ORIGINAL PAPER

DOI: 10.26794/2587-5671-2023-27-6-44-53
UDC 336.763.331(045)
JEL G12, E43

Treatment of Missing Market Data: Case of Bond Yield Curve Estimation

M.S. Makushkin, V.A. Lapshin
National Research University Higher School of Economics, Moscow, Russia

ABSTRACT

Missing observations in market data is a frequent problem in financial studies. The problem of missing data is often overlooked in practice. Missing data is mostly treated using ad hoc methods or just ignored. Our **goal** is to develop practical recommendations for treatment of missing observations in financial data. We illustrate the issue with an example of yield curve estimation on Russian bond market. We compare three **methods** of missing data imputation — last observation carried forward, Kalman filtering and EM–algorithm — with a simple strategy of ignoring missing observations. We **conclude** that the impact of data imputation on the quality of yield curve estimation depends on model sensitivity to the market data. For non-sensitive models, such as Nelson-Siegel yield curve model, final effect is insignificant. For more sensitive models, such as bootstrapping, missing data imputation allows to increase the quality of yield curve estimation. However, the **result** does not depend on the chosen data imputation method. Both simple last observation carried forward method and more advanced EM–algorithm lead to similar final results. Therefore, when estimating yield curves on the illiquid markets with missing market data, we **recommend** to use either simple non-sensitive to the data parametric models of yield curve or to impute missing data before using more advanced and sensitive yield curve models.

***Keywords*:** yield curve; term structure of interest rates; bond market; Nelson-Siegel method; liquidity level; missing data; emerging markets

For citation: **Makushkin M.S., Lapshin V.A. Treatment of missing market data: Case of bond yield curve estimation. *Finance: Theory and Practice*. 2023;27(6):44-53. (In Russ.). DOI: 10.26794/2587-5671-2023-27-6-44-53**

INTRODUCTION

Missing data is a common problem in many empirical studies. It is also found in finance, especially when it comes to emerging, low-liquidity markets. The presence of gaps in the data complicates the assessment of financial models and may distort conclusions. Market data gaps can arise for a variety of reasons, for example, due to low liquidity of the instrument, data censoring, or the filtering of outliers.

In liquid developed markets, the problem of missing data is generally ignored because missing observations are not frequent. Researchers focus more on model specification rather than on data quality issues.

In less liquid markets, missing observation are more frequent. Simply removing gaps in data can lead to the loss of important information. Therefore, researchers often pre-process the data. The choice of the method for missing data treatment is usually made ad-hoc, and the processing of the data itself is an auxiliary step on the way to answering the main research questions.

In the paper, we explore the treatment of missing observations in market data in more detail in relation to the task of estimating the bond yield curve. Our goal is to develop practical recommendations on how to deal with gaps in market data when estimating the term structure of interest rates.

To illustrate the problem and its possible solutions in practice, we use data on the trading of Russian Federal Loan Bonds (further — FLBs). We investigate how filling gaps in the trading data affects the estimation quality of the yield curve. This is an important topic subject since yield curve estimation in emerging markets is sometimes negatively impacted by gaps in market data [1].

The novelty of the paper is the application of statistical tools from the field of incomplete data analysis to the tasks of financial engineering. We show that the issues of data quality and completeness in emerging markets require no less attention than the questions of financial model choice.

LITERATURE REVIEW

In statistical science, there is a separate area dedicated to the formal analysis of missing data [2–4]. Data gaps are dangerous for two reasons. First, they can lead to shifted estimates of model parameters. Second, they increase the standard error of model coefficients and reduce the power of statistical tests [5].

One can safely remove the gaps from the analysis only if they are completely random, i.e. they are not dependent on their own non-observed values or on the values of other observations. If missing observations are simply random, i.e. they do not depend on their own missed value, but on the values of other parameters, then missed data can be restored by means of conditional imputation. If gaps are not random, then to fill them, you need to know the process that generates the gaps [3].

Gaps in market data can be both random and non-random. Non-random gaps can arise, for example, when trade in a stock or bond stops due to an unexpected decline in the price of the stock or bond (for example, suspension of trading in the Russian market in February–March 2022). Such cases must be examined independently, with special emphasis on the reasons behind missing observations.

With random missing data, on the contrary, you can work effectively. There are two ways to do this. The first option is to adapt the model to address data gaps. This is a more correct way, but it complicates the model, depends on its properties and specifications, and is not universal. Please refer to [6–8] for the description of missing data treatment in yield curve estimation tasks.

The second option to work with gaps is to pre-fill gaps in the data. The advantages of the approach are its ability to retain a simple original model, as well as to use the processed data for other purposes. As part of our study, we are considering this more general option.

