Missing Data Bias on a Selective Hedging Strategy

▪ *Kiss Gábor Dávid, Sávai Marianna, Udvari Beáta*

Abstract

Foreign exchange rates affect corporate profitability both on the macro and cash-flow level. The current study analyses the bias of missing data on a selective hedging strategy, where currency options are applied in case of Value at Risk (1%) signs. However, there can be special occasions when one or some data is missing due to lack of a trading activity. This paper focuses on the impact of different missing data handling methods on GARCH and Value at Risk model parameters, because of selective hedging and option pricing based on them. The main added value of the current paper is the comparison of the impact of different methods, such as listwise deletion, mean substitution, and maximum likelihood based Expectation Maximization, on risk management because this subject has insufficient literature. The current study tested daily closing data of floating currencies from Kenya (KES), Ghana (GHS), South Africa (ZAR), Tanzania (TZS), Uganda (UGX), Gambia (GMD), Madagascar (MGA) and Mozambique (MZN) in USD denomination against EUR/USD rate between March 8, 2000 and March 6, 2015 acquired from the Bloomberg database. Our results suggested the biases of missingness on Value at Risk and volatility models, presenting significant differences among the number of extreme fluctuations or model parameters. A selective hedging strategy can have different expenditures due to the choice of method. This paper suggests the usage of mean substitution or listwise deletion for daily financial time series due to their tendency to have a close to zero first momentum.

1. INTRODUCTION

Recently, participating in international trade has become an essential factor for economic growth and development not only for developed but also for developing countries, too. However, the success may depend on export competitiveness in which exchange rates play a crucial role and may improve or worsen export competitiveness in a short time. However, long term analyses are important, because export competitiveness may be worsened by the Dutch disease with its distortion effect on exports through the appreciation of a domestic currency. Developing countries are more vulnerable to this phenomenon because of a large amount of aid they receive.

Corporate sector in developing countries is exposed to a high level of uncertainty due to international economic and political conditions, or exchange rate risks. This paper analyses later one, defining the sensitivity of corporate contractual transactions in foreign currencies to exchange rate movements as transaction exposure (Madura 2008). Three policies are available for management: choosing to hedge most of its exposure, not hedging its exposure, or to hedge selectively. Using the option techniques, first one is the most expensive while the last one requires an ap-

propriate benchmark. In order to be able to analyse the long-run effects of the selective hedging strategy, exchange rates must be available for many years, or there should be some methods to be used by a researcher to overcome the missing data challenge.

Regarding multivariate time series analyses, they require synchronized and continuous data for modelling. However, there can be special occasions when one or some data is missing due to lack of a trading activity. This paper focuses on the impact of different missing data handling methods on GARCH and Value-at-Risk (VaR) model parameters, namely the volatility persistence and asymmetry and the fat-tailness of the corrected data. This paper was motivated by the importance of GARCH models during the determination of export-competitiveness (Cheunga- Sengupta 2013, Demir 2013), while VaR models are used in risk management to determine the limitations of acceptable risk towards an investment-position (several times combined with GARCH models, see Fernandez 2005). That is why GARCH and VaR outputs can determine the access to financial resources on micro-level or can be processed to measure macro-competitiveness. A selective hedging strategy is based on the outputs of these econometric models, making corporate competitiveness biased by procedures to overcome missing data problems.

Missing data (or missing values) is defined "as the data value that is not stored in a variable in the observation of interest" (Kang 2013), where time series can be affected by wave nonresponse cases as the suspension of data generating process is only a temporary issue (Graham 2012). There is a strong supposition about the time series: they should not contain missing observations without biasing the ARIMA and GARCH parameters, reduction of representativeness or statistical power due to their impact on the mean, variance, and autocorrelation (Juan Carlos et al. 2010, Kang 2013).

The main added value of the current paper is a comparison of the impact of different methods, such as listwise deletion, mean-substitution, and maximum-likelihood-based Expectation Maximization, on option pricing and costs of a selective hedging strategy because this subject has insufficient literature.

This study tested daily closing data of floating currencies from Kenya (KES), Ghana (GHS), South Africa (ZAR), Tanzania (TZS), Uganda (UGX), Gambia (GMD), Madagascar (MGA) and Mozambique (MZN) in USD denomination against EUR/USD rate between March 8, 2000 and March 6, 2015 acquired from the Bloomberg database. Sample countries have limited resilience towards foreign exchange exposure due to their weak fundamentals and mono-export economies (Cypher – Dietz 2004, Udvari 2014, Vijayakumar 2013), making them even more vulnerable to the results of GARCH models or VaR-based risk management.

The paper has the following structure: Theoretical Background section summarizes the main assumptions behind missingness of data, pointing to some differences between query and time series data. Then, listwise deletion, mean substitution, and Expectation Maximization approaches were introduced in Methods section as well as GARCH models, a basic Value-at-Risk and currency option pricing. The Results and Data section presents the statistical properties of raw, unsynchronized time series, and compares it with the three approaches to identify their tendencies for bias.

