and the worst one in period 4 which is in line with results achieved by Demertzidis and Jeleskovic (2016). NSM and SVM have a better performance in period 1 than in period 3 while for DLM it is vice versa. However, using the NSM no significant difference between period 1 and period 3, and using DLM between period 1 and 4 can be found. This is not the case for SVM which detects significant differences among all four periods at least at 10% level of significance. Thus, it implies that the SVM shows again the higher ability, also based on R^2 , to distinguish between

¹⁷ However, Demertzidis and Jeleskovic (2016) demonstrate that also the NSM possesses this ability when it is applied to tick-by-tick data on e-MID.

periods of a properly working interbank credit market and the odd market states. From the economic point of view, this may be an interesting and important finding.

The question arises as to why we get partially inconsistent results when we use MAE and RMSE on the one side and R^2 on the other side. The reason may rely on the variation of the dependent variable in the first period which is small within a day so that daily SIYC-s look quite flat. Whereas MAE and RMSE do not take directly into account the variation of the dependent variable, R^2 does. Given the empirical fact that the variation of interest rates in the first period is very small, compared to other periods before and during the financial crisis, the MAE and RMSE may be *per se* relatively lower in the period 1. On the other hand, the lower variation of the dependent variable has a relevant and direct impact on R^2 . This might cause the results that the best fit was achieved in period 1 according MAE and RMSE, and in period 2 according R^2 . After all, one must recognize that when different measures for the goodness-of-fit are used, different qualitative results can be achieved.

Regarding the direct comparison between the three different models, we can state that the SVM dominates the NSM and DLM in each different sub period regarding all three applied statistics. So, SVM may be the advanced model for modeling SIYC on an interbank credit market.

The comparison of results between the NSM and the DLM do not provide overall clear results. Regarding the R^2 , we can state that the NSM dominates the DLM only in periods 1 and 2, when the market is still functioning well, whereas the differences in the means of the R^2 are not statistically significant when the market is not functioning properly in periods 3 and 4. Regarding the comparison of these two models based on the MAE we can state that the differences in the sub periods are not statistically significant. This is also the case when comparing the models based on the RMSE.

Therefore, by taking into account these facts, we can conclude that the statistical justification is given to assume that SVM dominate both other models in terms of the direct comparisons based on different statistics for the goodness-of-fit. Hence, given the fact that SVM is able to model two humps, and thus higher non-linearities, which is, on the other hand, not the case with NSM and DLM we can state that alone the strong nonlinearities in SIYC are ground for such better performance of SVM. However, the NSM and the DLM are able to capture non-linearities in the SIYC as well as what is proven in section 4.1. Moreover, in terms of statistical tests, one

cannot see the NSM in favor of DLM although the means of three measures of goodness-of-fit are slightly higher for NSM.¹⁸

5. Conclusion

This paper represents the first analysis of the in-sample comparison of three standard models, namely NSM, SVM, and DLM, for the estimation of the non-linear SIYC on the e-MID market and on interbank credits markets on the intraday frequency in general. We apply estimations of the models' parameters based on the half-hourly means of interest rates. Regarding that, this procedure is in line with other comparable studies even though they use hourly intervals of interest rates on e-MID. Moreover, we split the data into four periods before, after the outbreak of financial crisis, after the collapse of Lehman Brothers and after the financial crisis to analyze the effects of the financial crisis of 2007.

We find out that all three models are suitable for the estimation of an intraday yield curve on e-MID. This is based on the fact that the goodness-of-fit of all three models for the SIYC is remarkably high in each period, and thus, these models can be used for the modelling of the SIYC on the e-MID market independently of the state of interbank credit market. For the measure of goodness-of-fit, R^2 , Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used. Furthermore, compared with the results from other studies, where linear regressions were applied, these three models seem to be highly dominant over all other linear models when comparing the goodness-of-fit measured by R^2 . To statistically justify our results and to compare them among different periods, and thus, among different states of interbank credit markets and among these three models, we use corresponding t-tests based on these three measures.