In practice, the most common solution to the problem is to remove gaps [9]. This is a simple solution, but it leads to the loss of some important information. In some situations, it may be useful to impute the missing data. A general overview of possible data imputation methods is presented in the papers [10, 11]. With regards to financial problems, both advanced methods of filling gaps in the data are used (EM-algorithm [12, 13], Bayesian model with Markov chain Monte Carlo [14]), as well as simpler methods (last observation carried forward [15] and the last weighted value [16]). In the cited examples methods for missing market data treatment are chosen without much of justification. There is usually no description of how the strategy used to fill gaps influences the results achieved. Data imputation is mentioned only as an intermediate step towards answering key research questions. In this regard, it is important to address the problem of missing market data in greater detail and to compare the methods of processing them for financial objectives.

METHODOLOGY

Yield curve estimation models

We illustrate the importance of the problem of processing gaps in market data with an example of constructing a yield curve in a bond market. The yield curve shows the relationship between interest rates and maturity. It has multiple practical uses ranging from macroeconomic forecasting to risk management and pricing of financial instruments. In emerging markets, estimation of the yield curve is often complicated by gaps in market data [1]. Therefore, the task of implying the yield curve from the data with missing observations seems important and relevant.

There are many models for yield curve estimation. All of them are based on the general principle that information about the unobserved yield curve is derived from observed market information, for example, from the prices of coupon bonds. An overview of yield curve models can be found in the paper [17].

When choosing the current interest rate term structure model, the quantity and quality of data available for calibration must be taken into account [18]. The model should be as complex and accurate as the available data allows. Where the number of available data is limited, conservative assumptions and simpler models should be used. Therefore, for market data with gaps, it is advisable to consider two models of the yield curve with varying degrees of data quality sensitivity. We select a very simple but stable Nelson-Siegel yield curve model [19] and a more flexible but less stable interest-rate bootstrapping method [20].

The Nelson-Siegel model in the factor specification [21] describes the yield $y(\tau)$ for maturity τ with the following equation:

$$y(\tau)=\beta_1+\beta_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)+\beta_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}-e^{-\lambda\tau}\right)+\varepsilon(\tau), \quad (1)$$

where $\theta=\{\beta_1,\beta_2,\beta_3,\lambda\}$ — model parameters vector, and; $\varepsilon(\tau)$ — error.

In fact, the rate of $y(\tau)$ in the model is the weighted sum of three factors, where factors are weighted by β_1,β_2,β_3. These weights can be interpreted as the level, slope and curvature of the yield curve. The fourth parameter λ describes the relative position of the curvature on the chart.

The Nelson-Siegel model captures quite well the empirically observed shapes of the yield curve. Therefore, it is used some way by many financial institutions.[1] In particular, the model of the G-curve of the Moscow Stock Exchange[2] is based on the Nelson-Siegel variation of the yield curve. The model is also popular in emerging markets and low-liquidity markets, as it has only four parameters that can be conveniently calibrated to a small set of market data available [22, 23].

Estimation of model parameters can be done by minimizing the average square error of revaluation of bond yields, or directly from the bond prices. A description of technical details behind the yield curve estimation can be found, for example, in paper [24]. We choose to calibrate the yield curve model directly to the coupon bond prices, because zero-coupon yields on the FLB market are not observed directly.

If the vector P, of coupon bond prices is given, the task of estimating the parameters of the yield curve is described by the following equations:

$$\hat{\theta}=\arg\min_{\theta}\sum_{i}^{N}(\hat{P}_i(\theta)-P_i)^2, \quad (2)$$

$$\hat{P}_i=\sum_{j=1}^{J_i}CF_{ij}e^{-y(\tau_{ij})\tau_{ij}}, \quad (3)$$

where j — serial number of cash flow; J — total number of cash flows, and; $y(\tau_{ij})$ — zero-coupon rate, corresponding to the time τ_{ij} till payment of CF_{ij} on the bond i. In fact, the model is estimated as a classical regression, so all data instabilities are smoothed at the calibration stage.

Bootstrap is a more data-sensitive method. It is based on the idea of sequential calculation of the zero-coupon rate in order of increasing the maturity of bonds. Given the values of bonds at various maturities, you may iteratively find zero-coupon interest rates using the bootstrap. The produced curve is an exact fit to the original market data.