2. THEORETICAL BACKGROUND

Competitiveness of companies is accepted and widely analysed, but the competitiveness of territorial unites (e.g., cities, regions or countries) is a subject of a strong debate: some scholars accept its existence but some deny (Krugman 1994, Lengyel 2016, Lukovics 2008). However, recently, it has been broadly accepted that competitiveness of regions can also be analysed. Regarding competitiveness, there are several definitions (e.g., Lukovics 2008, Samuelson – Nordhaus 2000, Sala-i-Martin 2010) and reports published by international organizations (e.g. World Economic Forum or IMD, but the EU has also published its own competitiveness reports). These reports rank countries according to their competitiveness level through measuring it by a set of indices. All these refer to the fact that competiveness is a complex concept. In this study, we rely on the definition of the EU published in 1999 in its sixth periodical report. According to it, competitiveness is *'the ability of companies, industries, regions, nations and supra-national regions to generate, while being exposed to international competition, relatively high income and employment levels'* (EC 1999, p. 75.). Based on this definition, the entity which we analyse, must participate in international competition, therefore, we focus on export competitiveness of countries in our research. The reason is that fluctuation of exchange rates may hinder or boost export performance and competitiveness, and because of the definition, all these may result in better (or worse) overall competitiveness of a country.

2.1 Export competitiveness

Participating in international trade is a crucial point for development, especially in (African) developing countries. Many researchers accept that higher export performance can contribute to economic growth (Ekholm – Södersten 2002, Freund – Bolaky 2008), but some also add that export can lead to economic development and result in poverty reduction, too (Dollar – Kraay 2003, Hallaert – Munro 2009, UNCTAD 2005). This concept is also supported by numerous international organizations: in 2005, the United Nations declared in the framework of the Millennium Development Goals that a market access of developing countries should be developed; in 2005, the World Trade Organization together with the OECD launched the Aid for Trade initiative in order to improve the supply-side capacity in developing countries so that they participate in international trade more effectively; or in 2015, the Sustainable Development Goals (accepted in the United Nations) have the aim to significantly increase the exports of developing countries and facilitate a market access of least developed countries (Udvari 2014; UN 2016). However, there is still a massive debate on the size of the effects that export performance can have on economic growth, development and poverty reduction (Hoekman – Özden 2005, Lee 2005, Subasat 2002). The existence of this debate is due to export competitiveness: the impacts of export performance is highly influenced by its competitiveness.

Export competitiveness depends on several factors. For example, one should take into consideration the size of the country, whether the sector has a comparative advantage or not (Yanikkaya 2003); how strong, strict and flexible the institutions are in a country (Freund – Bolaky 2008), how diversified the export is (Haddad et al. 2010), how large the trade costs are (Lombardi et al. 2016) and how complex the exported product is (Erkan – Yildirimci 2015). In this study we are focusing on the African countries, therefore, we have to mention the role of resource-en-

dowment as an important factor of export competitiveness. This is in close relationship with diversified export. It is shown that most of the African regions depend on the export of natural resources: more than 40 percent of their export consist of oil, palm oil or diamond (Vollmer et al. 2009). Furthermore, developing – and mainly African – countries depend on aid, since the proportion of aid from their GDP exceeds 10 percent, which may cause problems in export competitiveness because of the Dutch disease (Collier 2008; Doucouliagos – Paldam 2007; Rajan – Subramanian 2011). Dutch disease means that a large amount of foreign currency flows into a country (for example, due to aid), which results in real appreciation of the domestic currency. This leads to more expensive exports resulting in decreasing export competitiveness (McKinley 2005), that is, the country may lose its comparative advantage in the traded sector. As a result, as Collier (2008) points out, aid and export promotion cannot be competitors and only one of them can be supported to achieve economic growth and development. In this way, it is important to consider: Aid for Trade-type assistance can contribute to the reduction of trade costs (Lanz et al. 2016; Melo – Laurent 2016) which may result in better export competitiveness, but being aid and foreign currency flow, Aid for Trade may also result in worsening export competitiveness through Dutch disease and the appreciation of a domestic currency. Therefore, it is really important to be able to analyse long-term time series in the case of exchange rates and have a tool used by a researcher to overcome the missing data problem.

2.2 Missingness in financial time series

Financial time series, like daily closing currency data, can be missing due to lack of a trading activity on the specific data – while other markets are open. Therefore, the phenomena have a multivariable-dimension. This temporary suspension of market data can be a result of national differences, holidays and weekends, or by market forces like illiquid situations (in small-cap shares mostly) or when a trading activity is suspended due to a sudden collapse in pricing. A huge literature considers how pricing and market efficiency is affected by such brakes as the most cited "weekend effect" appears (Keim – Stambaugh 1984, Robins – Smith 2015, Zafar et al. 2012).

The literature distinguishes among three forms of mechanism behind missingness (Graham 2012, Junger – Leon 2015): one can assume that data is missing completely at random (MCAR), when missingness does not depend on the values of data or other observed particular variable and their exclusion do not bias our estimations due to their homogeneity (Enders 2010, Junger – Leon 2015, Kang 2013). Missing at random (MAR) happens when dropout is conditionally independent of the variable (Kang 2013), but we can assume some sort of mechanism behind the missingness (Graham 2012). Their exclusion may corrupt temporal structures, such as autocorrelation, trends, and seasonality (Junger – Leon 2015). Missing no at random (MNAR) case occurs when it is possible to make and unbiased estimation to model the missing data. When missingness is beyond researcher's control (their distribution is unknown), MAR is only an assumption (Graham 2012).

