Journal of Applied Mathematics, Statistics and Informatics (JAMSI), **4** (2008), No. 1

Simulations of the Correlated Financial Risk Factors[1]

MÁRIA BOHDALOVÁ AND ĽUDOMÍR ŠLAHOR

Abstract

Simulation is a widely used tool for generating draws from a lot of stochastic models. In this paper we will describe useful technique - Monte Carlo with dynamic risk factor modeling for construct probability distribution of risk factors. This technique demands the joint distribution to be known (risk factors models can be fitted using different distributional specifications, including nonnormal distributions) and it uses copula theory as a fundamental tool in modeling multivariate distribution. The copula theory allows the definition of the joint distribution through the marginal distributions and the dependence between the variables. It is obvious, that the accuracy of the forecasts of the behavior of the risk factors depends on the sources of error. Usually, the residuals follow normal distribution. However, it has been observed that the Gaussian models do not fit very well the real-life data, e. g. do not allow so called fat tails and asymmetry of observed log returns. Fat tail is a property of some probability distributions (alternatively referred to as heavy tail distributions) exhibiting extremely large kurtosis. In financial world, fat tails of the distributions of the risk factors imply the additional risk. In this paper, we propose to use of a multivariate t-distribution as a simple and powerful tool to simulate heavy tail distributions.

Mathematics Subject Classification 2000: 91B28, 65C05, 91B70
General Terms: Dynamic risk factor modeling
Additional Key Words and Phrases: multivariate t-distribution, Monte Carlo simulation, copula functions, fat tails

1. INTRODUCTION

Modeling and consequently simulating of the risk factors in the correlated markets requires knowledge of their joint (multivariate) distribution. However, the available data regarding the association between risk factors is often limited to some summary statistics (e.g. correlation matrix). In the special case of multivariate normal distributions, the covariance matrix and the mean vector, as summary statistics, completely specify the joint distribution. Generally, specific dependency models have to be used in conjunction with summary statistics. When we take into account higher dimensions, the variety of dependency structures dramatically increases. For given fixed marginal distributions and correlation matrix, we can construct infinitely many joint distributions.

In practice, a wide array of distributions can be used for different risk factors. Some of the commonly used distributions are the normal, the lognormal, geometric Brownian motion, GARCH, and others. The key to a meaningful modeling of

[1]This research was supported by Science and Technology Assistance Agency under contracts No. VEGA-1/3014/06, APVV-1/4024/07 and VEGA-0375-06

the risk factors is making reliable judgments about which statistical distribution is appropriate for which risk factors and estimating the parameters of the selected distributions. Another important issue is the specification of a modeling structure that meaningfully takes into account the interrelationships between different risk factors.

Of course, the accuracy of the forecasts of the behavior of the risk factors depends on the sources of error of the model. Usually, the residuals follow the normal distribution. However, it has been observed that the Gaussian models do not fit very well the real-life data, e. g. do not allow fat tails and asymmetry of observed probability distributions.

The purpose of this paper is twofold: firstly, will be detailed the steps of the Monte Carlo simulation with dynamic risk factor modeling; and secondly, will be compared, in practical example, how the distribution of the error vectors (we use normal and t-distribution) effects to the probability distribution of the risk factors.

This paper is organized as follow. Section 2 introduces some basics definitions, copula function and presents Sklar's theorem. Monte Carlo simulation with dynamic risk factor modeling is presented in Section 3. Finally, section 4 contains an application of the method to a practical example.

2. SOME BASICS DEFINITIONS AND COPULAE

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be the random vector of the n risk factors historical values, with marginal cumulative distribution functions (C.D.F.) $F_1, \ldots, F_n$. The multivariate C.D.F. $F(x_1, \ldots, x_n) = P[X_1 \leq x_1, \ldots, X_n \leq x_n]$, completely determines the dependence structure of random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$. However, its analytic representation is often too complex, making practically impossible its estimation and consequently its use in simulation models.

The most common methodologies use the multivariate conditional Gaussian distribution to simulate risk factor values due to its easy implementation. Unfortunately, empirical evidence underlines its inadequacy in fitting real data.