Formally, the bootstrap logic can be written as follows. Suppose we have a set of bonds $P_i, i\in 1\ldots N$ at with maturities $\tau\in 1\ldots J$. Then the price of each bond can be represented as the sum of its discounted future streams:

$$\begin{cases}\hat{P}_1=CF_{11}e^{y(\tau_{11})\tau_{11}}\\ \hat{P}_2=CF_{21}e^{y(\tau_{21})\tau_{21}}+CF_{22}e^{y(\tau_{22})\tau_{22}}\\ \ldots\\ \hat{P}_N=CF_{N1}e^{y(\tau_{N1})\tau_{N1}}+CF_{N2}e^{y(\tau_{N2})\tau_{N2}}+\ldots+CF_{NJ}e^{y(\tau_{NJ})\tau_{NJ}}\end{cases} \quad (4)$$

The system of equations (4) can be solved by the iterative method by replacing the found zero-coupon yields from the first equations in the subsequent equations.

In reality, the set of available bond prices P is limited. An assumption of the form of the yield curve is required at the intervals between the

[1] Bank of International Settlements. Zero-coupon yield curves: Technical documentation. BIS Papers. 2005;(25).

[2] MOEX (2021). Zero-coupon Yield Curve for Sovereign Bonds. https://www.moex.com/a3642 (accessed on 09.11.2022).

maturities of the available securities. We use the basic assumption of a piecewise constant form of the yield curve, but we can also assume more complex dependencies [25].

A bootstrapped yield curve $y(\tau)$ depends on each point of market data and ideally reproduces the initial data. Any noise in the source data can significantly distort the shape of the curve, and gaps in the data will make it less smooth [26]. Bootstrap is commonly used in developed liquid markets where enough bonds are traded and structural market inefficiencies are minimized.

Methods for missing data imputation

We examine two strategies of dealing with missing data: simply remove the gaps in data and fill gaps. The first option serves as a benchmark strategy, as it is most commonly found in financial research that deals with missing data. The second option is more advanced. It is less common in the literature, but sometimes it allows to improve the quality of model evaluation. We explore how the gap filling strategy improves the quality of yield curve estimation compared to the gap removal strategy.

We have chosen three methods for filling gaps: 1) a simple heuristic method of filling in the last value (last observation carried forward); 2) a Kalman filter that takes into account the previous dynamics of observations; 3) an EM-algorithm that takes into account the aggregate dynamics. This selection of methods makes it possible to compare how much the complexity of missing data imputation method impacts the quality of the evaluation of the curve.

The easiest way to fill a gap is to fill it with the last value. This method is simple to implement, but it uses the strong assumption that in the absence of data, a previous observation is the best estimate of a missing value.

The Kalman filter is a more advanced option for filling gaps. When filling the gaps, this method takes into account the historical dynamics of a data point. Based on noisy observations, the Kalman filter evaluates the unobserved state of the system. The system state can then be used to estimate the possible values for the missed observations. Technical aspects of the method are presented in the paper [27].

In contrast to Kalman filter, EM-algorithm considers the dynamics of not just one data point, but of the entire dataset. It takes into account the dynamics of other observations and the covariation structure of the data. This is an iterative algorithm consisting of two steps. In the E-step, the expected value (expectation) of the vector of non-observed variables is calculated on the basis of the first approximation of the model coefficients estimated using available information. The M-step solves the problem of maximization and is the next approximation of the vector of the model parameters. It is then used to estimate non-observed values, and the process repeats. [28].

Comparison of estimation quality

We compare the quality of the yield curve estimation using cross-validation. The idea of the method is that all observations are used for tests. This is useful when available data sets are small. An observation is excluded from the training dataset, the model is evaluated without this observation, and then the error is calculated for the excluded test observation. The procedure is performed iteratively for each observation in the sample, and the result is then averaged [29].

When calculating cross-validation errors, we use only real observations. The imputed data is thus only needed for a more accurate estimation of the curve. It does not need to be taken into account in the calculation of error, as the ultimate goal is to improve the quality of reproduction of real observations rather than recovered observations.

DATA

We use closing prices of standard FLB with fixed coupon in the period from May 2012 to December 2015 (approximately 1000 observations). We do not include amortizing, floating and inflation-linked FLBs in the sample, as their pricing principles differ from standard FLB coupons [30]. Prices were obtained from the Finam analytical platform.

More up-to-date data on FLB trades up to 2022 was also collected and studied. However, the percentage of gaps in the data sample is small, so they would not be representative for the study. In this regard, we are restricted to looking at older data.

An alternative could be to generate data gaps, but this approach involves an artificial, exogenous missing data generating process that may distort the conclusions of the study. That is why we compromise data speed in return for an actual and non-distorted gap structure. This method doesn't reduce the practical value of the produced results. They can still be applied to other bond markets where data gaps are still a problem — to the corporate debt market, as well as to less liquid sovereign — obligation markets, where data gap is still a common phenomenon.