Following Baraldi et al. (2015), there are three different approaches to assess the missing data problem. First, we can remove those time intervals when there is at least one missing data for a specific date. Listwise deletion or last observation carried forward a scheme can make time series more fragmented or may introduce bias in the estimation of the parameters unless there is

a chance that our missingness is MCAR (Kang 2013). The second approach substitutes the missing data by the unconditional mean value or the median (for skewed data, suggested by Junger – Leon 2015) of the historical data available. It has a similar impact like the last observation carried forward a scheme for the calculated logarithmic returns for time series with zero mean and mode. This solution is not recommended by Graham (2012) due to its distortions to make a higher concentration around the mean and underestimate errors and variance at MCAR states (Junger – Leon 2015, Enders 2010). Third, there are modern, computation-based approaches to reconstruct missing data trough minimization of an error function, derived from mean, variance or a likelihood ratio (Baraldi et al. 2015, Ceylan et al. 2013, Juan Carlos 2010). Expectation maximization (EM) models applying maximum likelihoods to estimate variance, covariance matrixes of the data, while there are also neural networks-based and genetic structure-based approaches as well (Ceylan et al. 2013, Juan Carlos 2010). The expectation maximization takes more computation time, because EM algorithm may be as difficult to compute as the likelihood function itself (Ruud 1991) and they require more specification of a data generation model (Houari et al. 2013), but they do not rely on the MCAR requirement which is a feature that remains to be fully exploited. The unbiasedness under MAR and higher efficiency under MCAR make maximum likelihood the method of choice in a situation with incomplete multinormal data (Wothke 1998). They are less biased than listwise and pairwise deletion and mean-imputation methods, but this advantage depends on the missing data rate, the covariance structure of the data and size of the sample (Wothke 1998).

Missing data problems can affect daily time series under multivariate applications, volatility spillover, extreme fluctuation or contagion modelling, where assumptions about conditional variance, covariance and correlation are critical.

3. METHODS

This paper applies and compares three different missing value handling methods to capture their ability to maintain central moments, autocorrelation, volatility persistence and extreme movements. All methods were used on daily closing data of African floating currencies and EUR in USD denomination between March 8, 2000, and March 6, 2015. After summarizing the theoretical aspects of missingness in the previous section (following Graham (2012) and Junger – Leon (2015)), this chapter defines the operative approaches to overcome this problem by Baraldi et al. (2015). It is also important to test the influence of these methods on the underlying data as it is the reason of the introduction of volatility models, following Cappeiello, Engle and Sheppard (2006). If the data is biased by missing data handling methods, volatility can be different with an impact on risk management as well as hedging, where the aim is to transfer an unwanted risk to another party (Heckinger, 2013). If the frames of this transfer are not clear, then hedging expenditures can be higher, with an adverse impact on corporate export competitiveness.

Let us assume n foreign exchange rates (1), where the ith ($1 \leq i \leq n$) currency has the following p price for every y trading day with v sample size:

$$P_i = [(y_1 \cdot p_{i,1} / y_v \cdot p_{i,v}].\tag{1}$$

There is also a kth ($1 \leq k \leq n$, and k≠i) currency (2) with w data, and χ ($\chi \neq y$) time indices:

$$P_k = [z_1 \cdot p_{k,1} \, / \, z_w \cdot p_{k,w}]. \tag{2}$$

Upper $P_{1,...,i,k,...,n}$ matrices should be united for the purposes of multivariate analysis which requires the synchronization of time indices.

Listwise deletion (3) means a T cap of specific time indices to exclude all cases where at least one value is missing:

$$T = Y \cap Z. \tag{3}$$

Mean substitution (4) can be applied only on logarithmic returns due to their near zero mean and mode. Assuming that we need returns in the end, the Last Observation Carried Forward (LSCF) scheme has similar benefits on prices – with zeroed differentials:

$$T = (Y \cup Z) \text{ with } p_{i,o} = p_{i,o-1} \text{ for } T \notin (Y \cap Z) \text{ where } r_{i,o} = p_{i,o} - p_{i,o-1} \, . \tag{4}$$

The LSCF procedure requires an addition of a very small positive $\varepsilon = 10^d$, $d \rightarrow +\infty$ number to satisfy the $e^{\, p_{i,o} - p_{i,o-1}} \neq 0$ requirement for a $p_{i,o} = p_{i,o-1}$ case if we would like to use logarithmic returns. The inclusion of ε will provide an asymptotical result $e^{\varepsilon + p(i,o) - p(i,o-1)} \neq 0$ for $p_{i,o} = p_{i,o-1}$ cases as well: $e^\varepsilon \approx 0$.

Regularized expectation-maximization (EM) algorithm is based on iterated linear regression analyses, but it replaces the conditional maximum likelihood estimation of regression parameters for Gaussian data (5), following Schneider (2001). For each $p_{t,i} \in P$ with missing values, the relationship between the prices with missing values (trading days) and the prices with available values is modelled by a linear regression model:

$$p_{NaN} = \mu_{NaN} + (p_a - \mu_a)B + \varepsilon \tag{5}$$

Where a represents the data available, and $B \in R^{\, n_a \times n_{NaN}}$ is a matrix of regression coefficients with covariance matrix with missing and available data from n all sample markets. The $\varepsilon R^{1 \times n_{NaN}}$ residual is assumed to be a zero-mean and $C \in R^{n_{NaN} \times n_{NaN}}$ unknown covariance matrix vector. In each iteration of the EM algorithm, estimates of the mean $\mu \in R^{1 \times n}$ and of the $\Sigma \in R^{n \times n}$ covariance matrix are taken as given, and from these estimates, the conditional maximum likelihood estimates of the matrix of regression coefficients B and of the covariance matrix C of the residual are computed for each record with missing values.