New simulation methodologies use copula functions - new tool to handle in a flexible way the comovement between risk factors and other relevant variables studied in finance (see [Cherubini 2004]).

Copula have been broadly used in statistical literature. The books of [Joe 1997] and [Nelsen 1999] presented a good introduction to the copula theory. Although copula have been only recently used in the financial area, there are already several applications in this area. The papers of [Bouyé et al. 2000], [Embrechts et al. 2002], [Embrechts et al. 2003] and [Cherubini 2004] provided general examples of applications of copulae in finance. While the method is borrowed from the theory of statistics, it has been gathering strong popularity both among academics and practitioners in the financial world mainly because of the increase of volatility and erratic behavior of financial markets.

The use of copula functions enables the task of specifying the marginal distributions to be decoupled from the dependence structure of variables. This allows us to exploit univariate techniques at the first step, and is directly linked to nonparametric dependence measures at second (see [Cherubini 2004]).

This section, is an introduction to the general definition of copula and Sklar's

Theorem.

Definition 2.1. An n - dimensional copula[2] is a multivariate C.D.F. C, with uniformly distributed margins on $\langle 0,1\rangle$ ($Uniform\,(0,1)$) and the following properties:

(1) $C:\langle 0,1\rangle^n \to \langle 0,1\rangle$;

(2) C is grounded and n-increasing[3];

(3) C has margins C_i which satisfy $C_i\,(u) = C\,(1,\dots,1,u,1,\dots,1) = u$ for all $u \in \langle 0,1\rangle$.

It is obvious, from the above definition, that if $F_1,\dots,F_n$ are univariate distribution functions, $C\,(F_1\,(x_1)\,,\dots,F_n\,(x_n))$ is a multivariate C.D.F. with margins $F_1,\dots,F_n$, since $U_i = F_i\,(X_i)$, $i = 1,\dots,n$, is a uniform random variable. Copula functions are a useful tool to construct and simulate multivariate distributions.

The following theorem is known as Sklar's Theorem. It is the most important theorem regarding to copulae functions since it is used in many practical applications.

Theorem 2.2. *Let F be an n-dimensional C.D.F. with continuous margins $F_1,\dots,F_n$. Then F has the following unique copula representation (canonical decomposition):*

$$F\,(x_1,\dots,x_n) = C\,(F_1\,(x_1)\,,\dots,F_n\,(x_n)) \tag{1}$$

The proof can be found in ([Nelsen 1999], p.18).

The theorem of Sklar [Sklar 1996] is very important, because it provides a way to analyse the dependence structure of multivariate distributions without studying marginals distributions. From Sklar's theorem we see, that for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated. The dependence structure can be represented by an adequate copula function. Moreover, the following corollary is attained from eq. 1.

Corollary 2.3. *Let F be an n-dimensional C.D.F. with continuous margins $F_1,\dots,F_n$ and copula C (satisfying eq. 1). Then, for any $u = (u_1,\dots,u_n)$ in $\langle 0,1\rangle^n$:*

$$C\,(u_1,\dots,u_n) = F\left(F_1^{-1}\,(u_1)\,,\dots,F_n^{-1}\,(u_n)\right) \tag{2}$$

where F_i^{-1} is the generalized inverse of F_i.

To illustrate the idea behind the copula function, we can think about the multivariate Gaussian distribution that is a "standard" assumption in risk management.

Corollary 2.4. *The Gaussian (or normal) copula is the copula of the multivariate normal distribution. In fact, the random vector $\boldsymbol{X} = (X_1,\dots,X_n)$ is multivariate normal iff:*

(1) the univariate margins $F_1,\dots,F_n$ are Gaussians;

[2]The original definition is given by [Sklar 1996].

[3]These properties mean that C is a positive probability measure.

(2) the dependence structure among the margins is described by a unique copula function C (the normal copula) such that[4]*:*

$$C_R^{Ga}(u_1, \dots, u_n) = \Phi_R\left(\phi^{-1}(u_1), \dots, \phi^{-1}(u_n)\right), \tag{3}$$

where Φ_R is the standard multivariate normal C.D.F. with linear correlation matrix R and ϕ^{-1} is the inverse of the standard univariate Gaussian C.D.F.