The average share of missing data in the collected sample is 10%. There were more missing observations in the beginning of the sample (around 30% in 2012). The share of gaps has then fallen down to 20%. By the end of 2015 there are almost no gaps in the sample. A large number of gaps are concentrated around weekends and public holidays (New Year's Eve, May Holidays). Gaps are distributed relatively randomly by individual securities. Long series of gaps are rare. Usually, their length does not exceed two to five days. Some securities are more "prone" to gaps, but generally they don't stand out much from the sample.

We rearrange the data as follows. From the closing prices we calculate coupon yields. Next, we fill the gaps in yields using data imputation methods described earlier. Then we go back from processed returns again to bond prices. The yield curve is then calibrated to new price data sample. For observations with no gaps, such a transition has no effect because yield and price are inseparable. As a result, all valid observations are preserved, while missed observations are recovered.

A transition from price to yield and backwards is necessary to take into account the pull-to-par effect (the convergence of the bond value to the nominal as it approaches repayment data). In addition, some methods of filling gaps, such as EM-algorithm, require normally distributed data. Yields distribution is closer to normal than distribution of bond prices.

To apply the EM-algorithm, we additionally calculate the average arithmetic yield on all securities for each day. This is necessary because the size of the sample is not fixed. Some securities are expired, others, on the contrary, are issued. As a result, estimating the covariance between the yields of various securities is difficult. This issue can be avoided by calculating the covariance between yield of a particular bond and an average yield, rather than the covariance matrix of yields. Of course, this simplification leads to the loss of some information. In fact, when filling gaps with the EM-algorithm, we only take into account the relationship of observation with the overall "level" of the yield curve. However, this simplification is acceptable, as parallel shifts explain most of the yield curve dynamics [31].

RESULTS

On the basis of the methodology described, gaps in data on bond trading were filled. A graphical comparison of different ways of filling gaps in yields and prices is shown on below on the example of FBL 25065 (*Fig. 1, 2*).

When the gaps in returns are filled, the results may vary greatly. Filling with the last value (shaped dots) looks most inaccurate. It creates long, constant sections in the data. Filling with the Kalman filter (cross) is slightly more effective. Because the model reads from previous yield dynamics, the constant areas are replaced by sloping ones. When using the EM-algorithm (plus), the results are quite close to expectations. In fact, the average yield dynamics of all securities is applied to the yield of a bond with missing prices.

As we transition from yields to prices, the difference between the methods of filling gaps becomes less noticeable. This is due to the fact that as the maturity decreases, the bond's sensitivity to the change in yield also decreases, and its price tends to par.

For data with filled gaps, we then constructed the yield curves using the Nelson-Siegel and

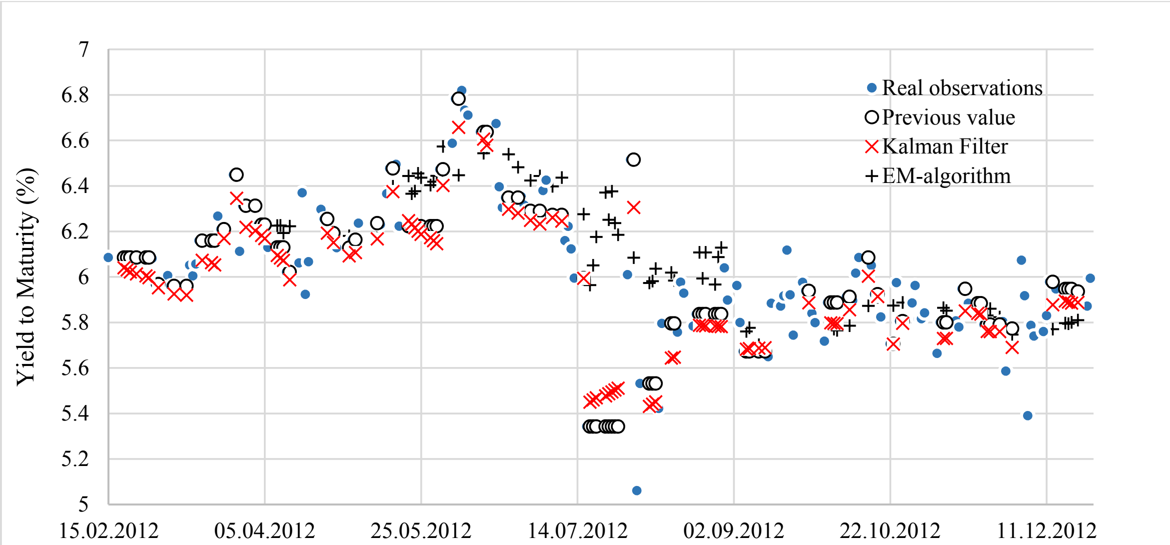

Fig. 1. **An Example of Processing Yields Gaps for FLB 25065**

Source: Author's calculations.

