РЭШ
Российская
экономическая
школа

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

MASTER THESIS

Тема: Динамическое моделирование многомерных распределений: последовательное построение многомерных распределений из распределений меньшей размерности

Title: Dynamic Modeling of Multivariate Distributions: Reconstructing Higher Dimensional Distributions from Smaller Ones

Студент / Student:

Renat Khabibullin
(Ф.И.О. студента, выполнившего работу)

Научный руководитель / Advisor:

Stanislav Anatolyev, PhD in Economics, Tenured Professor at NES
(ученая степень, звание, место работы, Ф.И.О.)

Оценка/ Grade:

_____

Подпись/ Signature:

_____

Москва 2011

Contents

1 Abstract. The problem statement

The problem of dynamic joint distribution estimation is very important from both theoretical and practical points of view: econometricians would be interested in developing new techniques and approaches to model dynamic joint distributions, whereas practitioners (especially, risk and asset managers) would be interested in obtaining dynamic distributions for computing risk measures and making optimal portfolio choices. This paper uses the principles that are similar to Engle's (2009) approach of estimating a vast-dimensional DCC model by merging estimates of pairwise models and introduces a new sequential methodology for dynamic joint distributions modeling based on combining small-dimensional distributions into higher-dimensional ones through compounding and aggregating functions. The new proposition uses marginal and bivariate distributions as inputs, combines them to capture the dependence between one marginal and one bivariate, and then aggregates all of the dependencies to obtain trivariate distributions. Higher-dimensional distributions are built in a similar manner from one-dimension-smaller distributions and univariate ones through compounding and then aggregating them into a single distribution. Additionally, the paper demonstrates how to apply this new sequential technique to model five-dimensional distribution of five DJIA constituents (as of June 8, 2009). Two different types of compounding functions are considered. Kolmogorov-Smirnov goodness-of-fit tests are conducted. Moreover, the paper compares this new methodology to copula-type modeling of distributions based on single five-dimensional time-varying t-copula.

2 Introduction. A brief literature review

The studies of the last two decades show that log-returns of stocks are non-normally distributed and their distribution experience fat tails and skewness (for example, a comprehensive analysis of the properties of financial data is given by Rydberg, 2000). Bollerslev (1987) was first to demonstrate the advantage of combining GARCH models with Student's t-Distribution of innovations to take into account heavy tails. In particular, GARCH models with normal errors do exhibit unconditional excess kurtosis, however it

is not large enough to explain most financial time series, and errors distributed according to Student's t-Distribution or Generalized Error Distribution (GED) typically resolve the problem. However, these approaches do not capture the distribution asymmetry that may appear in financial time series.

A useful method of addressing the asymmetry of financial returns is introducing skewness into the well-known distributions. Hansen (1994) was first to use a Skewed Student's t-Distribution for modeling GARCH innovations in financial time series. His approach is based on rescaling the distribution separately on the left and on the right from the threshold selected in a way to make the expectation to be zero. Fernández and Steel (1998) propose a similar method of introducing skewness to a symmetric distribution by explicit scale changing of the negative and positive its parts. Gallant and Tauchen (1989) use Gram-Charlier series expansion to describe the deviations from normality of errors in a GARCH model, introducing skewness and excess kurtosis. Theodossiou (2000) derives Skewed GED distribution. It also accounts for leptokurtosis and skewness, however like many skewed distributions it is obtained by considering two parts of the symmetric distribution separately, and hence, lacks smoothness. Azzalini and Capitanio (2003) generalize Student's t-Distribution in order to take into account the asymmetry. This approach incorporates both the advantage of using t-Distribution for modeling fat tails and the benefit of allowing for distribution skewness. Their proposal has superiority in the fact that it does not separate negative and positive parts of a distribution and does not restrain the smoothness property of the density.

In addition, Engle and Ng (1993) propose Nonlinear Asymmetric GARCH model (NAGARCH) for taking into account the leverage effect, which is usually observed on the stock markets: negative returns increase future volatility by a larger amount than positive ones of the same magnitude. NAGARCH specification implies introducing asymmetry not in the distribution of innovations, but in the standard GARCH model specification in order to reflect the leverage effect.

To model bivariate joint distributions, copulas are applied as proposed by Sklar (1959). According to his theorem, any multivariate distribution can be decomposed to marginal distributions and the function called copula that describes the dependence

between them. A straightforward extension of Sklar's (1959) theorem to conditional joint distributions is provided by Patton (2006). This theorem is successfully applied to the time-varying multivariate series. The main advantage of this approach is the fact that one can separately describe marginal distributions taking into account their specific features and the dependence structure between them. The most popular copulas are Gaussian copula, Student's t-copula, and the class of Archimedean copulas (see a detailed survey by Nelsen, 2006). Ausin and Lopes (2010) suggest using t-copula instead of normal one, because normal copula implies that there is no interdependence in the tails of the distribution and t-copula is able to capture that dependence.

The main value of this paper is the introduction of a new sequential methodology for dynamic joint distributions modeling based on combining small-dimensional distributions into higher-dimensional ones by compounding and aggregating functions. This idea uses the principles that are similar to the Engle's (2009) proposition of estimating vast dimensional Dynamic Conditional Correlation (DCC) model through merging the estimates of pairwise models. However, the problem of modeling joint distributions is much more complex one, because it takes into account not only the first two moments (like GARCH or DCC model) of the joint distribution, but the whole distribution that incorporates all of the moments. Similar to Engle (2009), on each step of the high-dimensional distribution estimation procedure the new sequential method allows operating only with two objects: a marginal distribution and a one-dimension-smaller distribution. So, the number of parameters used in optimization problem on each compounding step is the same across all dimensions, but the number of parameters that is obtained after aggregating these estimated compounding functions is rather large to ensure the flexibility of the parametric form of the model. Whereas, this number in the standard approaches usually either grows as $O(k^2)$ with increasing dimension k (as they may contain correlation matrix as a parameter) and the estimation is computationally very difficult (or even impossible), or remains constant for all dimensions, and hence, the parametric model is too restrictive.

To perform the first two steps of the new methodology and estimate its basic blocks, this paper combines the advantages of the approaches described above for estimation of

marginal distributions and bivariate copulas. It uses Skew t-Distribution proposed by Azzalini and Capitanio (2003) for innovations in a NAGARCH (Engle and Ng, 1993) structure for modeling marginal distributions of several DJIA-constituents. Besides this, it applies time-varying t-copula suggested by Ausin and Lopes (2010) for modeling dynamic joint distribution of the stocks' log-returns.

For implementing the subsequent steps of the new methodology, (1) asymmetrized time-varying t-copula and (2) asymmetrized time-varying Clayton-copula are used as the compounding functions and compared to each other. Arithmetic mean is used as the aggregating function as it is proved to be the best one among all considered forms.

Goodness-of-fit tests based on Diebold et al. (1998) and Breymann et al. (2003) are conducted to assess the approach for the estimation of both marginals and joint distributions.

To sum up this section, the new sequential approach allows for dynamic modeling of the joint distributions of a number of stocks' log-returns. This means that the whole joint behavior of the log-returns is obtained, rather than just a few moments of the distribution are modeled (as it is done in most of the papers on financial time series modeling). Moreover, this new sequential approach makes it possible to model joint distributions for vast dimensional cases, because it splits a single huge problem into smaller ones and solves them sequentially, and hence, makes the whole problem computationally feasible. Probably, a lot of financial institutions (hedge funds, brokers, banks, investment banks and many others) would be interested in such attractive procedure because they have a lot of open positions in a very huge number of assets. Modeling the whole distribution allows computing not only values at risk, variances and other simple risk measures, but also a lot of complex ones, such as expected shortfalls, that may depend not only on the first few moments of the distribution, but on the whole distribution itself. Additionally, it allows making complex portfolio choices that are based not only on the variance minimization criteria, but also on some more complicated ones, that takes into account higher moments of the joint distribution or even the whole distribution.

The paper is organized as follows. Part 3 briefly outlines Engle's (2009) approach. Part 4 describes the new proposal for dynamic joint distributions modeling based on

transition from small-dimensional distributions to higher-dimensional ones. Part 5 describes models used for the estimation of marginal distributions and joint distributions: the concept of Skew t-Distribution is introduced, GARCH structure for marginals is defined, t-copula along with likelihood function for the estimation is given. Additionally, the two parametric forms for compounding functions are described. Part 6 shows the application of the new sequential technique to five DJIA-constituents and provides with the goodness-of-fit tests for both the marginal distributions and the joint distributions, comparing the new methodology to the standard single-copula based approach. Part 7 outlines further research directions and part 8 concludes.

3 Engle's (2009) approach

Engle (2009) is interested in measuring risk in a highly multivariate framework. In particular, in order to forecast vast-dimensional correlation matrices he uses a simple multivariate GARCH model proposed by Engle (2002) and called Dynamic Conditional Correlation (DCC). The model is used in a standard framework of conditionally normal returns:

$$\begin{aligned} r_t|F_{t-1} &\sim N(0, H_t), \\ H_t &:= D_t R_t D_t, \\ D_t &:= \operatorname{diag}\left(\sqrt{h_{it}}\right), \end{aligned}$$

where r_t is a vector of time t returns, H_t and R_t are respectively its conditional covariance and correlation matrices, F_{t-1} denotes information available at time $t-1$, and h_{it} is conditional variance of an individual return. Define $\varepsilon_t := D_t^{-1} r_t$. Hence, ignoring some constants the log-likelihood function can be expressed as

$$\begin{aligned} L &= -\frac{1}{2T}\sum_{t=1}^{T}\left\{\log \mid H_t \mid + r_t' H_t^{-1} r_t\right\} \qquad (3.1) \\ &= -\frac{1}{2T}\sum_{t=1}^{T}\left\{2\log \mid D_t \mid + r_t' D_t^{-2} r_t\right\} - \frac{1}{2T}\sum_{t=1}^{T}\left\{2\log \mid R_t \mid + \varepsilon_t' R_t^{-2} \varepsilon_t\right\} + \frac{1}{2T}\sum_{t=1}^{T}\varepsilon_t'\varepsilon_t. \end{aligned}$$

The first step of Engle's (2009) approach implies estimation of standard univariate GARCH(P_i, Q_i) model for each individual asset:

$$h_{it} = \omega_i + \sum_{p=1}^{P_i} \alpha_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{i,t-q}.$$

Hence, the DCC model could be written as follows:

$$Q_t = (1 - \sum_{m=1}^{M} \alpha_m - \sum_{n=1}^{N} \beta_n)\overline{Q} + \sum_{m=1}^{M} \alpha_m \varepsilon_{t-m} \varepsilon'_{t-m} + \sum_{n=1}^{N} \beta_n Q_{t-n},$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}, \qquad Q_t^* = \operatorname{diag}\left(\sqrt{q_{iit}}\right),$$

where $\overline{Q}$ can be estimated as $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \varepsilon'_t$.