Although the normal copula does not have a simple analytical expression, it lends itself to simple Monte Carlo simulation techniques.

3. MONTE CARLO SIMULATION

Monte Carlo simulation is a general method of modeling stochastic processes (i.e. processes involving human choice or processes for which we have incomplete information). It simulates such a process by means of random numbers drawn from probability distributions which are assumed to accurately describe the uncertain components of the process being modeled ([Picoult 2002], p.22). Monte Carlo simulation has become a key technology in the financial sector. It can be applied in a variety of settings.

The traditional Monte Carlo simulation method is based on the following.

Definition 3.1. Assume that random variable X has a cumulative distribution function (C.D.F.) F_X. We define F_X^{-1} as

$$F_X^{-1}(q) = inf\{x : F_X(x) \geq q\}, \; 0 < q < 1. \tag{4}$$

Remark that F_X^{-1} is non-decreasing.

LEMMA 3.2. *For any random variable X and any random variable U which is uniformly distributed on* $(0,1)$, *we have that X and $F_X^{-1}(U)$ have the same C.D.F.*

For the proof see [Wang 1998], p.3.

A Monte Carlo simulation of a random variable X can be achieved by first drawing a random uniform number u from $U \sim Uniform(0,1)$, and then inverting u by $x = F_X^{-1}(u)$.

In a similar way, a Monte Carlo simulation of the random vector $\boldsymbol{X} = (X_1, \dots, X_n)$, usually starts with uniform random vector $\boldsymbol{U} = (U_1, \dots, U_n)$. If the random variables $X_1, \dots, X_n$ are independent (correlated), then we need n independent (correlated) uniform random variables $U_1, \dots, U_n$. For a set of given marginals, the correlation structure of the random variables $X_1, \dots, X_n$ is completely determined by the correlation structure of the uniform random variables $U_1, \dots, U_n$ (see [Wang 1998], p.3).

Monte Carlo simulation of the random vector $\boldsymbol{X} = (X_1, \dots, X_n)$ usually assumes that this vector has a multivariate normal distribution with normal marginals $X_i \rightarrow N\left(\mu_i, \sigma_i^2\right)$ (μ_i and σ_i^2 denote the mean and the variance of X_i) and a positive definite covariance matrix Σ:

$$\mu_i = \mathbb{E}(X_i), \; \sigma_i^2 = \mathbb{E}\left((X_i - \mu_i)^2\right), \tag{5}$$

[4] As one can easily deduce from eq. 2.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & c_{1,2} & c_{1,3} & \cdots & c_{1,n} \\ c_{1,2} & \sigma_2^2 & c_{2,3} & \cdots & c_{2,n} \\ c_{1,3} & c_{2,3} & \sigma_3^2 & \cdots & c_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_{1,n} & c_{2,n} & \cdots & c_{n-1,n} & \sigma_n^2 \end{pmatrix}, \tag{6}$$

where $c_{i,j}$ is the covariance between X_i and X_j:

$$c_{i,j} = \mathbb{E}\left((X_i - \mu_i)(X_j - \mu_j)\right). \tag{7}$$

Detailed information about Monte Carlo simulation is e.g. in [Joe 1997], [Jorion 2000], [Jílek 2000], [Crouhy et al. 2001], [Graeme 2007].

In a more complex form, statistical simulation would depend on more complex assumptions than normality and stable correlations. It would incorporate kurtosis and skewness in the probability distributions of changes in market rates, dynamic correlations of changes in market rates (e.g. the correlations could depend on the magnitude of changes of market rates) and fat tails.