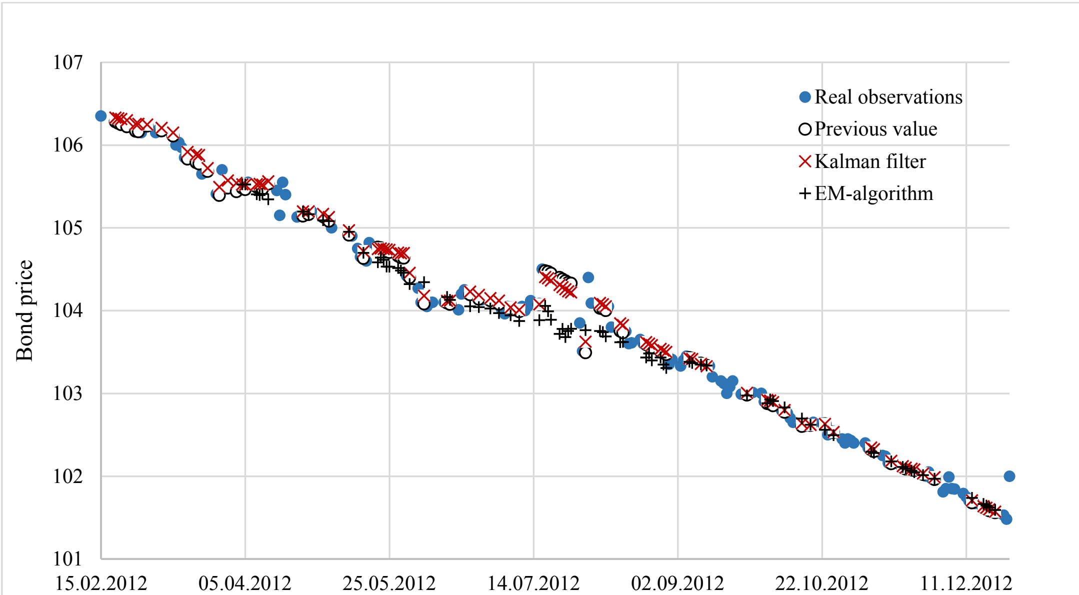

Fig. 2. **An Example of Processing Prices Gaps for FLB 25065**

Source: Author's calculations.

bootstrap models and calculated the average absolute fitting error with cross-validation. By comparing the error values, we counted the percentage of days when missing data imputation improved the quality of yield curve estimation compared to the simple deletion of the gaps. Only the days with recorded gaps were considered.

The gap filling strategy improves the fitting quality of the bootstrapped yield curve compared to the simple gap removal strategy (*Table*). The share of the days when data imputation has improved quality is approximately 65%. This is slightly higher for the EM-algorithm imputation, but generally speaking, the differences between the imputation methods are insignificant. The quality improvement resulting from the filling of gaps is statistically significant at a 95% confidence level.

For the Nelson-Siegel parametric model, missing data imputation does not significantly

Table

Percentage of the Days when Processing Data Gaps Provides a Better Yield Curve Fit than Simple Removal Processing of Gaps

	Bootstrapping		Nelson-Siegel Mode	
	Removal of gaps	Missing data imputation	Removal of gaps	Missing data imputation
Last value	35%	65%	52%	48%
Kalman filter	34%	66%	59%	41%
EM-Algorithm	31%	69%	49%	51%

Source: Author's calculations.