The second step of Engle's (2009) approach is the pairwise estimation of the DCC model, imposing the following dynamics on the conditional covariance of the assets i and j (here the case of $M = N = 1$ is considered):

$$q_{ijt} = (1 - \alpha - \beta)\overline{R}_{ij} + \alpha \varepsilon_{it-1} \varepsilon_{jt-1} + \beta q_{ij,t-1},$$

where the $\overline{R}_{ij}$'s are the elements of the matrix $\overline{R}$, which can be estimated as $\overline{R} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \varepsilon'_t$. To estimate the model, Engle (2009) suggest a method of approximate log-likelihood maximization of (3.1) by maximizing separately the first sum by volatility parameters and the second sum by the correlation parameters. Hence, the log-likelihood function for this pair become:

$$\begin{aligned} L_{ij} &= -\sum_t \left(\log(1 - \rho_{ijt}^2) + \frac{\varepsilon_{it}^2 + \varepsilon_{jt}^2 - 2\rho_{ijt}\varepsilon_{it}\varepsilon_{jt}}{1 - \rho_{ijt}^2} \right), \\ \rho_{ijt} &:= \frac{q_{ijt}}{\sqrt{q_{iit} q_{jjt}}}. \end{aligned}$$

After maximizing log-likelihood parameters for each pair, one will obtain a set of estimated parameters $\left\{\widehat{\theta}_{ij}\right\}$, where $\widehat{\theta}_{ij} = \left(\widehat{\alpha}_{ij}, \widehat{\beta}_{ij}\right)$. Engle (2009) suggests to aggregate it using so-called *blend function*:

$$
\begin{aligned}
\widehat{\theta} &= b(\widehat{\theta}_{12}, \widehat{\theta}_{13}, \ldots, \widehat{\theta}_{k-1k}), \\
b(\theta, \theta, \ldots, \theta) &= \theta \quad \forall \theta \in \Theta.
\end{aligned}
$$

The corresponding estimate is called *MacGyver estimator*. Engle (2009) shows that a componentwise median as a blend function seems to be the most appropriate one among the functions he considered.

Additionally, Engle et al. (2008) propose a similar method of estimating vast dimensional conditional correlations. The method combines pairwise log-likelihoods into the composite log-likelihood function. As a part of the proposal, they suggest using not all pairs of the assets in the composite log-likelihood, but contiguous pairs or even a certain number of random pairs, if extremely high computational efficiency is required. They demonstrate that the efficiency loss of considering a fixed number of random pairs instead of all pairs can be extremely small, and the computational time benefit is very large.

4 New sequential approach

Engle (2009) proposes the method of estimating vast dimensional DCC model through merging the estimates of pairwise models. This paper applies the same principles for more complicated problem of dynamic joint distributions modeling and suggests combining small-dimensional distributions into higher-dimensional ones by compounding and aggregating functions. The methodology uses the estimates of marginal distributions and that of bivariate copulas as building blocks, and hence, the first two steps of it are in choosing appropriate models for them. The essence of the new sequential approach is described below.

4.1 Algorithm

The basic idea of the whole methodology can be described in the following sequence of steps:

1. estimate the marginal distributions, for example, by Skew-t-NAGARCH model

proposed in this paper below (log-returns of the DJIA-constituents and Russian stocks are proved to be fitted well by this model) or by some non-parametric estimators to improve the fit:

$$\widehat{F}_1, \widehat{F}_2, \ldots, \widehat{F}_K;$$

2. transit from the estimates of the univariate distributions to bivariate ones by estimating them for all pairs of the stocks, for example, through copula-based approach (time-varying t-copula considered in this paper proved to be an appropriate one for modeling joint distributions of DJIA-constituents):

$$\widehat{F}_{12}, \widehat{F}_{13}. \ldots, \widehat{F}_{1K}, \widehat{F}_{23}, \ldots, \widehat{F}_{K-1,K};$$

3. derive trivariate distributions for all groups of three stocks using the following formula:

$$\widehat{F}_{ijk} = \frac{\widetilde{C}^{(3)}\left(\widehat{F}_i, \widehat{F}_{jk}; \widehat{\theta}_{ijk}\right) + \widetilde{C}^{(3)}\left(\widehat{F}_j, \widehat{F}_{ik}; \widehat{\theta}_{jik}\right) + \widetilde{C}^{(3)}\left(\widehat{F}_k, \widehat{F}_{ij}; \widehat{\theta}_{kij}\right)}{3},$$

where $\widehat{F}_{ijk}$ is the distribution function of interest, $\widehat{F}_i$ and $\widehat{F}_{jk}$ are the distribution functions estimated from the first and the second steps respectively, $\widetilde{C}^{(3)}\left(\widehat{F}_i, \widehat{F}_{jk}; \widehat{\theta}_{ijk}\right)$ is a parametrized "distribution" function that captures the dependence between an ith stock and a pair of (j^{th}, k^{th}) stocks (this paper considers two specifications for this function: (1) asymmetrized time-varying t-copula and (2) asymmetrized time-varying Clayton-copula), and $\widehat{\theta}_{ijk}$ is the set of unknown parameters that should be estimated;

4. similarly, estimate the m-dimensional distribution functions for all groups of m stocks: $\forall i_1 < i_2 < \ldots < i_m$,

$$\widehat{F}_{i_1,i_2,\ldots,i_m} = \frac{\sum_{l=1}^{m} \widetilde{C}^{(m)}\left(\widehat{F}_l, \widehat{F}_{i_1,\ldots,l-1,l+1,\ldots,i_m}; \widehat{\theta}_{l,i_1,\ldots,l-1,l+1,\ldots,i_m}\right)}{m};$$

5. finally, the joint distribution for all K stocks:

$$\widehat{F}_{1,2,\ldots,K} = \frac{\sum_{l=1}^{K} \widetilde{C}^{(K)}\left(\widehat{F}_l, \widehat{F}_{1,\ldots,l-1,l+1,\ldots,K}; \widehat{\theta}_{l,1,\ldots,l-1,l+1,\ldots,K}\right)}{K}. \tag{4.1}$$

4.2 Discussion

This section provides motivation behind compounding and aggregating functions and makes some notes on flexibility of the methodology and possible computational improvements.

Naturally, the intuition for the third step of the procedure is the following: similar to Engle's (2009) pairwise consideration, compounding functions $\widetilde{C}^{(3)}\left(\widehat{F}_i, \widehat{F}_{jk}; \widehat{\theta}_{ijk}\right)$ capture the dependence of a pair (an ith stock and a pair of (j^{th}, k^{th}) stocks), whereas similar to Engle's (2009) blend function, the aggregation merges three estimated compounding functions for the three stocks into a single estimate of their trivariate distribution function. Note, arithmetic mean as an aggregating function is used in this paper because it is proved to be the best among considered ones, however, some other functions can be considered.

Alternatively, for m-dimensional case the following intuition for compounding and aggregating functions can be provided: $\widetilde{C}^{(m)}$ can not capture "all of the dependence" between every stock (like Engle's (2009) pairwise estimation can not take into account all influence of all assets on its first step), however, it tries to capture the dependence between a stock and a group of stocks. After applying these functions, aggregation is conducted in order to capture "all of the dependence" (similar to Engle's (2009) blend function that merges pairwise estimates in the single one), and may be, in some sense aggregate out the error (like taking mean of several estimates in order to integrate out the error).

Again, as it was stated in the introduction, on each step of the high-dimensional distribution estimation procedure the new methodology operates only with two objects: a marginal distribution and a one-dimension-smaller distribution. This allows the number of parameters used in each optimization problem on each step of compounding function estimation to be the same across all dimensions, but overall number of parameters that

is obtained after aggregating these estimated compounding functions is rather large to ensure the sufficient flexibility of the parametric form of the model. As for the standard single-copula approaches, this number is usually either grows as $O(k^2)$ with increasing dimension k and the estimation becomes computationally very difficult, or remains constant across all dimensions, and hence, the parametric model for distribution estimation is too restrictive (additionally, see Section 6.5.1 for discussion on number of parameters while using the new sequential approach and the standard one).

Additional notes:

- the aggregation function can be selected different from the arithmetic mean (this requires further research) and it may even include some parameters in order to be more flexible;
- for reducing computational costs one can use the approach similar to one of the approaches by Engle et al. (2008), and if the number of stocks is very large, consider choosing pairs, triples, etc. randomly, instead of considering all of them in the aggregating functions.

This section has demonstrated the ideas of the new sequential approach for joint distributions modeling. It seems to be a reasonable instrument to make the estimation of vast dimensional joint distributions feasible and sufficiently flexible. The next section proposes models for the entities that have been considered in this section.

5 The models

In this section the basic model capturing heavy tails and skewness in marginal distributions along with the models for bivariate copulas and for compounding functions is described. First of all, the Skew t-Distribution used for innovations is described. Then, the model for marginal distributions with NAGARCH structure is shown. After that, the t-copula model is demonstrated. And finally, the two specifications for compounding functions based on asymmetrized time-varying t-copula and Clayton-copula are considered.

5.1 Skew t-Distribution

Azzalini and Capitanio (2003) propose one of the possible generalizations of Student's t-Distribution. It is able to capture fat tails and skewness observed in the marginal distributions of financial data. Another advantage of their generalization is the fact that their transformation does not restrain the smoothness of the density function obtained, which is useful for quasi-maximum likelihood optimization problem. Here is the proposed p.d.f. of Skew t-Distribution for univariate case:

$$f_Y(y) = 2\, t_\nu(y)\, T_{\nu+1}\left(\gamma \frac{y-\xi}{\omega}\left(\frac{\nu+1}{\nu+Q_y}\right)^{1/2}\right),$$

where

$$\begin{aligned} Q_y &= \left(\frac{y-\xi}{\omega}\right)^2, \\ t_\nu(y) &= \frac{\Gamma((\nu+1)/2)}{\omega\,(\pi\nu)^{1/2}\Gamma(\nu/2)}\,(1+Q_y/\nu)^{-(\nu+1)/2}, \end{aligned}$$

and $T_{\nu+1}(x)$ denotes the c.d.f. of standard t-distribution with $\nu+1$ degrees of freedom. The parameter γ reflects the skewness of the distribution. Denote later

$$Y \sim \mathrm{St}_1(\xi, \omega, \gamma, \nu).$$

It is worth displaying here the first three moments of the distribution when $\xi = 0$ (Azzalini and Capitanio, 2003):

$$\begin{aligned} \overline{\mu} &:= \frac{\gamma}{\sqrt{1+\gamma^2}}\left(\frac{\nu}{\pi}\right)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}, \\ E(Y) &= \omega\,\overline{\mu}, && (5.1) \\ E(Y^2) &= \overline{\sigma^2} = \omega^2\frac{\nu}{\nu-2}, && (5.2) \\ E(Y^3) &= \overline{\lambda} = \omega^3\overline{\mu}\,\frac{3+2\gamma^2}{1+\gamma^2}\,\frac{\nu}{\nu-3}. \end{aligned}$$

The last equation indicates that varying γ one can vary skewness of the distribution.