A sensitivity analysis is required to examine the bias of an uncertain input on the model, where the maintenance of the central tendencies, autocorrelation is studied as well as the patterns of the percentage of missing data (Kang 2013, Graham 2012). Variance models can be affected by missing data, making a model selection and parameterization biased. Different GARCH models were fitted to the data (with Oxford MFE and UCSD toolboxes) to analyze patterns of volatility persistence, following Cappeiello, Engle and Sheppard (2006). The applied GARCH(p,q), GJR GARCH(p,o,q), TARCH(p,o,q) and APARCH(p,o,q) (1-5) models can be useful to capture volatility developments and their clustering in time (heteroscedasticity).

$$\text{GARCH (p,q):} \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_i \, \sigma_{t-j}^2. \tag{6}$$

where σ_t^2 represents present variance, ω is a constant term, p denotes the lag number of squared past ε_{t-i}^2 innovations with α_i parameters, while q denotes the lag number of past σ_{t-j}^2 variances with β_i parameters to represent volatility persistence. Asymmetric GARCH models can be introduced via

$$\begin{cases} S_{t-i}^- = 1, if\ \varepsilon_{t-i} < 0 \\ S_{t-i}^- = 0, if\ \varepsilon_{t-i} \geq 0 \end{cases} \text{ as a sign asymmetric reaction to decreasing returns.} \tag{7}$$

GJR GARCH (p,o,q): $\sigma_t = \omega + \sum_{i=1}^p \alpha_i |\varepsilon_{t-i}| + \sum_{i=1}^o \gamma_i S_{t-i}^- |\varepsilon_{t-i}| + \sum_{j=1}^q \beta_i \sigma_{t-j}$, $\tag{8}$

TARCH (p,o,q): $\quad\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^o \gamma_i S_{t-i}^- \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_i \sigma_{t-j}^2$, $\tag{9}$

APARCH (p,o,q): $\quad\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta$, $\tag{10}$

where $\alpha_i > 0\ (i=1,…,p)$, $\gamma_i + \alpha_i > 0\ (i=1,…,o)$, $\beta_i \geq 0\ (i=1,…,q)$, $\alpha_i + 0.5\ \gamma_j + \beta_k + < 1\ (i=1,…,p, j=1,…,o, k=1,…,q)$ and δ index parameter can be between 1 and 2.

The model selection was made with a focus on homoscedastic residuals (using a 2 lagged ARCH-LM test), searching for the lowest Bayesian Information Criteria (BIC).

The missing values have an impact on the density function of the data – listwise deletion assumed to make more data on the tails, while mean substitution can increase the representation of the expected value. The EM should produce data between these mean and extremes. An extreme fluctuation of the data was tested with ordinary Value-at-Risk (1%) (11) and (5%) models (12), where the weight of extreme data and the kurtosis of non-extreme data were the variables of my sensitivity analysis.

$$VaR\ (1\%): r \in (r_n \cup r_x^- \cup r_x^+),\ where\ r_x^- < \mu - 2.326 * \sigma\ and\ r_x^+ > \mu + 2.326 * \sigma \tag{11}$$

$$VaR\ (5\%): r \in (r_n \cup r_x^- \cup r_x^+),\ where\ r_x^- < \mu - 1.65 * \sigma\ and\ r_x^+ > \mu + 1.65 * \sigma \ , \tag{12}$$

where r is a logarithmic return, μ unconditional mean, σ conditional standard deviation from a GARCH model, r_x^- represents extreme negative, r_x^+ extreme positive returns and r_n denotes a non-extreme subset of data (Madura 2008). VaR (5%) has the tendency to define more return as extreme (~5% of the data on each tail), so it can be used better to highlight the difference between missing data approaches. However, a selective hedging requires a low amount or signals, that is why VaR (1%) approach will be used there.

The Biger and Hull model was used to calculate call (13) and put (14) currency option fees:

$$European\ call = e^{-r^*T} SN\left\{ \frac{ln\frac{S}{X} + T[r-r^* + \frac{\sigma^2}{2}]}{\sigma\sqrt{T}} \right\} - e^{-rT} XN\left\{ \frac{ln\frac{S}{X} + T[r-r^* - \frac{\sigma^2}{2}]}{\sigma\sqrt{T}} \right\} \tag{13}$$

$$European\ put = Europen\ call + e^{-rT} X - e^{-r^*T} S \tag{14}$$

where r represents domestic interest rate (in USD, downloaded from stooq.com), r* is foreign interest rate (in African sample currency, downloaded from tradingeconomics.com), S spot exchange rate, X target exchange rate, T remaining time till maturity in years, e natural logarithm, N(.) is standard normal cumulative distribution function and σ conditional standard deviation from GARCH model (Madura 2008, pp. 136).

This paper applies the following setup when comparing of three different methods: GARCH models are fitted at first to provide conditional standard deviations both for VaR (1%) and option pricing models, then, VaR (1%) signals are detected. Sample currencies tend to depreciate in the analysed period, therefore, appreciation can be an unwanted phenomena from operation

point of view. By using applied selective hedging strategy, the company decides to hedge with one year options when an appreciation VaR (1%) signal is detected – then keeps waiting until this option expires. Overall costs of the strategy compared by using of different missing data handling methods.

4. RESULTS AND DATA

Statistical properties of raw and refined data were compared in this section to present the underlying differences among missing value handling approaches and their impact on a model parameterization.