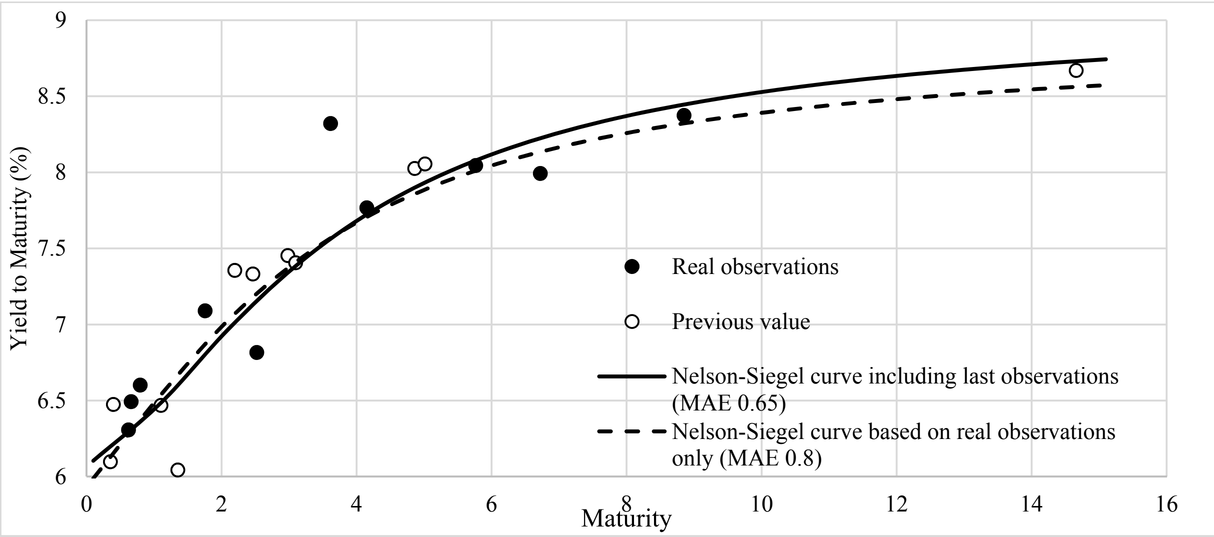

Fig. 3. **Nelson-Siegel Yield Curve for Russian FLB Constructed Using Only Real Data and Using Both Real and Gaps Data as of June 9, 2022**

Source: Author's calculations.

improve the quality of the curve estimation. The share of days when data imputation helped improve the quality of the estimation and the days when the curve is better to evaluate only on the available data are roughly equal.

The difference in the impact of missing data imputation on the quality of the yield curve estimation using the Nelson-Siegel model and bootstrapping can be explained by the different sensitivity of the models to market data. The Nelson-Siegel model has only 4 parameters. They can be estimated quite well from available observations. Adding two or three more recovered points to the 16 real points will not have a significant impact on the outcome. Bootstrap, on the contrary, depends on each point of market data. Adding even one observation makes the yield curve more smooth.

An example of the improvement in the quality of the estimation of the yield curve is shown in *Fig. 3–4* based on the trade data for 9 June 2012. It's a pre-holiday day, so trading activity was lower and there were more gaps in the data. The yield curve was first calibrated only to real observations and then to a combination

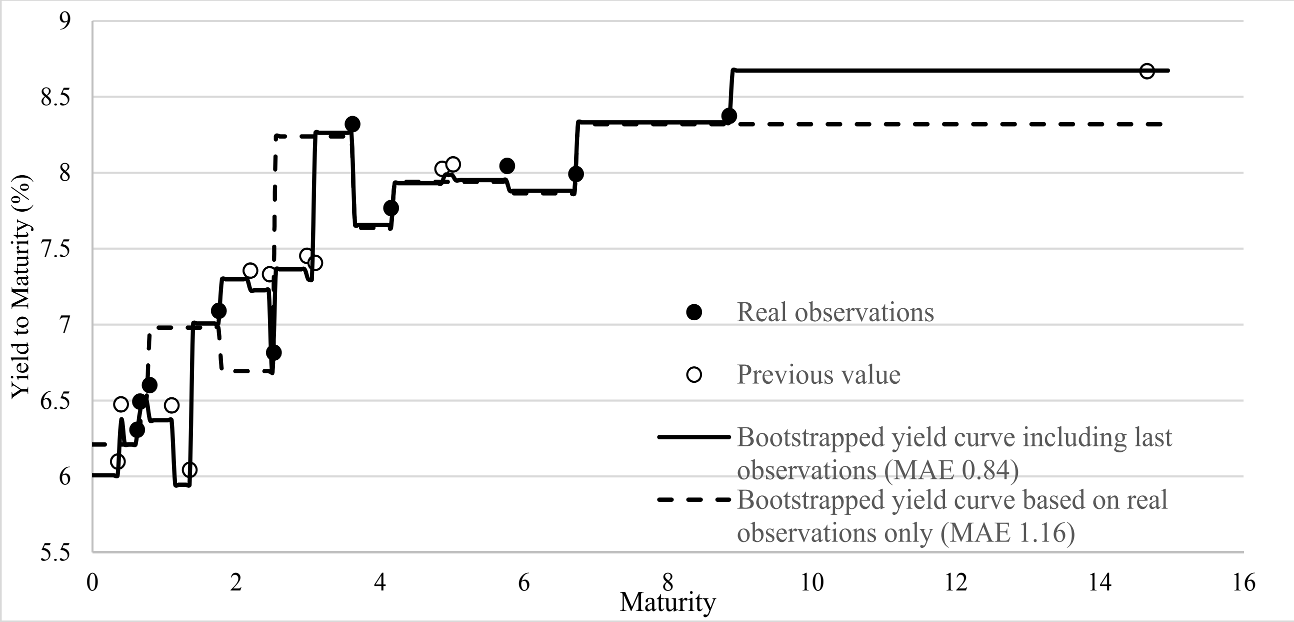

Fig. 4. **Bootstrapped Zero-Coupon Yield Curve for Russian FLB Constructed Using Only Real Data and Using Both Real and Gaps Data as of June 9, 2022**

Source: Author's calculations.

of real data and data with filled gaps. Both for the Nelson-Siegel model (*Fig. 3*), and for the bootstrap (*Fig. 4*) missing data imputation allowed to describe more accurately the term structure of interest rates, especially at the far end. The average absolute error on cross-validation was significantly reduced (the error values are shown in figures).