One can see, however, that the first moment of Y is different from 0 and its second central moment is not equal to 1. It is useful to define the standardized Skew

t-Distribution by adjusting for zero expectation and unit variance through setting ξ and ω in the following way:

$$\omega = \left(\frac{\nu}{\nu-2} - \overline{\mu}^2\right)^{-1/2},$$
$$\xi = -\omega\,\overline{\mu}.$$

Denote below the standardized Skew t-Distribution as $\mathrm{St}\,(\gamma, \nu)$. Additionally, assume now that the following data sample of log-returns is available: $\{\mathbf{y}_t = \{y_{it}\}_{i=1}^{p}\}_{t=1}^{T}$, where y_{it} is the individual log-return of the ith stock at the time moment t, p is the total number of stocks, and T is the length of the data sample.

5.2 Marginal distributions

The standard NAGARCH structure is imposed and the following model for marginal distributions is proposed:

$$\begin{aligned} y_{it} &= \mu_i + \sqrt{h_{it}}\varepsilon_{it}, \qquad \varepsilon_{it} \sim i.i.d.\,\mathrm{St}(\gamma_i, \nu_i), \\ h_{it} &= \omega_i + \alpha_i\left(\{y_{i,t-1} - \mu_i\} + \kappa_i\sqrt{h_{i,t-1}}\right)^2 + \beta_i h_{i,t-1}, \end{aligned}$$

where i denotes the number of an asset, y_{it}'s are log-returns of the asset, h_{it}'s are the conditional variances of y_{it}'s, and $(\mu_i, \gamma_i, \nu_i, \omega_i, \alpha_i, \beta_i, \kappa_i)$ is the set of parameters. It is worth noting here that the parameter κ_i reflects the leverage effect and is expected to be negative.

Using this structure one can derive the conditional distribution of y_{it}:

$$F_i\left(y_{it}|h_{it}\right) = F^{\mathrm{St}}_{\gamma_i,\nu_i}\left(\frac{y_{it}-\mu_i}{h_{it}^{1/2}}\right),$$

where $F^{\mathrm{St}}_{\gamma_i,\nu_i}$ denotes the distribution function of the standardized Skew t-Distribution $\mathrm{St}\,(\gamma_i, \nu_i)$.

The log-likelihood function for estimation of each of the marginal distributions will have the following form:

$$\ln L_i = \sum_{t=2}^{T} \left\{ \ln f^{\mathrm{St}}_{\gamma_i,\nu_i} \left(\frac{y_{it} - \mu_i}{h_{it}^{1/2}} \right) - \frac{1}{2} \ln h_{it} \right\}.$$

Note, that it is assumed that the univariate data follows the process defined in this subsection and the goodness-of-fit tests below do not reject the hypothesis that the data do follow this process.

5.3 The t-copula

The model for copula is modified from the basic one of Ausin and Lopes (2010) and is assumed to be the following one:

$$C_{\eta,R}\left(u_1,\ldots,u_p\right) = \int_{-\infty}^{T_\eta^{-1}(u_1)} \cdots \int_{-\infty}^{T_\eta^{-1}(u_p)} \frac{\Gamma\left(\frac{\eta+p}{2}\right)\left(1+\frac{\mathbf{v}'R^{-1}\mathbf{v}}{\eta-2}\right)^{-\frac{\eta+p}{2}}}{\Gamma\left(\frac{\eta}{2}\right)\sqrt{(\pi(\eta-2))^p\left|R\right|}} d\mathbf{v},$$

where $T_\eta^{-1}(\cdot)$ is the inverse of the c.d.f. of the standardized Student's t-Distribution, η is its degrees of freedom and R is correlation matrix. Denote the expression under the integral as $f_{\eta,R}(\mathbf{u})$. It is the p.d.f. of the standardized multivariate Student's t-Distribution.

Following Ausin and Lopes (2010) the dynamics of the correlation matrix R is assumed to be the following one:

$$R_t = (1-a-b)\overline{R} + a\Psi_{t-1} + bR_{t-1},$$

where $a \geq 0$, $b \geq 0$, $a+b \leq 1$, $\overline{R}$ is positive definite constant matrix with unit diagonal, and Ψ_{t-1} is a matrix with the following elements:

$$\begin{aligned} \Psi_{ij,t-1} &= \frac{\sum_{h=1}^{m} x_{it-h}x_{jt-h}}{\sqrt{\sum_{h=1}^{m} x_{it-h}^2 \sum_{h=1}^{m} x_{jt-h}^2}}, \\ x_{it} &= T_\eta^{-1}\left(F^{\mathrm{St}}_{\gamma_i,\nu_i}\left(\frac{y_{it}-\mu_i}{h_{it}^{1/2}} \right) \right). \end{aligned}$$

The advantage of such defined R_t lies in the fact that this form of matrix is positive definite and a well-defined correlation matrix, so no additional transformations (like

logistic transformation in the work of Patton, 2006) are needed to ensure that.

Substituting the marginal distributions to the assumed copula function the model for the joint c.d.f. of a vector of financial log-returns $\mathbf{y}_t = (y_{1t}, \ldots, y_{pt})$ will take the following form:

$$F\left(\mathbf{y}_t|\mathbf{h_t}\right) = C_{\eta,R_t}\left(F_1\left(y_{1t}|h_{1t}\right), \ldots, F_p\left(y_{pt}|h_{pt}\right)\right). \tag{5.3}$$

In order to find the log-likelihood function one first need to find the joint p.d.f. by differentiating the above equation of the joint c.d.f. (5.3):

$$\begin{aligned} f\left(\mathbf{y}_t|\mathbf{h}_t\right) &= f_{\eta,R_t}\left(T_\eta^{-1}\left(F_1\left(y_{1t}|h_{1t}\right)\right), \ldots, T_\eta^{-1}\left(F_p\left(y_{pt}|h_{pt}\right)\right)\right) \\ &\quad \prod_{i=1}^{p}\left\{\frac{1}{t_\eta\left(T_\eta^{-1}\left(F_i\left(y_{it}|h_{it}\right)\right)\right)} f_{\gamma_i,\nu_i}^{\mathrm{St}}\left(\frac{y_{it}-\mu_i}{h_{it}^{1/2}}\right)\frac{1}{h_{it}^{1/2}}\right\}. \end{aligned}$$

Hence, the log-likelihood function for the estimation procedure could be written as the following one:

$$\begin{aligned} \ln L &= \sum_{t=m+1}^{T} \ln f_{\eta,R_t}\left(T_\eta^{-1}\left(F_1\left(y_{1t}|h_{1t}\right)\right), \ldots, T_\eta^{-1}\left(F_p\left(y_{pt}|h_{pt}\right)\right)\right) \\ &\quad + \sum_{t=m+1}^{T}\sum_{i=1}^{p}\left\{-\ln t_\eta\left(T_\eta^{-1}\left(F_i\left(y_{it}|h_{it}\right)\right)\right) + \ln f_{\gamma_i,\nu_i}^{\mathrm{St}}\left(\frac{y_{it}-\mu_i}{h_{it}^{1/2}}\right) - \frac{1}{2}\ln h_{it}\right\}. \end{aligned}$$

This paper applies the sequential approach and the marginal distributions are estimated on the first step (note: this may lead to inefficiency of the estimates). Hence, the last two summands in the last sum of the log-likelihood written above do not depend on the copula parameters, and the log-likelihoods for pairwise copulas can be written in the following form:

$$\begin{aligned} \ln L_{ij} &= \sum_{t=m+1}^{T} \ln f_{\eta,R_t}\left(T_\eta^{-1}\left(\widehat{F}_i\left(y_{it}|h_{it}\right)\right), T_\eta^{-1}\left(\widehat{F}_j\left(y_{jt}|h_{jt}\right)\right)\right) \\ &\quad - \sum_{t=m+1}^{T}\left\{\ln t_\eta\left(T_\eta^{-1}\left(\widehat{F}_i\left(y_{it}|h_{it}\right)\right)\right) + \ln t_\eta\left(T_\eta^{-1}\left(\widehat{F}_j\left(y_{jt}|h_{jt}\right)\right)\right)\right\}. \end{aligned}$$

5.4 Models for compounding functions

Again, the basic idea of compounding functions is to capture the dependence between the marginal distribution of log-returns of a single asset and the joint distribution of a group of assets. For this purpose the author considers modeling them as asymmetrized copulas. One of the reasons is described below.

5.4.1 Asymmetrized copulas

When one estimates bivariate distributions using copulas, he or she deals with two a priori similar objects of the same nature: namely, marginal distributions of two assets under consideration. Hence, it might make sense to use symmetric[1] copulas for distribution modeling in that case. In contrast, compounding functions operate with two objects of similar, but different nature: namely, a marginal distribution of an asset and a joint distribution of a group of assets. Thus, it seems to be better to use asymmetric copulas rather than symmetric ones as compounding functions.

Khoudraji (1995) in his Ph.D. thesis proposes a theorem that allows constructing asymmetric bivariate copulas from any symmetric one:

Theorem 1. *Any symmetric bivariate copula $C^{(sym)}(u,v)$ can be transformed to the asymmetric bivariate copula $C^{(asym)}(u,v)$ by the following transformation:*

$$C^{(asym)}(u,v) = u^{\alpha} v^{\beta} C^{(sym)}\left(u^{1-\alpha}, v^{1-\beta}\right), \quad 0 \leq \alpha, \beta \leq 1.$$

This theorem is used below in order to build asymmetric copulas from t-copula and Clayton-copula. Then, these asymmetric copulas are used for constructing compounding functions for the new sequential approach.

5.4.2 Asymmetrized time-varying t-copula

Symmetric standardized t-copula has already been described in the Subsection 5.3. Applying Theorem 1, one can obtain the functional form of the asymmetrized bivariate standardized t-copula:

[1]Here and further symmetric means that $C(X,Y) = C(Y,X)$ as functions and asymmetric means that $C(X,Y) \neq C(Y,X)$ as functions. Do not confuse with asymmetric tail dependence, like in Clayton-copula.

$$C_{\eta,\rho}^{(t,asym)}(u,v) = u^{\alpha}\, v^{\beta} \int\limits_{-\infty}^{T_{\eta}^{-1}\left(u^{1-\alpha}\right)} \int\limits_{-\infty}^{T_{\eta}^{-1}\left(v^{1-\beta}\right)} \frac{\Gamma\left(\frac{\eta+2}{2}\right)\left(1+\frac{x^2+y^2-2\,\rho\,x\,y}{(\eta-2)(1-\rho^2)}\right)^{-\frac{\eta+2}{2}}}{\Gamma\left(\frac{\eta}{2}\right)\pi\,(\eta-2)\sqrt{1-\rho^2}} dx\, dy,$$

where u denotes the marginal distribution of an asset, v is the distribution of a group of assets, and similar time-varying structure on the correlation coefficient as in the Subsection 5.3 is applied. The form of the compounding function on the m-th step in this case will be the following one: $\widetilde{C}^{(m)}\left(\widehat{F}_l, \widehat{F}_{i_1,\ldots,l-1,l+1,\ldots,i_m}; \theta_{l,i_1,\ldots,l-1,l+1,\ldots,i_m}\right) = C_{\eta,\rho}^{(t,asym)}\left(\widehat{F}_l, \widehat{F}_{i_1,\ldots,l-1,l+1,\ldots,i_m}\right)$, where $\theta_{l,i_1,\ldots,l-1,l+1,\ldots,i_m} = \{\alpha, \beta, \eta, \overline{\rho}, a, b\}$ is the set of parameters (the last three parameters come from the time-varying structure of the correlation matrix R, that is reduced to the correlation coefficient ρ in bivariate case). One can see, that in the case of asymmetrized time-varying t-copula there are only six parameters to estimate in each optimization problem on each step of the new sequential approach regardless of the dimension of the whole problem.