4.1 Raw data

Floating African currencies, the Euro-fixed CFA Franc (XAF) and Euro in USD denomination was tested between March 8, 2000, and March 6, 2015. CFA Franc (XAF) followed strictly the euro only, due to its fixed regime, showing an appreciation against US dollar during the entire time set in Figure 1. Kenyan Shilling (KES) and South African Rand (ZAR) presented an appreciating trend before the subprime crisis only, otherwise, the entire currency set suffered from depreciation.
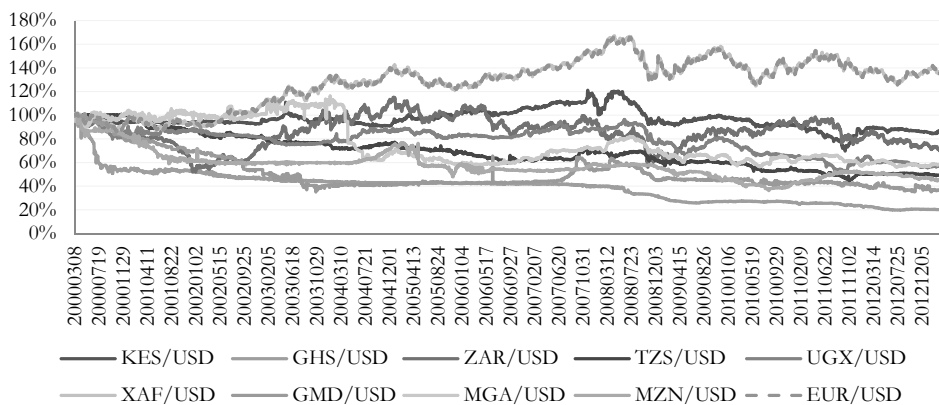

Fig. 1 – Developments of Selected African Currencies between 2000 and 2015 (March 8 2000=100%). Source: Bloomberg

Logarithmic returns of the raw data had zero mean and a low standard deviation, while symmetry appeared only for EUR and GMD (Table 1). The excess kurtosis presented a higher-than-expected occurrence of extreme fluctuation – pegged XAF and EUR presented a moderated level only. None of the currencies followed a normal distribution and most of the data suffered from autocorrelation (except EUR) and heteroscedasticity (except KES, ZAR, and EUR) at 2 lags. The entire sample was weak stationary.

Tab. 1 – Descriptive statistics of the raw data. Source: author's calculation, using Kevin Sheppard's MFE Toolbox for Matlab

currency	mean	std	skewness	kurtosis	normal distribution	auto-correlation	hetero-scedasticity	station-arity
					Jarque-Bera (p)	Ljung-Box (p)	ARCH-LM (p)	ADF (p)
KES/USD	0.00	0.01	-0.32	20.51	0.00	0.00	0.10**	0.00
GHS/USD	0.00	0.01	-1.23	33.77	0.00	0.00	0.00	0.00
ZAR/USD	0.00	0.01	-1.05	15.74	0.00	0.01	0.14**	0.00
TZS/USD	0.00	0.01	0.82	39.78	0.00	0.00	0.00	0.00
UGX/USD	0.00	0.01	-0.47	16.76	0.00	0.00	0.02	0.00
XAF/USD	0.00	0.01	0.13	5.14	0.00	0.00	0.00	0.00
GMD/USD	0.00	0.02	0.05	169.41	0.00	0.00	0.03	0.00
MGA/USD	0.00	0.01	-1.63	54.41	0.00	0.00	0.00	0.00
MZN/USD	0.00	0.01	-0.80	42.11	0.00	0.00	0.00	0.00
EUR/USD	0.00	0.01	-0.02	4.39	0.00	0.52*	0.59**	0.00

Notes: *: non-autocorrelated at 2 lags, **: homoscedastic at 2 lags

Four (GARCH, TARCH, GJR-GARCH, APARCH) GARCH models with 13 different lag-number setup were fitted on the raw dataset to find the most fitting model trough searching for the lower BIC. Half of the sample preferred asymmetric models (except GHS, TZS, UGX, EUR), but previous volatilities had a major role in the estimation of present volatility – innovations were important at GMD only (Table 2).

Tab. 2 – GARCH models of the raw data. Source: author's calculation, using Kevin Sheppard's UCSD Toolbox for Matlab

currency	model	con-stant	alpha 1	alpha 2	gamma	beta 1	beta 2	BIC
KES/USD	TARCH(1,1,2)	0.00	0.26		-0.03	0.46	0.29	-4.20
GHS/USD	GARCH(1,2)	0.00	0.11			0.37	0.51	-3.70
ZAR/USD	GJR GARCH(1,1,1)	0.00	0.12		-0.10	0.93		-3.21
TZS/USD	GARCH(1,1)	0.00	0.21			0.79		-3.91
UGX/USD	GARCH(1,1)	0.00	0.20			0.80		-3.82
XAF/USD*								
GMD/USD	TARCH(2,1,1)	0.00	0.30	0.30	-0.16	0.48		-2.81
MGA/USD	GJR-GARCH(1,1,2)	0.00	0.02		0.03	0.47	0.49	-3.18

MZN/USD	GJR-GARCH(1,1,2)	0.00	0.26		-0.11	0.29	0.51	-3.32
EUR/USD	GARCH(1,1)	0.00	0.04			0.96		-3.71

Notes: *: none of the models were able to provide homoscedastic residuals with normal distribution

Value-at-Risk (Table 3) was able to create a close-to-symmetric set of non-extreme returns, while kurtosis dropped under 5. Extreme fluctuations had a lower weight than 10% (except the 11% of XAF and EUR), so the method was able to capture those rare cases, which were responsible for most of the fat tailness of the data.