CONCLUSION

A comparison of different methods for missing market data imputation was carried out. The results are illustrated by an example of yield curve estimation in the Russian bond market. We have shown that ignoring gaps can lead to distorted estimates of the yield curve model. On the contrary, missing data imputation could improve the quality of estimating the yield curve compared to removing missed observations from the sample.

The effect of filling gaps in the data on the quality of estimation of the yield curve depends on the selected curve model. For the Nelson-Siegel parametric model, the positive effect of filling gaps is minimal. For bootstrapping, a statistically significant improvement in evaluation quality is recorded when filling gaps in the data. The observed differences are related to the degree of sensitivity of yield curve models to market data. The Nelson-Siegel parametric model can be efficiently calibrated even when only a small number of data points is available. For bootstrap, however, every additional observation is important.

In practice, when selecting an approach for missing data pre-processing, we propose evaluating the financial model's sensitivity to market data. It will be useful to pre-fill gaps in data if data-sensitive models are used. The specific method of dealing with gaps in relation to the task of yield curve estimation is less significant. Both simple last observation carried forward method, and more sophisticated methods of filling gaps based on the Kalman filter or EM-algorithm give a similar result. If it is not possible to fully process data gaps, then it is necessary to use simpler and less data-sensitive models.

The paper findings can be useful for yield curve estimation in low-liquidity markets and for other financial studies that deal with incomplete market data.

REFERENCES

1. Makushkin M., Lapshin V. Yield curve estimation in illiquid bond markets. *Ekonomicheskii zhurnal Vysshei shkoly ekonomiki = The HSE Economic Journal*. 2021;25(2):177–195. (In Russ.). DOI: 10.17323/1813–8691–2021–25–2–177–195
2. Haitovsky Y. Missing data in regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1968;30(1):67–82. DOI: 10.1111/j.2517–6161.1968.tb01507.x
3. Rubin D.B. Inference and missing data. *Biometrika*. 1976;63(3):581–592. DOI: 10.1093/biomet/63.3.581
4. Little R.J.A. Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics.* 1988;6(3):287–296. DOI: 10.1080/07350015.1988.10509663
5. Newman D.A. Missing data: Five practical guidelines. *Organizational Research Methods*. 2014;17(4):372–411. DOI: 10.1177/1094428114548590
6. Cortazar G., Schwartz E.S., Naranjo L.F. Term-structure estimation in markets with infrequent trading. *International Journal of Finance & Economics*. 2007;12(4):353–369. DOI: 10.1002/ijfe.317
7. Poghosyan K., Poghosyan A. Yield curve estimation and forecasting in Armenia. *Armenian Journal of Economics*. 2019;(4):1–19. URL: https://www.aea.am/files/papers/a2019_1.pdf
8. Nagy K. Term structure estimation with missing data: Application for emerging markets. *The Quarterly Review of Economics and Finance*. 2020;75:347–360. DOI: 10.1016/j.qref.2019.04.002
9. Kofman P., Sharpe I.G. Using multiple imputation in the analysis of incomplete observations in finance. *Journal of Financial Econometrics.* 2003;1(2):216–249. DOI: 10.1093/jjfinec/nbg013
10. Briggs A., Clark T., Wolstenholme J., Clarke P. Missing… presumed at random: Cost-analysis of incomplete data. *Health Economics*. 2003;12(5):377–392. DOI: 10.1002/hec.766
11. Graham J.W. Missing data analysis: Making it work in the real world. *Annual Review of Psychology*. 2009;60:549–576. DOI: 10.1146/annurev.psych.58.110405.085530
12. Warga A. Bond returns, liquidity, and missing data. *Journal of Financial and Quantitative Analysis*. 1992;27(4):605–617. DOI: 10.2307/2331143
13. Zhu H. An empirical comparison of credit spreads between the bond market and the credit default swap market. *Journal of Financial Services Research*. 2006;29(3):211–235. DOI: 10.1007/s10693–006–7626-x
14. Kos'yanenko A.V. Missing market data imputation using Bayesian inference. Financial engineering, risk management and actuarial science. Preprint WP16/2007/02. Moscow: SU HSE; 2007. 32 p. (In Russ.).
15. Teplova T., Budanova D. Pricing effectiveness at the Russian corporate bonds market. *Vestnik Moskovskogo universiteta. Seriya 6: Ekonomika = Moscow University Economics Bulletin*. 2017;(4):3–28. (In Russ.). DOI: 10.38050/01300105201741
16. Khakimzhanov S., Mustafin Y., Kubenbayev O., Atabek D. Constructing a yield curve in a market with low liquidity. *Russian Journal of Money and Finance*. 2019;78(4):71–98. DOI: 10.31477/rjmf.201904.71 (In Russ.: *Den'gi i kredit*. 2019;78(4):71–98. DOI: 10.31477/rjmf.201904.71).
17. Lapshin V., Tereshchenko M. The choice of the model of the term structure of interest rates on the basis of its properties. *Korporativnye finansy = Journal of Corporate Finance Research*. 2018;12(2):171–187. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.12.2.2018.171–187
18. Kurbangaleev M., Lapshin V. An analysis of the risk-free term structure estimation requirements posed by various financial problems. *Upravlenie finansovymi riskami = Financial Risk Management Journal*. 2015;(1):50–60. (In Russ.).
19. Nelson C.R., Siegel A.F. Parsimonious modeling of yield curves. *The Journal of Business*. 1987;60(4):473–489. DOI: 10.1086/296409
20. Fama E.F., Bliss R.R. The information in long-maturity forward rates. *The American Economic Review*. 1987;77(4):680–692.
21. Diebold F.X., Li C. Forecasting the term structure of government bond yields. *Journal of Econometrics*. 2006;130(2):337–364. DOI: 10.1016/j.jeconom.2005.03.005