5.4.3 Asymmetrized time-varying Clayton-copula

Similarly, using Theorem 1, one can obtain the form of asymmetrized bivariate Clayton-copula. It will have the following form:

$$C^{(Clayton,asym)}(u,v) = \begin{cases} u^{\alpha} v^{\beta}\left(\max\left\{u^{-\theta(1-\alpha)} + v^{-\theta(1-\alpha)} - 1, 0\right\}\right)^{-1/\theta}, & \theta \in [-1; +\infty] \setminus \{0\} \\ u\, v, & \theta = 0 \end{cases},$$

where parameter θ can be rewritten through Kendall's τ: $\theta = \frac{2\tau}{1-\tau}$, and similar to the t-copula's correlation parameter (except the facts that x's are defined as just u's and v's and instead of Pearson correlation Kendall's τ is used for computing Ψ's) time-varying structure on the Kendall's τ is imposed and θ is recomputed through it. In this case, the compounding function on the m-th step will have the following form: $\widetilde{C}^{(m)}\left(\widehat{F}_l, \widehat{F}_{i_1,\ldots,l-1,l+1,\ldots,i_m}; \theta_{l,i_1,\ldots,l-1,l+1,\ldots,i_m}\right) = C^{(Clayton,asym)}\left(\widehat{F}_l, \widehat{F}_{i_1,\ldots,l-1,l+1,\ldots,i_m}\right)$, where $\theta_{l,i_1,\ldots,l-1,l+1,\ldots,i_m} = \{\alpha, \beta, \overline{\tau}, a, b\}$ is the set of parameters (the last three parameters come from the time-varying structure of τ). One can see, that in the case of asymmetrized

time-varying Clayton-copula there are just five parameters to estimate in each optimization problem on each step regardless of the dimension of the whole joint distribution estimation problem.

5.5 Aggregating functions

As it has been already stated, one can consider different non-parametric and even parametric aggregators. The basic idea of them is to aggregate the parts of the dependencies that are caught by the compounding functions in order to obtain the whole dependence of all assets, or alternatively, integrate out the error that is introduced by each individual compounding function. The author has considered two non-parametric aggregating functions: geometric mean and arithmetic mean. As the intuition suggests, arithmetic mean is proved to be much more suitable according to the conducted goodness-of-fit tests for the joint distributions, and thus, in this paper the results only for arithmetic mean aggregating function are included (for the results see plots in the Appendix).

The next section demonstrates how to apply the new sequential method proposed in this section for modeling five-dimensional distribution of real financial data and compares it to the standard single-copula approach.

6 Empirical evidence and estimation

6.1 Data

Five DJIA-constituents (as of June 8, 2009) are chosen to conduct the empirical tests. The selection is based on the high level of liquidity and the availability of historical prices. In particular, GE – General Electric Co, MCD – McDonald's Corp, MSFT – Microsoft Corp, KO – Coca-Cola Co, and PG – Procter & Gamble Co stocks are used to estimate the models. Daily data for 1 year (from Jan 03, 2007 to Dec 31, 2007) were used for estimation to exclude the period of the recent financial crisis. The prices of the stocks are adjusted for splits and dividends, and then the log-returns are constructed and used in the models. The relative prices dynamics plots, histograms of the log-returns, and sample statistics of the log-returns for all five stocks are presented below.

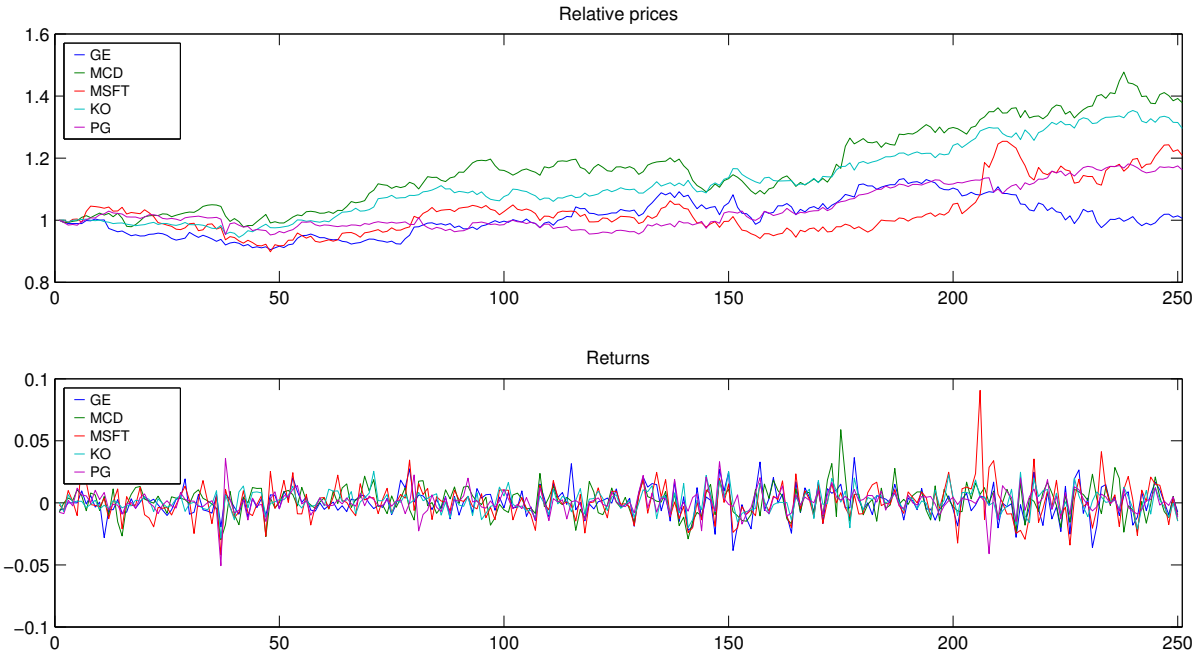

Figure 6.1: Relative prices and returns dynamics for GE, MCD, MSFT, KO, and PG on NYSE from Jan 03, 2007 to Dec 31, 2007

	GE	MCD	MSFT	KO	PG
minimum	-0.0384	-0.0298	-0.0421	-0.0285	-0.0506
maximum	0.0364	0.0589	0.0907	0.0254	0.0359
mean, $\times 10^{-3}$	0.0248	1.2831	0.7575	1.0363	0.5987
standard deviation	0.0115	0.0116	0.0143	0.0087	0.0091
skewness	-0.0349	0.2617	0.9461	0.0512	-0.6106
(zero-skewness: p-val)	(0.8216)	(0.0912)	(0.0000)	(0.7408)	(0.0001)
kurtosis	3.9742	4.8977	8.7270	3.6313	9.2954
(zero-ex.kurtosis: p-val)	(0.0017)	(0.0000)	(0.0000)	(0.0416)	(0.0000)

Table 1: Sample statistics on returns for GE, MCD, MSFT, KO, and PG on NYSE from Jan 03, 2007 to Dec 31, 2007

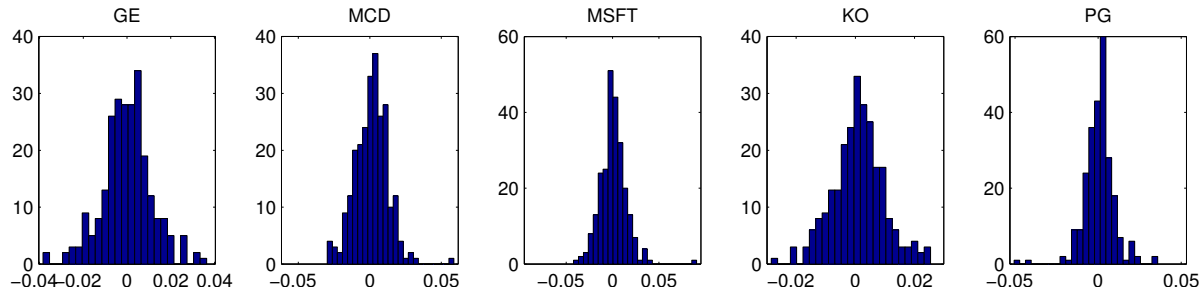

Figure 6.2: Histograms of GE, MCD, MSFT, KO, and PG returns on NYSE from Jan 03, 2007 to Dec 31, 2007

One can see that the unconditional sample distributions in some cases demonstrate skewness and heavy tails (the corresponding hypotheses about zero skewness and zero excess kurtosis can be rejected on the 95% confidence level). This at least partially justifies the selection of the Skew t-Distribution for modeling marginals.