Regarding the analysis among the periods, we find that the highest R^2 can be achieved in period 2 by all three models. The second-best result for R^2 is achieved in period 1 by SVM and NSM. These are periods which are assumed to have a properly functioning market. All three models achieved the smallest R^2 in period 4 when the market liquidity after interventions of ECB was very low. Hence, it is assumed that in this period the market was not functioning properly and due to that, all three models have the smallest R^2 . However, also in this odd market state all three analyzed models still achieve remarkably high R^2 which is statistically different from zero by a very high significance level. Using MAE and RMSE, the best goodness-of-fit is achieved in the first period whereas the lowest one is in period 3 by all three models. The reason for this

¹⁸ Again, the only significant difference in favor of NSM over DLM is given in periods 1 and 2 and based only on R^2 .

variation in results based on R^2 and other two one measures is that the last two do not directly consider the variation of the dependent variable. Thus, this variation was the lowest one in period 1 so that this fact may cause this discrepancy. Moreover, NSM and SVM achieved second best results regarding MAE and RMSE in period 2 whereas DLM achieves the second-best result in period 4. However, neither NSM nor DLM are able to distinguish a statistically significant difference between period 2 and 4. On the other hand, SVM is able to recognize statistically different results among all four periods using each measure of fit. This is a strong result and we strongly recommend using the SVM when one wants to analyze different market states as in periods before, during and after financial crises.

Furthermore, we find out that, that the SVM, based on the two-sample t-test, dominates the NSM and the DLM regarding in-sample performance measures in all four single periods and regarding all three applied measures. At first sight, the NSM model seems to be the second best model, due to the fact that it dominates the DLM through the different periods and due to the different in-sample statistics. However, these differences in term of goodness-of-fit regarding MAE and RMSE between NSM and DLM are not statistically significant. Hence, one can state in this context that the results from NSM and DLM are not statistically different. Regarding R^2 , NSM outperform DLM significantly only in periods 1 and 2. Again, these are states when the market was functioning properly.

Hence, our findings state that SVM is to be preferred when an economic analysis on interbank credit market should be conducted. NSM could be preferred over DLM if one conducts the in-sample analysis in interbank credit markets on condition that the market is working properly. However, this finding for NSM and DLM is based only on the goodness-of-fit-measurement given by R^2 and this statement cannot be given based on MAE and RMSE.

References

- [1] Abbassi, P., Fecht, F., & Tischer, J. (2017). Variations in market liquidity and the intraday interest rate. *Journal of Money, Credit and Banking*, 49(4), 733-765.

- [2] Afonso, A., & Martins, M. M. (2012). Level, slope, curvature of the sovereign yield curve, and fiscal behavior. *Journal of Banking & Finance*, 36(6), 1789-1807.

- [3] Aljinović, Z., Poklepović, T., & Katalinić, K. (2012). Best Fit Model for Yield Curve estimation. *Croatian Operational Research Review*, 3(1), 28-40.

- [4] Angelini, P. (2000). Are banks risk averse? Intraday timing of operations in the interbank market. *Journal of Money, Credit and Banking*, 54-73.

- [5] Angelini, P., Nobili, A., & Picillo, C. (2011): "The interbank market after August 2007: What has changed, and why?," in: *Journal of Money, Credit and Banking*, 43(5), 923-958.

- [6] Baglioni, A., & Monticini, A. (2008). The Intraday Price of Money: Evidence from the e-MID Interbank Market. *Journal of Money, Credit and Banking*, 40(7), 1533-1540.

- [7] Baglioni, A., & Monticini, A. (2010). The intraday interest rate under a liquidity crisis: the case of August 2007. *Economics Letters*, 107(2), 198-200.

- [8] Baglioni, A., & Monticini, A. (2013). Why does the interest rate decline over the day? Evidence from the liquidity crisis. *Journal of Financial Services Research*, 44(2), 175-186.

- [9] Bernadell, C., Coche, J., & Nyholm, K. (2005). *Yield curve prediction for the strategic investor*. ECB Working Paper Series No. 472 (2005)
- [10] BIS (2005): Zero- Coupon Yield Curves: Technical Documentation, BIS Papers, Nr. 25
- [11] Björk, T., & Christensen, B. J. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4), 323-348.
- [12] Bliss, R. R. (1996). *Testing term structure estimation methods* (No. 96-12a). Working paper, Federal Reserve Bank of Atlanta.
- [13] Brunetti, Celso, Jeffrey H. Harris, Shawn Mankad, & George Michailidis (2015). "Interconnectedness in the Interbank Market," Finance and Economics Discussion Series 2015-090. Washington: Board of Governors of the Federal Reserve System, <http://dx.doi.org/10.17016/FEDS.2015.09>
- [14] Chen, Y., & Niu, L. (2014). Adaptive dynamic Nelson–Siegel term structure model with applications. *Journal of Econometrics*, 180(1), 98-115.
- [15] Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2009). An arbitrage-free generalized Nelson–Siegel term structure model. *The Econometrics Journal*, 12(3), C33-C64.
- [16] Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics*, 164(1), 4-20.