In this paper, we describe Monte Carlo simulation with dynamic risk factor modeling ([SAS® 2007]) that is done in the following sequence:

(1) Suppose, that is given a set n of correlated risk factors $X_1, \ldots, X_n$. Each risk factor may be modeled by appropriate model with typical equation:

$$y_i = f_i\left(\boldsymbol{x}, \boldsymbol{y}, \theta_i\right) + \epsilon_i \tag{8}$$

where $i = 1, 2, \ldots, n$, $\boldsymbol{y}$ are the endogenous variables, $\boldsymbol{x}$ are the exogenous variables, θ_i are the estimated parameters, and ϵ_i is error vector:

$$\epsilon_i \sim F_i\left(\chi_i\right), \tag{9}$$

where $F_i\left(.\right)$ is a user-specified distribution function. (The estimation and inference of the parameter vectors θ and χ are based on standard maximum likelihood estimation (MLE) procedures and generalized method of moments (GMM) and they can be estimated using desktop econometric software such as SAS® software.)

(2) We estimate the parameters of the joint distribution $F\left(.\right)$ for each error vector, that is determined by the copula $C(.)$ (e.g. Gaussian copula, see eq. 3).
The estimate of the joint distribution $H(.)$ for the error vector can be constructed as: $F\left(\epsilon\right) = C\left(F_1^{-1}\left(\epsilon_1\right), \ldots, F_n^{-1}\left(\epsilon_n\right)\right)$ (see eq. 2).
(This fixes the distribution for $\boldsymbol{y}$, the endogenous variables, because the parameters of the model equations have already been estimated (in the last step), and the equations can be used to solve for $\boldsymbol{y}$.)

(3) —In the case of the multinormal error vector, we estimate correlation matrix Σ using the C.D.F.s $(F_i\left(.\right))$, along with the inverse standard normal C.D.F. (ϕ^{-1}), which uses the relationship $\phi^{-1}\left(F_i\left(\epsilon_i\right)\right)$.
—In the case of the multivatiate t-distribution of error vector, correlation matrix Σ is created by crossing the normally distributed residuals. The normally distributed residuals are created from the t-distribution residuals using the normal inverse C.D.F. and the t-C.D.F.

(4) This step consists of a generation m (m is the number of Monte Carlo iterations, usually m equals to 1000, 1500 and more) n-tuples of independent $N(0,1)$ variables, that are transformed to a correlated set by using Σ [5]. They are then transformed back to the uniform by using $\phi^{-1}(.)$. These u_i are transformed into a set of draws from the joint distribution by using $\epsilon_i = \phi_i^{-1}(u_i)$.

In the next section, we will show how this method can be used for forecasting of the four interest rate models a fifth-day-ahead.

4. APPLICATION

In the financial sector, interest rates are a very important subject. Interest rates, which can be loosely defined as the price of money, affect the livelihoods of individuals and businesses each and every day. In particular, we are interested in the LIBOR rates since they serve as a widely used reference interest rate for a variety of financial instruments. LIBOR stands for the London Interbank Offered Rate and is an interest rate at which banks can borrow funds, in marketable size, from other banks in the London interbank market. LIBOR rates are currently fixed in various currencies (USD, EUR, GBP, JPY and other major currencies) and maturities (usually 1, 3, 6 or 12 months) and is widely used as an interest rate index. LIBOR is the primary benchmark used by banks and investors worldwide. Aside from fixing the cost of borrowing money, LIBOR is used to calculate interest rates applied to a wide range of contracts including over-the-counter (OTC) instruments such as interest rate swaps, forward rate agreements, syndicated loans, floating rate notes, etc. In our investigation we deal exclusively with six and twelve monthly maturities in both USD and EUR rates.

4.1 Data description

Interest rates are frequently used risk factors, that have influence on some portfolios and therefore the risk managers need to know its future values. The presented theory, in this paper, may be used for prediction LIBOR's interest rates a some days ahead (for example, we will perform the predictions for fifth day ahead).

The database contains 845 daily interest rates of EUR 6M, EUR 12M, USD 6M, USD 12M, from January 2nd 2004 to April 18th 2007. Data follows from `http://www.bba.org.uk/bba/jsp`.

Figure 1 presents the plots of all series risk factors and Table I contains descriptive statistics.

From the figure 1, we can see that both, the EUR and USD time series, are correlated. Table I shows that means of all series are positive. All time series distributions are nearly symmetric, but they do not have normal distributions (p-value of the Shapiro-Wilk test of normality is less than < 0.0001 for all time series).