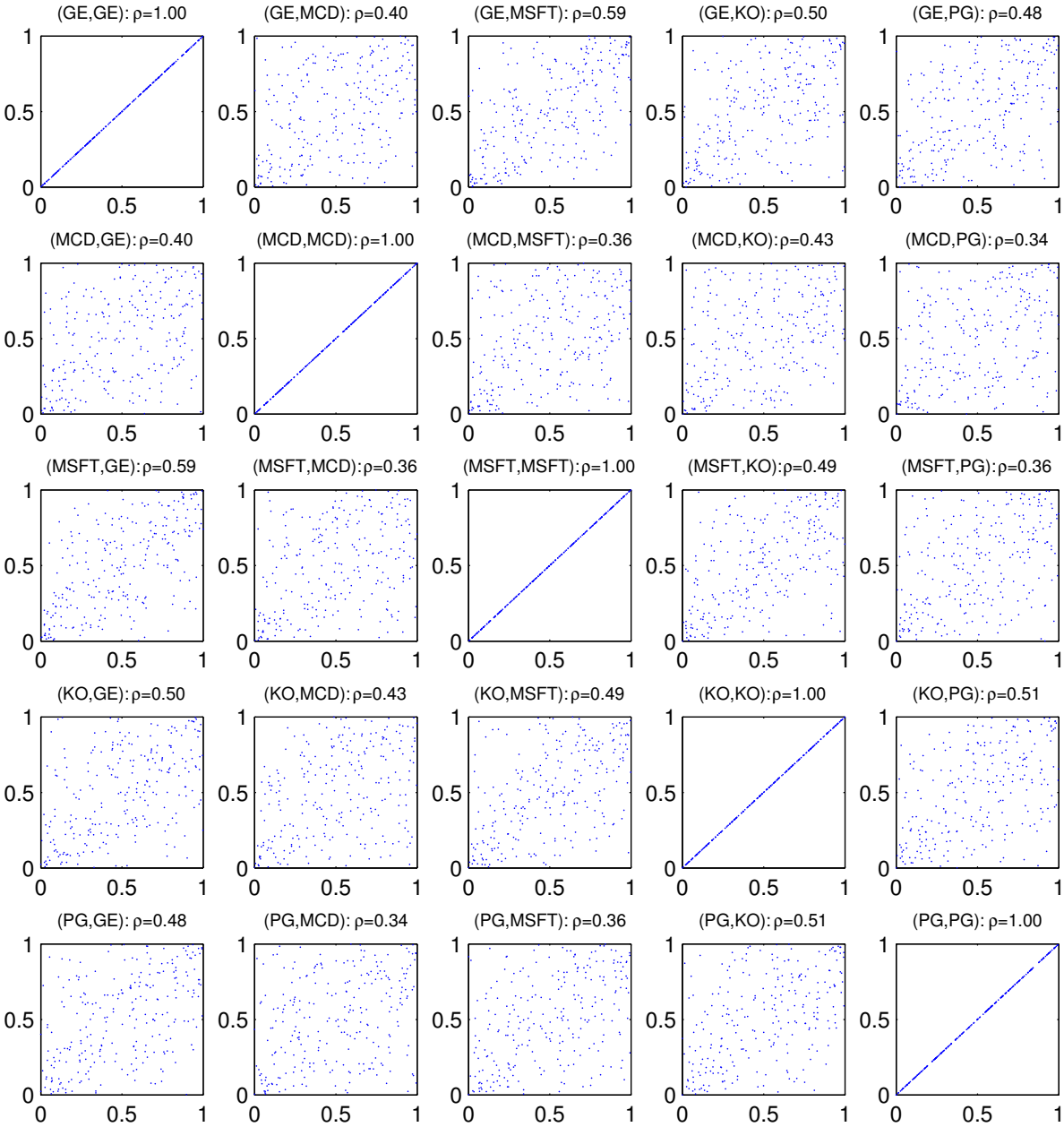

Figure 6.3: Pairwise scatter plots of the marginal distributions along with sample correlations of GE, MCD, MSFT, KO, and PG returns on NYSE from Jan 03, 2007 to Dec 31, 2007

The scatter plots and correlations show that as expected all of the stocks are positively correlated (because they are partially driven by the same economic environment). The pair of MSFT and PG, for example, has correlation lower than most of others because these stocks are from the completely different sectors of the economy: the first one is from the Technology sector and the second one is from the Consumer Goods sector (classification by http://finance.yahoo.com). At the same time, the pair KO and PG has greater correlation than others due to the fact that they are both from the Consumer Goods sector.

The estimation procedure is conducted sequentially: first, estimation of the marginal distributions; then, estimation of the bivariate copula parameters; and finally, stepwise estimation of compounding functions and aggregation of them into the distributions of interest. The author is aware about the inefficiencies that may arise in the procedure, however such approach allows the estimation for large dimensions to be computationally feasible.

6.2 Marginal distributions estimation

As it has been described in the Section 5, Skew-t-NAGARCH model is used for the marginals in order to take into account the asymmetry, heavy tails, the leverage effect, and volatility clustering that are all usually observed in the financial data, especially in the log-returns of stocks. For marginal distributions estimation the initial value of the conditional variance to start the GARCH process, h_{i1}, is selected to be the sample unconditional variance of log-returns, that is $h_{i1} = \mathrm{V}(y_{it})$. In addition, several constraints on the parameters are imposed in order to guarantee the stationarity of the processes for returns and conditional variances. The estimates of the parameters of marginal distributions are summarized in the table below. It is worth noting the economic meaning of the parameters presented here: μ accounts for mean return, ω – for unconditional variance, α – for ability to predict the conditional variance by current innovation, β – for persistence of the conditional variance, κ – for the leverage effect, ν – for heavy tails, and γ – for skewness.

	GE	MCD	MSFT	KO	PG
μ, $\times 10^{-3}$	-0.032	1.340	0.660	0.750	0.574
	(0.615)	(0.709)	(0.886)	(0.673)	(0.645)
ω, $\times 10^{-5}$	0.569	5.845	0.852	0.667	0.608
	(0.581)	(1.893)	(0.809)	(0.740)	(0.523)
α	0.106	0.153	0.041	0.142	0.107
	(0.062)	(0.090)	(0.024)	(0.082)	(0.052)
β	0.861	0.379	0.915	0.787	0.837
	(0.084)	(0.130)	(0.055)	(0.139)	(0.063)
κ	-0.074	-0.530	-0.174	-0.032	-0.168
	(0.364)	(0.394)	(0.796)	(0.740)	(0.606)
ν	6.482	9.672	5.898	8.098	3.305
	(2.325)	(4.976)	(2.360)	(4.046)	(0.734)
γ	-0.014	-0.431	0.176	-0.236	-0.106
	(0.077)	(0.706)	(0.533)	(0.950)	(0.482)

Table 2: Parameters estimates for marginal distribution of GE, MCD, MSFT, KO, and PG returns on NYSE from Jan 03, 2007 to Dec 31, 2007. Robust standard errors are in the round brackets

It is useful to compare the values of μ parameter in the model for marginal distributions with the sample means of the log-returns of the financial data under consideration: one can see that they are rather close to each other. This indicates that our model is rather good. The ν parameters reflect the extent to which the tails are fat and also correspond to the histograms and kurtosis values given above (note: the parameter ν is inversely related to the kurtosis). The γ parameters reflect the skewness of the distributions, and one can note that this parameter is insignificant for all stocks. However, the hypothesis of zero-skewness is rejected for MSFT and PG stocks on 95% confidence level (see Table 1 on page 21) and one can see that the signs of the γ parameter for the log-returns of these stocks correspond to the signs of their sample skewness. Thus, the author does not exclude this parameter from consideration. Moreover, the proposed sequential procedure is intended to be universal, and for example, these parameters are significant for most Russian stocks that were considered during the research. See below the goodness-of-fit tests for marginal distributions to determine how well the proposed model describes the data.

6.2.1 Goodness-of-fit tests for marginal distributions

The goodness-of-fit tests need to be conducted in order to assess the validity of the Skew-t-NAGARCH specification that was chosen for modeling the marginal distributions. For this purpose the following approach based on the paper of Diebold et al. (1998) is used. The approach is based on transforming the time series of log-returns into the new series that should have the known pivotal distribution in the case of correct specification and then testing the hypothesis that the transformed series indeed have that known distribution.

They use the following proposition:

Proposition 1. *Suppose a sequence* $\{y_t\}_{t=1}^T$ *is generated from the distributions* $\{F_t(y_t|\Omega_t)\}_{t=1}^T$, *where* $\Omega_t = \{y_{t-1}, y_{t-2}, \ldots\}$. *Then under the usual condition of a non-zero Jacobian with continuous partial derivatives, the sequence of transformations* $\{F_t(y_t|\Omega_t)\}_{t=1}^T$ *is i.i.d.* $U(0,1)$.

They propose testing the uniformity property and the independence property separately by investigating the histogram and correlograms of the moments up to degree of 4. This paper follows this approach, however the statistical tests rather than graphical analysis are conducted in order to test the uniformity and independence properties separately, as they suggest.

In order to test the uniformity of the transformed series Kolmogorov-Smirnov tests for uniformity are conducted. The author is aware about the parameter estimation effect that arises in such kind of tests (when firstly, the parameters of the distribution functions are estimated, and only then the transformed series is constructed) and that the tests require adjustments due to it, however, in practice almost everybody neglects this effect while conducting such kind of tests, and so does the author. The conducted tests show that on 95% confidence level the hypotheses about uniformity of the transformed series is not rejected. The quantitative results along with the diagrams are presented below. The model passes this part of the tests successfully, and next, the independence property of the transformed series should be tested.

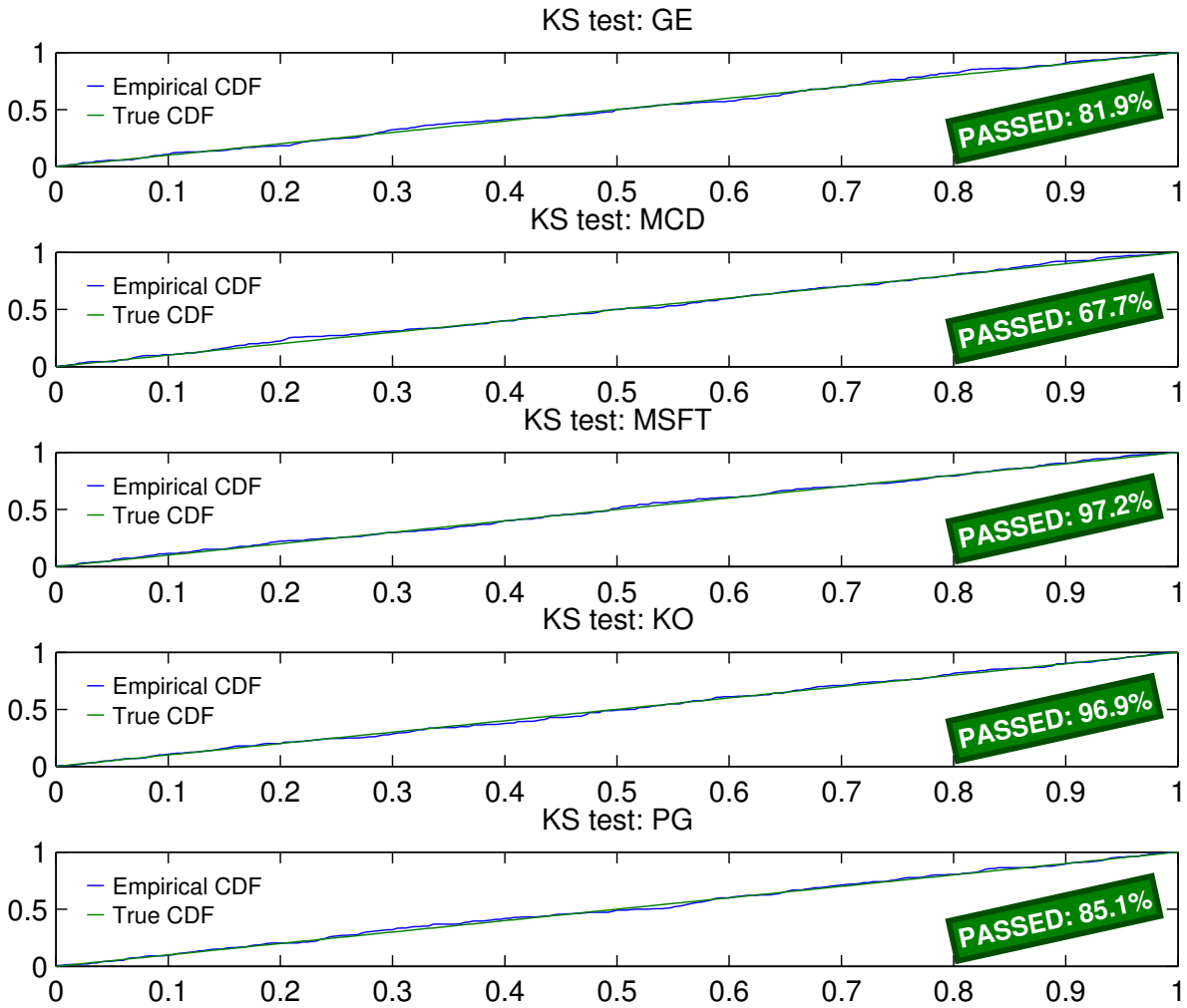

Figure 6.4: Kolmogorov-Smirnov tests for the marginals' GoF tests (p-value is presented after the "PASSED" word)