22. Zoricic D., Orsag S. Parametric yield curve modeling in an illiquid and undeveloped financial market. *UTMS Journal of Economics*. 2013;4(3)234–252. URL: https://www.utmsjoe.mk/files/Vol.%204%20No.%203/1–1B-Zoricic-Orsag_-_Croatia.pdf
23. Chou J.-H., Su Y.-S., Tang H.-W., Chen C.-Y. Fitting the term structure of interest rates in illiquid market: Taiwan experience. *Investment Management and Financial Innovations*. 2009;6(1):101–116. URL: https://www.researchgate.net/publication/288118114_Fitting_the_term_structure_of_interest_rates_in_illiquid_market_Taiwan_experience
24. Wahlstrøm R.R., Paraschiv F., Schürle M.A Comparative analysis of parsimonious yield curve models with focus on the Nelson-Siegel, Svensson and Bliss versions. *Computational Economics*. 2022;59(3):967–1004. DOI: 10.1007/s10614–021–10113-w
25. Hagan P.S., West G. Interpolation methods for curve construction. *Applied Mathematical Finance*. 2006;13(2):89–129. DOI: 10.1080/13504860500396032
26. Avdeeva O., Tsyplakov A. A method for adaptive estimation of the term structure of interest rates. *Ekonomicheskii zhurnal Vysshei shkoly ekonomiki = The HSE Economic Journal*. 2015;19(4):609–639. (In Russ.).
27. Kalman R. A new approach to linear filtering and prediction problems. *Transactions of the AMSE. Journal of Basic Engineering*. 1960;82(1):35–45. DOI: 10.1115/1.3662552
28. Dempster A.P., Larid N.M., Rubin D.B. Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1977;39(1):1–22. DOI: 10.1111/j.2517–6161.1977.tb01600.x
29. Lapshin V., Sohatskaya S. Choosing the weighting coefficients for estimating the term structure from sovereign bonds. *International Review of Economics & Finance*. 2020;70:635–648. DOI: 10.1016/j.iref.2020.08.011
30. Kurbangaleev M., Lapshin V., Shepeleva I. Consistency of Russian sovereign bond prices. *Upravlencheskii uchet i finansy = Management Accounting and Finance Journal*. 2016;(1):40–51. (In Russ.).
31. Falkenstein E., Hanweck J. Minimizing basis risk from non-parallel shifts in the yield curve. Part II: Principal components. *The Journal of Fixed Income*. 1997;7(1):85–90. DOI: 10.3905/jfi.1997.408199

ABOUT THE AUTHORS

Mikhail S. Makushkin — postgraduate student, Department of Financial Market Infrastructure, National Research University Higher School of Economics, Moscow, Russia
https://orcid.org/0000-0002-3491-1561
Corresponding author:
mikhailmakushkin@gmail.com

Victor A. Lapshin — Cand. Sci. (Phys.-Math.), Assoc. Prof., Financial Engineering and Risk Management Laboratory, School of Finance, National Research University Higher School of Economics, Moscow, Russia
https://orcid.org/0000-0002-9396-4161
vlapshin@hse.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 22.08.2022; revised on 26.10.2022 and accepted for publication on 26.11.2022. The authors read and approved the final version of the manuscript.