4.2 The model

In specifying the four variate model we must, at first, define the four models for the marginal variables and then the distribution of the error vectors. The models for the

[5] The standard techniques for simulating correlated changes in n risk factors are Cholesky decomposition and Principal Component Analysis - PCA. For a description see ([Míka 1985], p.80), ([Wang 1998], p.12), ([Bohdalová and Stankovičová 2006], p. 41-52)

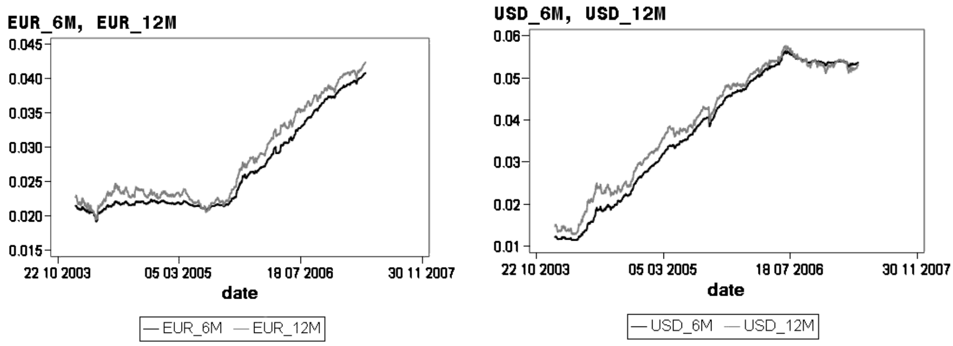

Fig. 1. Daily interest rates of EUR & USD 6M and 12M (LIBOR)

Variable	**Mean**	**Std.Dev**	**Min**	**Max**	**Median**	**Skewness**	**Kurtosis**	**5th Pctl**
EUR 6m	0.02654	0.00654	0.01920	0.04086	0.02210	0.88032	-0.76313	0.02075
EUR 12m	0.02796	0.00686	0.01943	0.04240	0.02383	0.74235	-0.99976	0.02116
USD 6m	0.03753	0.01512	0.01145	0.05640	0.04000	-0.36145	-1.31994	0.01175
USD 12m	0.03942	0.01393	0.01283	0.05766	0.04220	-0.46377	-1.16588	0.01376

Table I. Descriptive statistics of the risk factors

univariate variables must take into account the characteristics of the risk factors. We use known multivariate geometric Brownian motion model of the following form:

$$\begin{aligned} rate_t &= rate_{t-1} + \mu \cdot rate_{t-1} + \nu \\ \nu &= \sigma \cdot rate_{t-1} \cdot \epsilon \end{aligned} \tag{10}$$

For the purpose of this paper we perform Monte Carlo analysis of models that have residuals ϵ distributed as

(1) a multivariate t-distribution $\left(\epsilon \sim t\left(\sigma_1^2, \ldots, \sigma_n^2, df\right)\right)$ [6]
(2) a multivariate normal distribution $(\epsilon \sim N(0, \Sigma))$.

The distributions of interest rates on the fifth day ahead are shown in the figures 2–5 in Appendix. The two curves overlayed on the graphs are a kernel density estimation and a normal distribution fit to the results. The Table II shows real values of the interest rates. Tables III and IV shows selected descriptive statistics for modeled interest rates.

Date	**EUR 6M**	**EUR 12M**	**USD 6M**	**USD 12M**
25/04/07	0.041175	0.042761	0.053469	0.052419

Table II. Real values of the risk factors (April 25th 2007)

[6] A multivariate t-distribution serves the purpose to describe fat tails of the distributions of the risk factors.