Next, the tests for serial correlation are conducted. Diebold et al. (1998) claim that on practice it is sufficient to investigate the moments up to degree of 4. This paper follows their proposition and tests the hypothesis about the joint insignificance of the coefficients of the regression of each moment on its 20 lags using the F-test. The results are presented in the table below.

degree of the central moment	GE	MCD	MSFT	KO	PG
1	0.701	0.454	0.762	0.336	0.310
2	0.763	0.805	0.448	0.070	0.437
3	0.567	0.672	0.763	0.611	0.657
4	0.887	0.774	0.635	0.032	0.172

Table 3: P-values of the F-tests for serial correlation for the marginals' GoF tests

The hypotheses about the absence of serial correlation are not rejected almost everywhere on 95% confidence level. The exception is the 4th central moment of the KO stock: the hypothesis is not rejected there on 99% confidence level. This is considered to be satisfactory for the purposes of this paper.

Additionally, the Ljung-Box tests are carried out to test for autocorrelation in the residuals of the marginals' specification. All of the tests do not reject the hypothesis of the absence of serial correlation in the residuals of the models considered. This also indicates good fit of the marginal distributions.

6.3 Pairwise copula estimation

The estimates of the parameters of the pairwise copulas are summarized in the following table:

	GE,MCD	GE,MSFT	GE,KO	GE,PG	MCD,MSFT
η	9.627	7.948	6.107	14.236	9.883
	(7.732)	(3.006)	(2.390)	(11.290)	(8.488)
a	0.074	0.089	0.002	0.038	0.159
	(0.075)	(0.113)	(0.002)	(0.023)	(0.096)
b	0.399	0.001	0.486	0.913	0.385
	(0.241)	(0.130)	(0.306)	(0.031)	(0.226)
$\bar{\rho}$	0.418	0.625	0.513	0.557	0.429
	(0.062)	(0.042)	(0.050)	(0.076)	(0.073)

	MCD,KO	MCD,PG	MSFT,KO	MSFT,PG	KO,PG
η	11.825	6.011	5.672	6.926	10.760
	(9.397)	(2.476)	(2.226)	(2.968)	(9.417)
a	0.072	0.170	0.031	0.197	0.053
	(0.087)	(0.091)	(0.225)	(0.152)	(0.076)
b	0.447	0.394	0.462	0.000	0.342
	(0.288)	(0.266)	(2.057)	(0.169)	(0.209)
$\bar{\rho}$	0.417	0.368	0.556	0.469	0.504
	(0.059)	(0.080)	(0.056)	(0.065)	(0.050)

Table 4: Parameters estimates for the pairwise copulas of GE, MCD, MSFT, KO, and PG returns on NYSE from Jan 03, 2007 to Dec 31, 2007. Robust standard errors are in the round brackets

One can note that the parameter $\bar{\rho}$ reflects the historical correlation between the stocks' log-returns, and so it is expected their values to be close to the correlations computed from the data sample under consideration. Indeed, if one looks on the Figure 6.3 on page 22 where these values are shown, he or she can verify that the sample correlations are rather close to the $\bar{\rho}$ parameters estimated in the pairwise copulas models. This partially indicates the adequacy of the model applied for modeling pairwise distributions of the stocks considered.

6.3.1 Goodness-of-fit tests for bivariate copulas

The goodness-of-fit tests in this section are based on the approach proposed by Breymann et al. (2003). The approach relies on the definition of Probability Integral Transform that was first given by Rosenblatt (1952):

Definition 1 (Probability Integral Transform (PIT))**.** *Let* $\mathbf{X} = (X_1, \ldots, X_d)$ *denote a random vector with marginal distributions* $F_i(x_i)$ *and conditional distributions* $F_{i|i-1\ldots1}(x_i|x_{i-1}, \ldots, x_1)$ *for* $i = 1, \ldots, d$*. The PIT of* $\mathbf{x}$ *is defined as* $\mathbf{z} = T(\mathbf{x}) = T(x_1, \ldots, x_d) = (T_1, \ldots, T_d)$*:*

$$\begin{aligned} T_1 &= F_1(x_1), \\ T_p &= F_{p|p-1\ldots1}(x_p|x_{p-1}, \ldots, x_1), \quad p = 2, \ldots, d. \end{aligned}$$

One can show that $T(\mathbf{X})$ is uniformly distributed on the p-dimensional hyper-cube (for details, see Rosenblatt, 1952). This implies that $Z_1, \ldots, Z_p$ are uniformly and independently distributed on $[0, 1]$. Similar to the Section 6.2.1, this approach can be extended to the time series case (see for example, Patton, 2006). The same approach for testing as in that section is chosen, that is Kolmogorov-Smirnov tests for uniformity and F-tests for serial uncorrelation are conducted. However, here one can note that there exist $p!$ ways of choosing conditional distributions. For example, in the case of pairwise copulas there exist two ways of writing conditional distributions: $X_2 \mid X_1$ and $X_1 \mid X_2$. This paper examines them for all pairs of stocks considered. The results are presented in the two figures below (the marginal distributions are not included because they have already been examined in the Section 6.2.1).

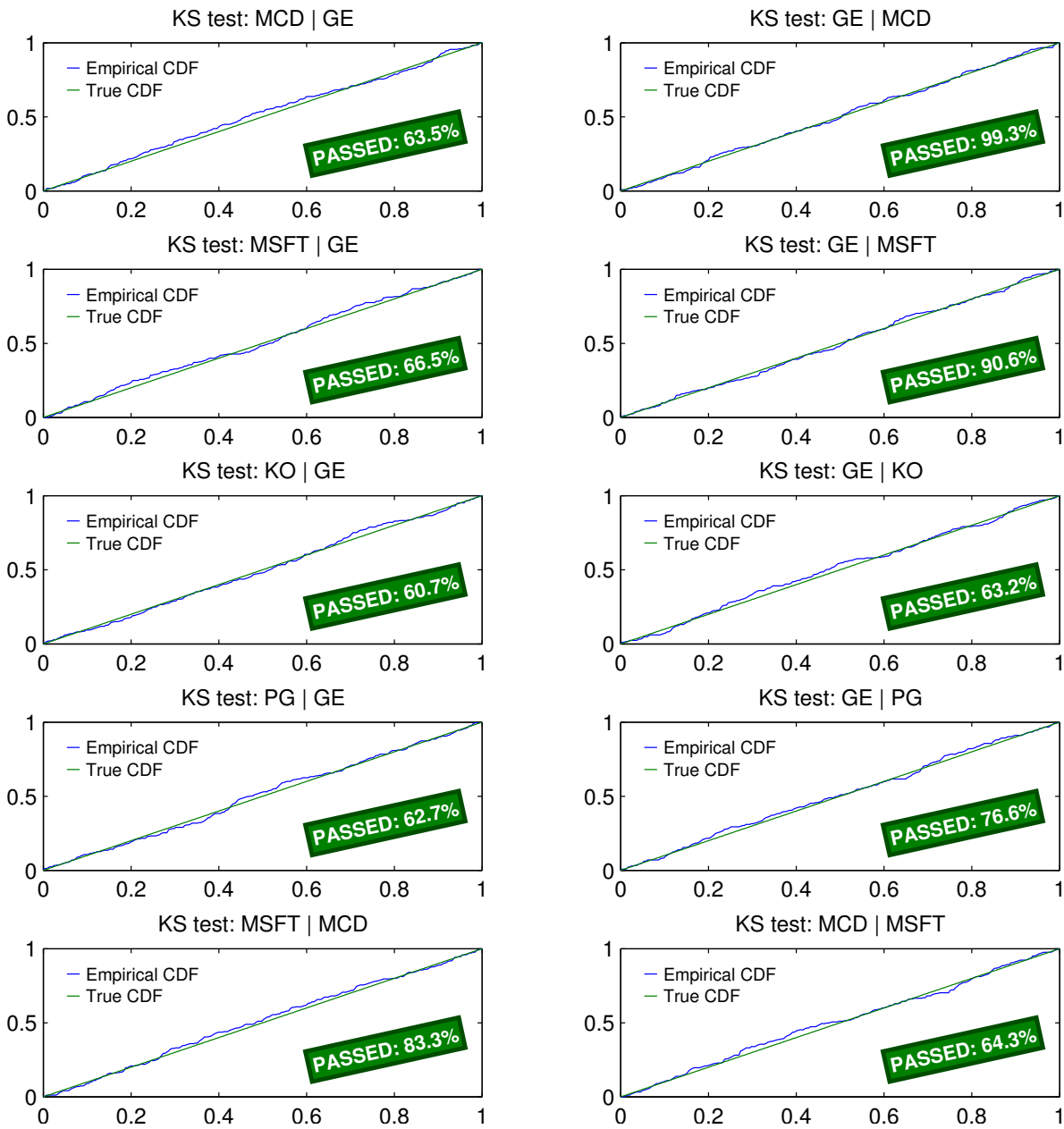

Figure 6.5: Pairwise Kolmogorov-Smirnov tests for conditional distributions for the first five pairs of GE, MCD, MSFT, KO, and PG – GoF tests for the bivariate copula specification (p-values are presented after the "PASSED" word, see overleaf for other pairs)

One can see that all Kolmogorov-Smirnov tests are passed on any reasonable confidence level. This indicates that the time-varying t-copula used for modeling bivariate distributions is good enough and can be selected as a second basic block for the new sequential approach. See also the figure overleaf for the tests for the complete set of pairs.