Tab. 3 – Value-at-Risk (5%) properties of the raw data (in USD). Source: author's calculation, following Madura (2008)

currency	KES	GHS	ZAR	TZS	UGX	XAF	GMD	MGA	MZN	EUR
mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
standard deviation	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00
skewness	0.07	-0.14	-0.04	-0.02	-0.02	0.02	-0.04	0.04	0.05	0.01
kurtosis	3.48	4.34	2.51	3.16	2.84	2.52	4.07	3.68	4.24	2.48
X- treshold	-0.01	-0.02	-0.02	-0.01	-0.01	-0.01	-0.03	-0.02	-0.02	-0.01
X+ treshold	0.01	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.01
Number of X-	148	121	179	136	149	190	105	100	123	189
Number of X+	138	128	145	121	127	178	95	90	107	207
Number of non-extreme returns	3606	3573	3589	3578	3563	3455	3589	3424	3598	3517

4.2 Comparison of Methods

MGA currency suffered from missingness at most (9%), while ZAR and EUR had none of them (KES: 1%, TZS, UGX, MZN: 2%, GHS, XAF, GMD: 3%). The listwise deletion was the most restrictive approach, while other two methods made a less dramatic reduction in the length of the entire dataset.

Descriptive statistics of refined data by three different approaches had no significant difference according to the paired t-test. The mean remained close to zero, but standard deviation doubled or tripled in 60% of the cases at EM method. The asymmetry of the data was completely distorted by all the methods, but kurtosis increased in 40% of the cases and remained at the previous level in the other 40% at listwise deletion. The kurtosis increased in half of the cases under mean substitution or remained stable. The kurtosis dramatically increased when using EM method.

The data remained non-normal distributed and weak stationer, and there were no significant changes in the autocorrelation of heteroscedasticity properties.

The results of Value-at-Risk (5%) have the same message as kurtosis, where the EM approach provided significantly less VaR (5%) signals, but the "non-extreme" subset suffered from the significant increase in kurtosis in 80% of the sample (except ZAR and EUR). It means that VaR-based risk management can be biased by the missing data if it is managed trough EM methodology. The listwise approach presented a significantly lower impact on VaR properties.

The listwise approach in volatility modelling had a moderate but significant impact on parameters only, and suggested a different model for MGA and MZN, while it was now possible to fit a GARCH model to XAF data. The innovation parameters increased while previous volatility decreased a bit, and the models presented a better fit – despite the expected higher fragmentation of the approach. Mean substitution pushed MGN and GHS currencies towards a more complicated APARCH model, but only GHS lost its former symmetric design. This approach increased the parameters of volatility persistence with similar BIC. The EM approach suggested asymmetric models instead of former symmetric models (for KES, GHS, TZS, UGX), while three former asymmetric preference decreased to symmetry (GMD, MGA, MZN) – but BIC increased almost everywhere, suggesting that it was harder to find well-fitting models with homoscedastic residuals. ZAR and EUR were completely unaffected by different approaches (despite that they had to lose the most value to meet listwise deletion standards), while MGA and MZN were completely the subject of missing data management.
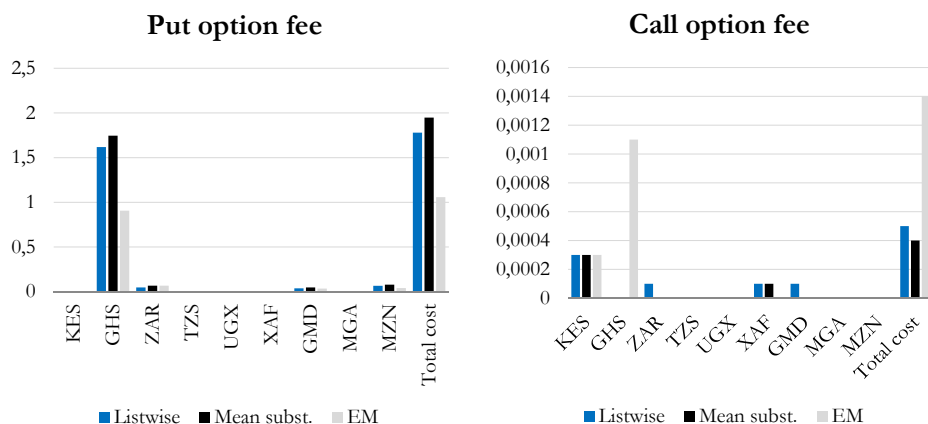

Fig. 2 – Cost of a selective hedging strategy (against USD). Source: author's calculation

Companies with transactional exposures towards local currencies and USD can follow a selective hedging strategy, when they decide to hedge their positions with 1-year-long call or put options under extreme appreciations – marked by VaR (1%). Assuming the company uses USD in its books and has the same positions in sample currencies, the hedging can have completely different cost – a 84% gap for put options and 250% for call options (figure 2)! The missing data can have a significant impact on profitability under these assumptions.

5. CONCLUSION

Floating African currencies were studied in this paper from the aspect of missing data handling methods and by their bias on selective hedging strategies. The analysed dataset has different properties in the case of the conditional volatility and extreme fluctuation, after the application of three different mainstream approaches to overcome missing trading days. It means that option pricing can have three different "fair" prices according to the outputs of Biger and Hull currency option pricing model. Export competitiveness can be adversely affected if selective hedging expenditures depend on decisions on a model selection.

Unfortunately, the missing data analysis literature focuses mostly on the aspects of query-type data, or time series were mostly tested on non-financial data. The missing data can appear on the markets of illiquid assets or due to differences among trading activities. The maximum-likelihood-based Expectation Maximization (EM) models are very popular nowadays to manage missing data in query data due to its ability to maintain the covariance matrix of the data. However, compared to listwise deletion or mean substitution methods, the EM method presented huge biases on daily closing data of financial time series. This application increased the second and fourth moment dramatically, providing poor VaR-signal performance and made conditional volatility more asymmetric. Risk management decisions or competitiveness studies can be also biased by the choice of a method, pointing to the significance of these results.