- [17] Clare, A., & Lekkos, I. (2000). An analysis of the relationship between international bond markets (No. 123). Bank of England.
- [18] Csajbok, A. (1998). *Zero-coupon yield curve estimation from a central bank perspective* (No. 1998/2). Magyar Nemzeti Bank (Central Bank of Hungary).
- [19] De Pooter, M. (2007). Examining the Nelson-Siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance. *Available at SSRN 992748*.
- [20] De Rezende, R. B. (2011). Giving flexibility to the Nelson-Siegel class of term structure models. *Available at SSRN 1290784*.
- [21] De Rezende, R. B., & Ferreira, M. S. (2008). Modeling and Forecasting the Brazilian Term Structure of Interest Rates by an Extended Nelson-Siegel Class of Models: A Quantile Autoregression Approach. In *Latin American Meeting of Econometric Society*.
- [22] Demertzidis, A. & Jeleskovic, V. (2016): " Empirical Estimation of Intraday Yield Curves on the Italian Interbank Credit Market e-MID,, MAGKS Joint Discussion Paper Series in Economics No. 49-2016
- [23] Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of econometrics*, 130(2), 337-364.
- [24] Diebold, F. X., & Rudebusch, G. D. (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of econometrics*, 131(1-2), 309-338.

- [25] Engler, M. & Jeleskovic, V. (2016): " Intraday volatility, trading volume and trading intensity in the interbank market e-MID,, MAGKS Joint Discussion Paper Series in Economics No.48-2016
- [26] Furfine, C. H. (2001). Banks as Monitors of Other Banks: Evidence from the Overnight Federal Funds Market*. *The Journal of Business*, 74(1), 33-57.
- [27] Furfine, C. (2002). The interbank market during a crisis. *European Economic Review*, 46(4), 809-820.
- [28] Gabbi, G., Germano, G., Hatzopoulos, V., Iori, G., and Politi, Mauro (2012): Market microstructure, bank's behavior and interbank spreads, Working Paper, City University of London.
- [29] Ganchev, A. (2009): Modeling the yield curve of spot interest rates under the conditions in Bulgaria, Working Paper, Academy of Economics, Department of Finance and Credit, Narodnostopanski Arhiv, International edition, 119-137.
- [30] Gilli, M., Grosse, S. & Schumann, E., Calibrating the Nelson-Siegel-Svensson Model. Available at SSRN: <https://ssrn.com/abstract=1676747>
- [31] GrØnborg, N. S., & Lunde, A. (2016). Analyzing Oil Futures with a Dynamic Nelson-Siegel Model. *Journal of Futures Markets*, 36(2), 153-173.
- [32] Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The US Treasury yield curve: 1961 to the present. *Journal of monetary Economics*, 54(8), 2291-2304.

- [33] Hautsch, N. & V. Jeleskovic (2008). Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models. In W. Härdle, N. Hautsch, and L. Overbeck (Eds.), *Applied Quantitative Finance* (2 ed.). pp. 379 – 397. Berlin, Germany: Springer.
- [34] Ioannides, M. (2003). A comparison of yield curve estimation techniques using UK data. *Journal of Banking & Finance*, 27(1), 1-26.
- [35] Iori, G., Kapar, B., & Olmo, J. (2012). *The cross-section of interbank rates: a nonparametric empirical Investigation* (Report No. 12/03). London, UK: Department of Economics, City University London.
- [36] Jurgilas, M., & Žikeš, F. (2014). Implicit intraday interest rate in the UK unsecured overnight money market. *Journal of Financial Intermediation*, 23(2), 232-254.
- [37] Kempf, A., & Mayston, D. (2005). Commonalities in the liquidity of a limit order book. In *Department of Finance and Centre for Financial Research, University of Cologne Germany Working paper*.
- [38] Laurini, M. P., & Hotta, L. K. (2010). Bayesian extensions to Diebold-li term structure model. *International Review of Financial Analysis*, 19(5), 342-350.
- [39] Merrouche, O., & Schanz, J. (2010). Banks' intraday liquidity management during operational outages: Theory and evidence from the UK payment system. *Journal of Banking & Finance*, 34(2), 314-323.
- [40] Mönch, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics*, 146(1), 26-43.

- [41] Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of business*, 473-489.
- [42] Schich, S. T. (1997). *Estimating the German term structure* (No. 1997, 04e). Deutsche Bundesbank, Research Centre.
- [43] Svensson, L (1994), 'Estimating and interpreting forward interest rates: Sweden 1992-94', *IMF working paper series*, No 114.
- [44] Tam, C. S. & Yu, I. W.: (2008): "Modelling sovereign bond yield curves of the US, Japan and Germany," in: *International Journal of Finance & Economics*, 13(1), 82-91