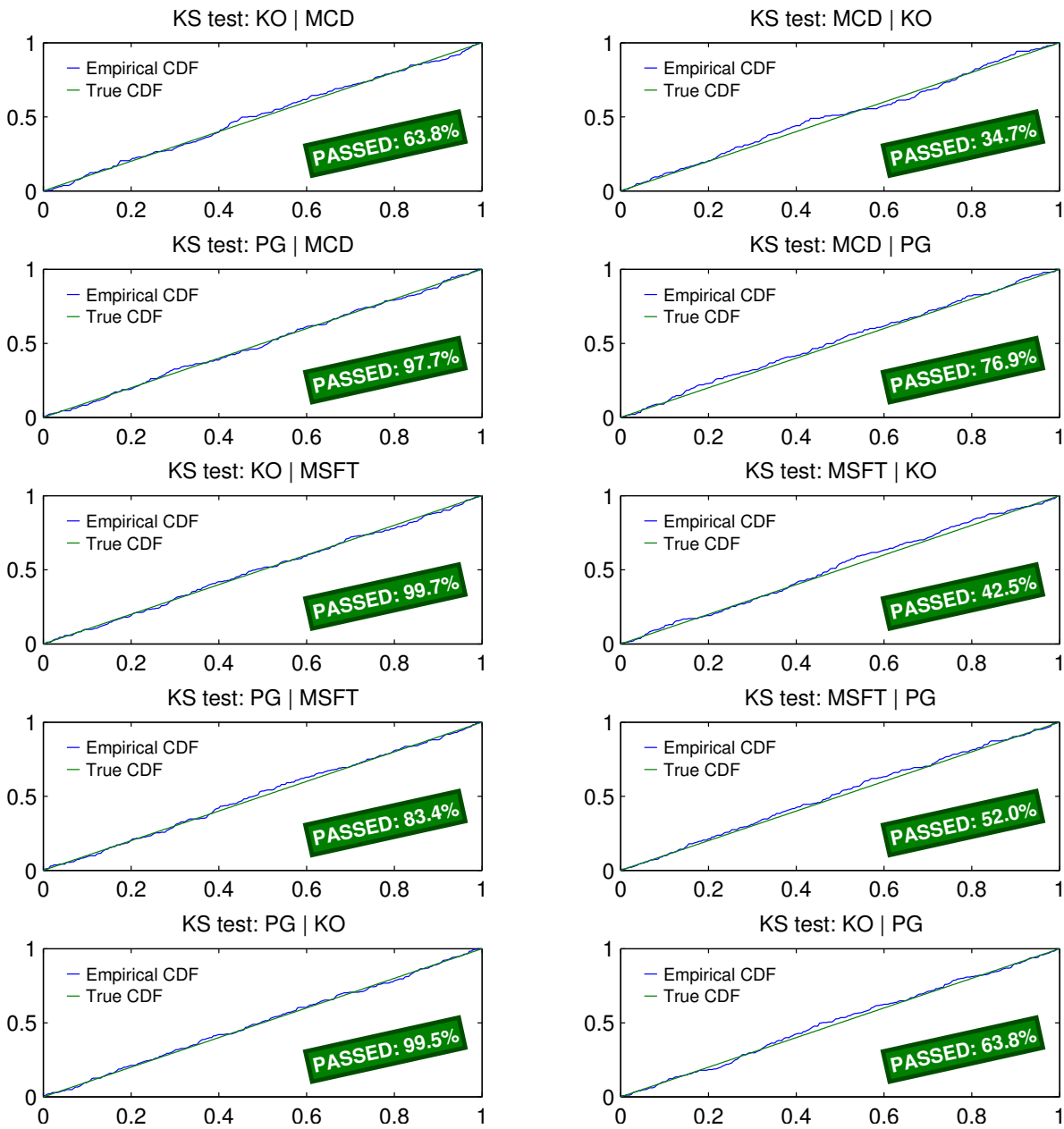

Figure 6.6: Pairwise Kolmogorov-Smirnov tests for conditional distributions for the second five pairs of GE, MCD, MSFT, KO, and PG – GoF tests for the bivariate copula specification (p-values are presented after the "PASSED" word)

As it was proposed above in the Section 6.2.1, the hypothesis about the absence of serial correlation is tested through testing the hypothesis about the joint insignificance of the coefficients of the regression of each of the 4 moments on its 20 lags using the F-test. Additionally to that method, here one needs to include the lags of the moments of other series of PIT in order to test for the independence of each of the PIT series. The results obtained suggest that the hypotheses about the absence of serial correlation

are not rejected on any reasonable confidence level almost in every case. The table with the results is not presented in this paper because of its huge size.

The results of this subsection indicate that the specification of the pairwise copulas is rather good and can be used in the next steps of the new sequential approach.

6.4 Compounding functions estimation

This paper proposes the following approach for estimating the parameters of the compounding functions. The equation (4.1) is interpreted as mean of the joint distribution estimators that is taken in order to cancel out the individual errors in each estimate. Hence, it might make sense to estimate each of the summands separately by, for example, Maximum Likelihood Estimator. Thus, in each optimization problem there are only fixed number of parameters (it is 5 for Clayton-based and 6 for t-based compounding functions as shown in the Section 5.4), because each compounding function operates only with two objects: a marginal distribution of an asset and a joint distribution of a group of assets. This makes it feasible the estimation of the distributions for any dimension.

In order to use MLE for estimating parameters of the compounding functions one needs to differentiate the function $\widetilde{C}^{(m)}(\cdot,\cdot)$ m times. There are two ways of differentiating: symbolic and numeric. The first one leads to huge formulas (because such a function is highly nested), and hence, may require great computations. The second one may require computing the $\widetilde{C}^{(m)}(\cdot,\cdot)$ function with a good accuracy, because numeric differentiating requires finding finite differences that are usually small, and hence, the function itself should be computed rather accurately. This paper adopts the second approach because the first approach requires more powerful computational resources.

Nevertheless, it seems that one could use even more computationally efficient estimation way: Maximum Spacings Estimator (MSPE, for more details on univariate case see among others Cheng & Amin, 1983; Ranneby, 1984; and Anatolyev & Kosenok, 2005).

6.4.1 Maximum Spacings Estimator (MSPE)

This subsection describes MSP estimator for the i.i.d. univariate case. The basic idea of the MLE lies in maximizing the product of the values of the density function in

the points of data (the red bars on the 6.7) on the parameters of that density. When number of observation grows, the lengths of red bars on the 6.7 tend to be equal to the corresponding hatched areas (surely, divided by the base of an area). Thus, if one defines the product of spacings as the product of hatched areas, when asymptotically the problem of likelihood maximization should be equivalent to the problem of product of spacings maximization. Anatolyev and Kosenok (2005) show, that in fact, in the univariate case the two problems give asymptotically first-order equivalent estimates of the distribution parameters.

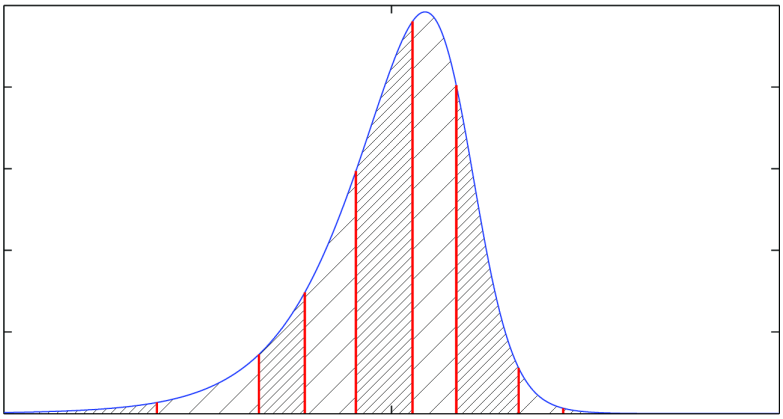

Figure 6.7: Maximum Spacings Estimator illustration

Note, that MSPE as opposed to MLE offers the estimation of parameters without computing the values of p.d.f.'s (and hence, the derivatives of compounding functions), but only with knowing the values of c.d.f.'s. This is what one needs to make the computations even more efficient.

Ranneby, Jammalamadaka, and Teterukovskiy (2005) extend MSPE approach for the multivariate i.i.d. case, however there does not exist that extension to the multivariate time series case. When such an extension will be devised, the Maximum Spacings Estimator may be a useful alternative to the Maximum Likelihood Estimator for using in the new sequential method of joint distributions modeling.

6.4.2 Estimates of the parameters

For the present, the new sequential approach uses Maximum Likelihood for estimation. The following tables summarize the estimates of the parameters of the compounding

functions considered. The first three tables show the values of the parameters in the t-based approach for different groups of assets, and the second three tables show the values in the Clayton-based approach.

$\widetilde{C}(\cdot;\cdot)$	α	β	η	$\overline{\rho}$	a	b
(GE; MCD, MSFT)	0.044	0.322	8.612	0.884	0.004	0.992
(MCD; GE, MSFT)	0.007	0.011	8.557	0.408	0.124	0.351
(MSFT; GE, MCD)	0.008	0.001	8.828	0.503	0.143	0.092
(GE; MCD, KO)	0.002	0.251	8.544	0.499	0.001	0.249
(MCD; GE, KO)	0.000	0.112	9.928	0.403	0.059	0.399
(KO; GE, MCD)	0.017	0.008	61.964	0.446	0.048	0.334
(GE; MCD, PG)	0.023	0.114	14.379	0.468	0.109	0.059
(MCD; GE, PG)	0.009	0.028	5.197	0.338	0.114	0.423
(PG; GE, MCD)	0.001	0.194	5.306	0.472	0.106	0.275
(GE; MSFT, KO)	0.076	0.224	8.508	0.670	0.076	0.128
(MSFT; GE, KO)	0.007	0.007	6.867	0.536	0.036	0.232
(KO; GE, MSFT)	0.031	0.162	8.459	0.602	0.004	0.314
(GE; MSFT, PG)	0.002	0.161	10.151	0.638	0.015	0.946
(MSFT; GE, PG)	0.030	0.001	10.296	0.498	0.101	0.034
(PG; GE, MSFT)	0.057	0.251	8.786	0.612	0.060	0.242
(GE; KO, PG)	0.046	0.069	13.185	0.496	0.001	0.033
(KO; GE, PG)	0.009	0.001	23.245	0.497	0.066	0.484
(PG; GE, KO)	0.001	0.234	7.505	0.612	0.136	0.016
(MCD; MSFT, KO)	0.005	0.219	6.018	0.469	0.119	0.362
(MSFT; MCD, KO)	0.003	0.001	9.917	0.436	0.054	0.402
(KO; MCD, MSFT)	0.129	0.025	19.673	0.541	0.026	0.286
(MCD; MSFT, PG)	0.133	0.315	8.481	0.588	0.358	0.136
(MSFT; MCD, PG)	0.001	0.000	6.562	0.407	0.078	0.606
(PG; MCD, MSFT)	0.001	0.128	5.165	0.425	0.120	0.561
(MCD; KO, PG)	0.000	0.004	6.686	0.320	0.154	0.324
(KO; MCD, PG)	0.231	0.020	19.575	0.577	0.066	0.605
(PG; MCD, KO)	0.000	0.405	4.553	0.587	0.164	0.365
(MSFT; KO, PG)	0.002	0.001	6.930	0.473	0.125	0.015
(KO; MSFT, PG)	0.002	0.001	23.721	0.530	0.067	0.251
(PG; MSFT, KO)	0.022	0.143	8.385	0.544	0.159	0.221