The results of application comparison in this paper suggest the usage of mean substitution or listwise deletion for daily financial time series due to their tendency to provide similar characteristics to the original time series both for univariate and multivariate cases.

References

1. Baraldi, P., Di Maio, F., Genini, D., & Zio, E. (2015). Reconstruction of Missing Data in Multidimensional Time Series by Fuzzy Similarity. *Applied Soft Computing Journal*, 26, 1–9. http://dx.doi.org/10.1016/j.asoc.2014.09.038

2. Cappeiello, L., Engle, R. F., & Sheppard, K., (2006). Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics*, 4 (4), 537–572. http://dx.doi.org/10.1093/jjfinec/nbl005

3. Cheunga, Y-W., Sengupta, R. (2013). Impact of exchange rate movements on exports: An analysis of Indian non-financial sector firms. *Journal of International Money and Finance*, 39 (12), 231–245. http://dx.doi.org/10.1016/j.jimonfin.2013.06.026

4. Collier, P. (2008). *The bottom billion: Why the poorest countries are failing and what can be done about it*. Oxford University Press, Oxford.

5. Cypher, J. M., Dietz, J. L. (2004). *The Process of Economic Development*. New York: Routledge

6. Demir, F. (2013). Growth under exchange rate volatility: Does access to foreign or domestic equity markets matter? *Journal of Development Economics*, 100 (1), 74–88. http://dx.doi.org/10.1016/j.jdeveco.2012.08.001

7. Dollar, D., Kraay, A. (2003). Institutions, trade, and growth. *Journal of Monetary Economics*, 50 (1), 133–162.

8. Doucouliagos, H., Paldam, M. (2007). The Aid Effectiveness Literature: the Sad Results of 40 Years of Research. *Economic Working Paper*, 2005-15, University of Aarhus.

9. EC (1999). *Sixth Periodic Report on the Social and Economic Situation and Development of Regions in the European Union*. European Commission, Luxembourg.

10. Ekholm, K., Södersten, B. (2002). Growth and Trade vs. Trade and Growth. *Small Business Economics*, 19(2), 147-162. doi:10.1023/A:1016241416227

11. Engle, R. F., (2002). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics*, 20(3), 377–389. http://dx.doi.org/10.2139/ssrn.236998

12. Erkan, B., Yildirimci, E. (2015). Economic Complexity and Export Competitiveness: The Case of Turkey. *Procedia - Social and Behavioral Sciences*, 193, 524-533. doi:10.1016/j.sbspro.2015.06.262

13. Fernandez, V. (2005). Risk management under extreme events. *International Review of Financial Analysis*, 14, 113– 148. http://dx.doi.org/10.1016/j.irfa.2004.06.012

14. Figueroa García, J. C., Kalenatic, D., & López Bello, C. A. (2010). An Evolutionary Approach for Imputing Missing Data in Time Series. *Journal of Circuits, Systems & Computers*, 19(1), 107–121. http://dx.doi.org/10.1142/S0218126610006050

15. Freund, C., Bolaky, B. (2008). Trade, regulations, and income. *Journal of Development Economics*, 87 (2), 309-321. http://dx.doi.org/10.1016/j.jdeveco.2007.11.003

16. Graham, J. W. (2012). *Missing Data Analysis and Design*. New York: Springer

17. Hallaert, J. J., Munro, L. (2009). Binding constraints to trade expansion: Aid for Trade objectives and diagnostic tools. *OECD* Trade Policy Working Paper, 94, OECD.

18. Heckinger, R. (2013). Derivatives Overview. In: Understanding Derivatives – Markets and Infrastrucuture. *Chichago: Federal Reserve Bank of Chichago,* 1-11.

19. Hoekman, B., Özden, C. (2005). Trade preferences and differential treatment of developing countries: a selective survey. *Policy Research Working Paper*, 3566, The World Bank.

20. Houari, R., Bounceur, A., Kechadi, T., & Euler, R. (2013). *A New Method for Estimation of Missing Data Based on Sampling Methods for Data Mining.* CCSEIT http://dx.doi.org/10.1007/978-3-319-00951-3_9

21. Junger, W.L., & Ponce de Leon, A. (2015). Imputation of Missing Data in Time Series for Air Pollutants. *Atmospheric Environment*, 102, 96–104. http://dx.doi.org/10.1016/j.atmosenv.2014.11.049

22. Kang, H. (2013). The Prevention and Handling of the Missing Data. *Korean J Anesthesiol*, 64(5), 402–406 http://dx.doi.org/10.4097/kjae.2013.64.5.402

23. Keim, D. B., & Stambaugh, R. F. (1984). A Further Investigation of the Weekend Effect in Stock Returns. *The Journal of Finance*, 39(3), 819–835. http://dx.doi.org/10.1111/j.1540-6261.1984.tb03675.x

24. Krugman, P. (1994). Competitiveness: A Dangerous Obsession. *Foreign Affairs*, 73(2), pp. 28-44.

25. Lanz, R., Roberts, M., Taal, S. (2016). Reducing trade costs in LDCs: The role of Aid for Trade. WTO Staff Working Papers, ERSD-2016-05, World Trade Organization.