Variable	Mean	Std.Dev	Min	Max	Median	Skewness	Kurtosis	5th Ptcl
EUR 6m	0.04116	0.00050	0.03450	0.04566	0.04116	-1.88399	36.47224	0.04049
EUR 12m	0.04269	0.00098	0.03586	0.05187	0.04271	0.18524	9.71065	0.04128
USD 6m	0.05388	0.00095	0.04045	0.06159	0.05389	-1.15021	32.67132	0.05255
USD 12m	0.05305	0.00189	0.01493	0.07685	0.05305	-4.01475	127.0683	0.05068

Table III. Descriptive statistics of the modeled risk factors with the multivariate t-distribution error vectors

Variable	Mean	Std.Dev	Min	Max	Median	Skewness	Kurtosis	5th Ptcl
EUR 6m	0.04123	0.00024	0.04040	0.04194	0.04123	-0.02949	-0.06575	0.04083
EUR 12m	0.04278	0.00044	0.04145	0.04408	0.04279	0.04371	-0.16164	0.04207
USD 6m	0.05388	0.00042	0.05265	0.05527	0.05388	0.08219	-0.09178	0.05321
USD 12m	0.05314	0.00077	0.05068	0.05564	0.05314	0.08129	-0.02712	0.05187

Table IV. Descriptive statistics of the modeled risk factors with the multinormal distribution error vectors

4.3 Concluding remarks

In this paper, we have shown one approach for forecasting the future values of the risk factors. We have used Monte Carlo simulation with dynamic risk factor modeling, that

(1) enables to the risk management practitioners efficiently fit many risk factors models simultaneously,
(2) enables to fit risk factors models using different distributional specifications, including non-normal distributions,
(3) uses the simulation engine based on the statistical concept of a copula theory,
(4) enables to explore the error vectors either by multi-normal or by multivariate t-distributions.

4.4 Acknowledgments

We greatly acknowledge grants from Tatra Banka and SAS Slovakia.

REFERENCES

Bohdalová, M. and Stankovičová, I. 2006. Využitie pca pri analýze rizikových faktorov investičného portfólia. *Forum Statisticum Slovakum* 3, 41–52.

Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G., and Roncalli, T. 2000. Copulas for finance, a reading guide and some applications. Working paper, Financial Econometrics Research Center, City University, London.

Cherubini, U. 2004. *Copula Method in Finance*. Wiley Finance Series.

Crouhy, M., Galai, D., and Mark, R. 2001. *Risk Management*. McGraw-Hill Companies, New York.

Embrechts, P., Lindskog, F., and McNeil, A. 2003. Modelling dependence with copulas and applications to risk management. Handbook of Heavy Tailed Distributions in Finance (Edited by S.T. Rachev), Elsevier, 329–384.

Embrechts, P., McNeil, A., and Straumann, D. 2002. Correlation and dependence in risk management: properties and pitfalls. Risk Management Value at Risk and Beyond (Edited by M. Dempster), Cambridge University Press, 176–223.

Graeme, W. 2007. Risk measurement for financial institutions. http://www.smealsearch.psu.edu/29776.html.

Jílek, J. 2000. *Finanční rizika.* Grada Publishing, Praha.

Joe, H. 1997. *Multivariate Models and Dependence Concepts.* Chapman and Hall.

Jorion, P. 2000. *Value at Risk: The Benchmark for Controlling Market Risk.* McGraw-Hill Professional Book Group, Blacklick, USA.

Míka, S. 1985. *Numerické metódy algebry.* SNTL, Praha.

Nelsen, R. B. 1999. *An Introduction to Copulas.* Springer Verlag, New York.

Picoult, E. 2002. *Risk Management: Value at Risk and Beyond.* Cambridge University Press, Chapter Quantifying the Risks of Trading, 1–59.

SAS®. 2007. Risk dimensions®: Dynamic risk factor modeling methodology. http://www.riskadvisory.com/pdfs/sasriskdimensionsriskfactor.pdf.

Sklar, A. 1996. Random variables, distribution functions and copulas - a personal look backward and forward, in distributions with fixed marginals and related topics. Institute of Mathematical Statistics, (Edited by L. Rüschendorff, B. Schweizer and M. Taylor), Hayward, CA., 1–14.

Wang, S. 1998. Aggregation of correlated risk portfolios: Models & algorithms. *Proceedings of the Casualty Actuarial Society*, 848–939.