26. Lee, Y.-S. (2005). Foreign direct investment and regional trade liberalization: a viable answer for economic development? *Journal of World Trade*, 39 (4), 701–717.

27. Lengyel, I. (2016). Competitiveness of Metropolitan Regions in Visegrad Counties. *Procedia - Social and Behavioral Sciences,* Volume 223, 10 June 2016, pp. 357–362

28. Lombardi, P., Dal Bianco, A., Freda, R., Caracciolo, F., Cembalo, L. (2016). Development and trade competitiveness of the European wine sector: A gravity analysis of intra-EU flows. *Wine Economics and Policy*, 5 (1), 50-59. http://dx.doi.org/10.1016/j.wep.2015.12.002

29. Lukovics, M. (2008). *Térségek versenyképességének mérése.* (Measuring the competitiveness of territorial units). Szeged: JATEPress.

30. Madura, J. (2008). *International Financial Management.* Mason: Thomson.

31. McKinley, T. (2005). Why is the 'Dutch Disease' always a disease? The macroeconomic consequences of scaling up ODA. Working Paper, 10, International Poverty Centre, United Nations Development Programme.

32. Melo, J., Laurent, W. (2016). Aid for Trade and the Trade Facilitation Agreement: What they can do for LDCs. Woking Paper, P153, FERDI.

33. Rajan, R., Subramanian, A. (2011). Aid, Dutch disease, and manufacturing growth. *Journal of Development Economics*, 94 (1), 106-118. http://dx.doi.org/10.1016/j.jdeveco.2009.12.004

34. Robins, R. P., Smith, G. P. (2016). No More Weekend Effect. *Critical Finance Review,* 5(2), Forthcoming.

35. Ruud, P. A. (1991). Extensions of Estimation Methods Using the EM Algorithm. *Journal of Econometrics*, 49(3), 305–341. http://dx.doi.org/10.1016/0304-4076(91)90001-T

36. Sala-i-Martin, X. (2010): The Economics behind the World Economic Forum's Global Competitiveness Index. In De Grauwe, P. (ed): *Dimensions of Competitiveness.* Cambridge: The MIT Press, pp. 1-18.

37. Samuelson, P.A., Nordhaus, W.D. (2012). *Economics.* New York: McGraw-Hill/Irwin.

38. Schneider, T. (2001). Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. *Journal of Climate*, 14, 853–871. http://dx.doi.org/10.1175/1520-0442(2001)014%3C0853:AOICDE%3E2.0.CO;2

39. Subasat, T. (2009). Does simultaneous implementation of import-substitution and exportpromotion neutralize each other? *Journal of Developing Areas,* 43 (1), 45-63.

40. Zafar, N., Urooj, S. F., Chughtai, S., Amjad, S. (2012). Calendar anomalies: Case of Karachi Stock Exchange. *African Journal of Business Management,* 6 (24), 7261-7271. http://dx.doi.org/10.5897/AJBM11.1847

41. Udvari B. (2013). *Az Európai Unió nemzetközi fejlesztéspolitikája: a kereskedelempolitikai eszközök eredményessége.* Szeged: JATEPress.

42. Udvari, B. (2014). Impacts of Aid for Trade on Trade with the EU The Role of Old and New Member States. *Journal of Global Policy and Governance*, 3 (1), 77-93. http://dx.doi.org/10.14666/2194-7740-3-1-006

43. UN (2016). Sustainable Development Goals. Accessed: www.un.org/ sustainabledevelopment/sustainable-development-goals/ (29 October 2016)

44. UNCTAD (2005). *Developing countries in international trade 2005: Trade and Development Index*. United Nations Conference on Trade and Development, New York and Geneva.

45. Vijayakumar, S. (2013). An Empirical Study on the Nexus of Poverty, GDP Growth, Dependency Ratio and Employment in Developing Countries. *Journal of Competitiveness*, 5 (2), 67-82. http://dx.doi.org/10.7441/joc.2013.02.05

46. Vollmer, S. – Martínez-Zarzoso, I. – Nowak-Lehmann D., F. – Klann, N-H. (2009): EU-ACP Economic Partnership Agreements. Empirical evidence for Sub-Saharan Africa. *Proceedings of the German Development Economics Conference*, Frankfurt a. M. 2009, pp. 29

47. Wothke, W. (1998). *Longitudinal and Multi-group Modelling with Missing Data*. Mahwah. NJ: Lawrence Erlbaum Associates.

48. Yanikkaya, H. (2003). Trade openness and economic growth: a cross-country empirical investigation. *Journal of Development Economics*, 72 (1), 57–89. http://dx.doi.org/10.1016/ S0304-3878(03)00068-3

49. Yozgatligil, C., Aslan, S., Iyigun, C., & Batmaz I. (2013). Comparison of Missing Value Imputation Methods in Time Series: the Case of Turkish Meteorological Data. *Theoretical & Applied Climatology*, 112(1/2), 143–167. http://dx.doi.org/10.1007/s00704-012-0723-x

Contact information

Gábor Dávid Kiss
University of Szeged, Faculty of Economics and Business Administration
Kálvária sgt. 1, Szeged, 6722, Hungary
E-mail: kiss.gabor.david@eco.u-szeged.hu

Marianna Sávai
University of Szeged, Faculty of Economics and Business Administration
Kálvária sgt. 1, Szeged, 6722, Hungary
E-mail: savai.marianna@eco.u-szeged.hu

Beáta Udvari
University of Szeged, Faculty of Economics and Business Administration
Kálvária sgt. 1, Szeged, 6722, Hungary
E-mail: udvari.beata@eco.u-szeged.hu