Mária Bohdalová,
Department of Information systems,
Comenius University, Faculty of Management,
Odbojárov 10,
820 05 Bratislava 25
e:mail: maria.bohdalova@fm.uniba.sk

Ľudomír Šlahor,
Department of Finance and Economics,
Comenius University, Faculty of Management,
Odbojárov 10,
820 05 Bratislava 25
e:mail: ludomir.slahor@fm.uniba.sk

Received September 2007

Appendix

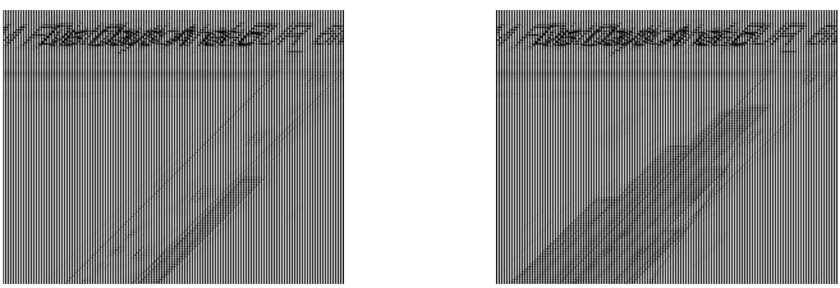

Fig. 2. Distribution of EUR 6M (residuals have multiv.t- and normal-distribution)

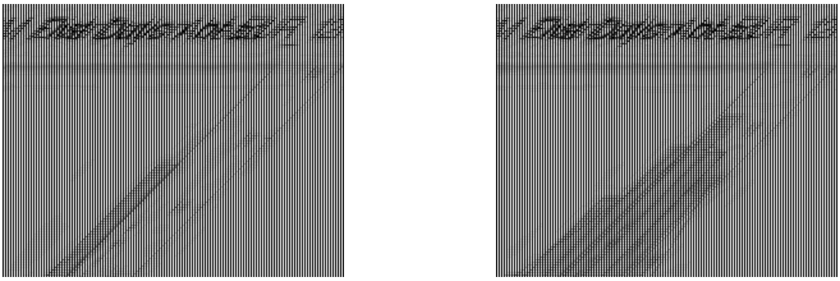

Fig. 3. Distribution of EUR 12M (residuals have multiv. t- and normal-distribution)

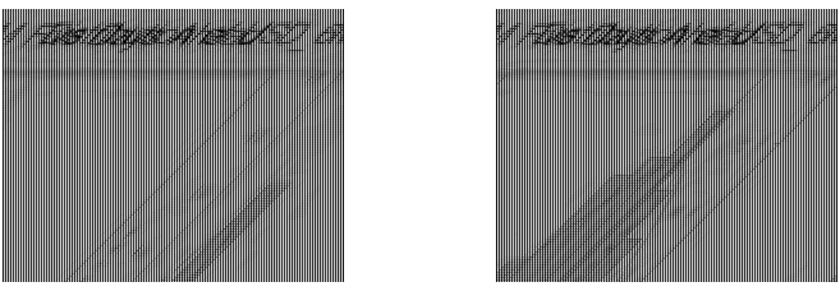

Fig. 4. Distribution of USD 6M (residuals have multiv. t- and normal-distribution)

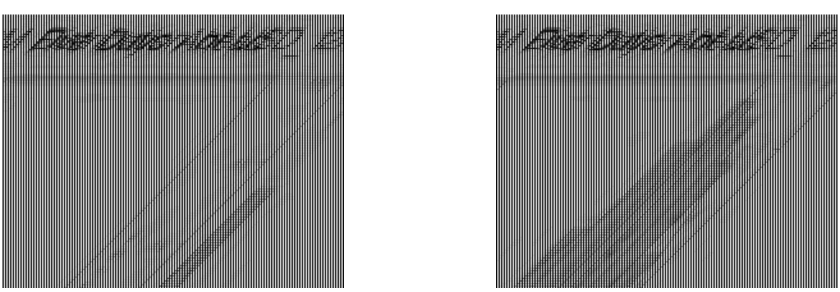

Fig. 5. Distribution of USD 12M (residuals have multiv. t- and normal-distribution)