Table 5: MSP estimates of the parameters of the t-based compounding functions for the group of three assets

$\widetilde{C}(\cdot;\cdot)$	α	β	η	$\overline{\rho}$	a	b
(GE; MCD, MSFT, KO)	0.057	0.326	8.580	0.630	0.007	0.646
(MCD; GE, MSFT, KO)	0.005	0.180	8.326	0.440	0.137	0.244
(MSFT; GE, MCD, KO)	0.022	0.110	8.443	0.508	0.058	0.238
(KO; GE, MCD, MSFT)	0.015	0.062	8.540	0.466	0.023	0.477
(GE; MCD, MSFT, PG)	0.026	0.137	8.868	0.523	0.096	0.103
(MCD; GE, MSFT, PG)	0.046	0.201	8.464	0.447	0.189	0.390
(MSFT; GE, MCD, PG)	0.007	0.008	8.666	0.423	0.185	0.047
(PG; GE, MCD, MSFT)	0.025	0.235	6.492	0.511	0.092	0.494
(GE; MCD, KO, PG)	0.026	0.269	10.768	0.447	0.092	0.188
(MCD; GE, KO, PG)	0.023	0.228	8.492	0.400	0.163	0.320
(KO; GE, MCD, PG)	0.140	0.073	8.825	0.440	0.101	0.304
(PG; GE, MCD, KO)	0.048	0.400	8.507	0.675	0.079	0.730
(GE; MSFT, KO, PG)	0.013	0.109	8.563	0.553	0.004	0.706
(MSFT; GE, KO, PG)	0.004	0.004	9.226	0.475	0.105	0.130
(KO; GE, MSFT, PG)	0.022	0.073	9.093	0.506	0.064	0.316
(PG; GE, MSFT, KO)	0.022	0.225	8.615	0.591	0.120	0.186
(MCD; MSFT, KO, PG)	0.022	0.356	8.839	0.481	0.229	0.112
(MSFT; MCD, KO, PG)	0.023	0.056	8.487	0.418	0.100	0.345
(KO; MCD, MSFT, PG)	0.041	0.022	8.652	0.450	0.066	0.757
(PG; MCD, MSFT, KO)	0.035	0.327	8.487	0.564	0.102	0.552

Table 6: MSP estimates of the parameters of the t-based compounding functions for the group of four assets

$\widetilde{C}(\cdot;\cdot)$	α	β	η	$\overline{\rho}$	a	b
(GE; MCD,MSFT,KO,PG)	0.175	0.265	10.818	0.546	0.157	0.227
(MCD; GE,MSFT,KO,PG)	0.189	0.410	9.088	0.578	0.247	0.199
(MSFT; GE,MCD,KO,PG)	0.394	0.301	8.580	0.463	0.309	0.285
(KO; GE,MCD,MSFT,PG)	0.285	0.585	8.680	0.433	0.312	0.321
(PG; GE,MCD,MSFT,KO)	0.045	0.252	8.499	0.520	0.051	0.782

Table 7: MSP estimates of the parameters of the t-based compounding functions for the group of five assets

$\widetilde{C}(\cdot;\cdot)$	α	β	$\overline{\tau}$	a	b
(GE; MCD, MSFT)	0.001	0.282	0.382	0.074	0.111
(MCD; GE, MSFT)	0.150	0.210	0.423	0.184	0.141
(MSFT; GE, MCD)	0.000	0.075	0.299	0.094	0.010
(GE; MCD, KO)	0.212	0.227	0.449	0.000	0.008
(MCD; GE, KO)	0.006	0.309	0.391	0.000	0.013
(KO; GE, MCD)	0.003	0.002	0.239	0.000	0.639
(GE; MCD, PG)	0.236	0.332	0.466	0.163	0.003
(MCD; GE, PG)	0.204	0.286	0.434	0.215	0.003
(PG; GE, MCD)	0.002	0.429	0.418	0.061	0.013
(GE; MSFT, KO)	0.086	0.013	0.340	0.002	0.366
(MSFT; GE, KO)	0.000	0.000	0.292	0.178	0.007
(KO; GE, MSFT)	0.149	0.291	0.513	0.015	0.524
(GE; MSFT, PG)	0.149	0.003	0.303	0.037	0.862
(MSFT; GE, PG)	0.185	0.000	0.343	0.070	0.719
(PG; GE, MSFT)	0.273	0.383	0.702	0.244	0.097
(GE; KO, PG)	0.358	0.337	0.637	0.176	0.000
(KO; GE, PG)	0.135	0.289	0.456	0.000	0.002
(PG; GE, KO)	0.269	0.323	0.578	0.151	0.018
(MCD; MSFT, KO)	0.001	0.298	0.370	0.019	0.309
(MSFT; MCD, KO)	0.000	0.262	0.385	0.088	0.027
(KO; MCD, MSFT)	0.011	0.014	0.259	0.007	0.292
(MCD; MSFT, PG)	0.004	0.144	0.249	0.256	0.010
(MSFT; MCD, PG)	0.146	0.216	0.370	0.126	0.212
(PG; MCD, MSFT)	0.355	0.353	0.606	0.324	0.010
(MCD; KO, PG)	0.078	0.289	0.345	0.050	0.007
(KO; MCD, PG)	0.236	0.001	0.332	0.009	0.929
(PG; MCD, KO)	0.295	0.338	0.506	0.000	0.001
(MSFT; KO, PG)	0.228	0.242	0.511	0.127	0.417
(KO; MSFT, PG)	0.160	0.251	0.449	0.136	0.023
(PG; MSFT, KO)	0.231	0.287	0.541	0.205	0.013

Table 8: MSP estimates of the parameters of the Clayton-based compounding functions for the group of three assets

$\widetilde{C}(\cdot;\cdot)$	α	β	$\overline{\tau}$	a	b
(GE; MCD, MSFT, KO)	0.073	0.186	0.369	0.002	0.242
(MCD; GE, MSFT, KO)	0.025	0.317	0.403	0.111	0.251
(MSFT; GE, MCD, KO)	0.012	0.238	0.405	0.073	0.204
(KO; GE, MCD, MSFT)	0.009	0.016	0.260	0.002	0.133
(GE; MCD, MSFT, PG)	0.101	0.176	0.345	0.013	0.264
(MCD; GE, MSFT, PG)	0.142	0.288	0.404	0.216	0.120
(MSFT; GE, MCD, PG)	0.083	0.248	0.392	0.094	0.119
(PG; GE, MCD, MSFT)	0.141	0.329	0.429	0.161	0.117
(GE; MCD, KO, PG)	0.177	0.287	0.434	0.091	0.008
(MCD; GE, KO, PG)	0.091	0.307	0.393	0.122	0.267
(KO; GE, MCD, PG)	0.016	0.036	0.246	0.002	0.382
(PG; GE, MCD, KO)	0.021	0.416	0.444	0.016	0.238
(GE; MSFT, KO, PG)	0.145	0.349	0.472	0.069	0.128
(MSFT; GE, KO, PG)	0.108	0.283	0.434	0.085	0.310
(KO; GE, MSFT, PG)	0.172	0.317	0.485	0.083	0.119
(PG; GE, MSFT, KO)	0.273	0.375	0.643	0.218	0.342
(MCD; MSFT, KO, PG)	0.011	0.272	0.317	0.195	0.164
(MSFT; MCD, KO, PG)	0.093	0.227	0.376	0.127	0.325
(KO; MCD, MSFT, PG)	0.027	0.051	0.256	0.023	0.297
(PG; MCD, MSFT, KO)	0.304	0.316	0.571	0.266	0.106

Table 9: MSP estimates of the parameters of the Clayton-based compounding functions for the group of four assets

$\widetilde{C}(\cdot;\cdot)$	α	β	$\overline{\tau}$	a	b
(GE; MCD,MSFT,KO,PG)	0.144	0.269	0.397	0.049	0.109
(MCD; GE,MSFT,KO,PG)	0.123	0.335	0.403	0.189	0.210
(MSFT; GE,MCD,KO,PG)	0.075	0.236	0.392	0.070	0.281
(KO; GE,MCD,MSFT,PG)	0.069	0.187	0.313	0.028	0.188
(PG; GE,MCD,MSFT,KO)	0.248	0.304	0.494	0.199	0.080

Table 10: MSP estimates of the parameters of the Clayton-based compounding functions for the group of five assets

The standard errors of the estimates are not presented on the tables due to their huge size.

6.5 Constructing distributions through aggregating functions

The overall distribution is constructed through taking the mean of the estimated compounding functions from the previous step. The intuition is to cancel out the individual errors obtained in each step of compounding functions estimation. The results are summarized in the next subsection as goodness-of-fit tests. Due to huge number of serial correlation tests conducted, their results are not included in this paper. However, it is worth saying that they are passed on 95% confidence level almost in all cases.

6.5.1 Goodness-of-fit tests. Comparison to five-dimensional t-copula

This subsection compares three approaches for dynamic modeling of joint distributions. The first one is using time-varying five-dimensional t-copula estimated through ML. The second and the third ones are the demonstrations of the new sequential approach based on (1) asymmetrized time-varying bivariate t-copula and (2) asymmetrized time-varying bivariate Clayton-copula respectively. The graphical results are presented in the Appendix. Note: two dimensional conditional densities are not included for 5D t-copula here, however all Kolmogorov-Smirnov tests are passed for them (as for that of the new sequential approach).

Before conducting goodness-of-fit tests, the benchmark, which the new sequential approach will be compared to, is estimated. Time-varying five-dimensional t-copula is used as such benchmark. Its parameters estimates are summarized in the following table:

η	13.426
	(4.380)
a	0.030
	(0.035)
b	0.157
	(0.308)

$\{\overline{\rho}_{ij}\}$	1	2	3	4	5
1	1.000	0.425	0.621	0.502	0.510
	(0.000)	(0.055)	(0.042)	(0.055)	(0.049)
2	0.425	1.000	0.415	0.398	0.367
	(0.055)	(0.000)	(0.056)	(0.055)	(0.062)
3	0.621	0.415	1.000	0.539	0.465
	(0.042)	(0.056)	(0.000)	(0.053)	(0.057)
4	0.502	0.398	0.539	1.000	0.495
	(0.055)	(0.055)	(0.053)	(0.000)	(0.049)
5	0.510	0.367	0.465	0.495	1.000
	(0.049)	(0.062)	(0.057)	(0.049)	(0.000)

Table 11: Parameters estimates for the time-varying five-dimensional t-copula, that models the distribution of (1) GE, (2) MCD, (3) MSFT, (4) KO, and (5) PG log-returns on NYSE from Jan 03, 2007 to Dec 31, 2007. Robust standard errors are in the round brackets

One can see that both five-dimensional copula-based approach and the new sequential approach pass the Kolmogorov-Smirnov tests for all distributions on 95% confidence level. However, time-varying five-dimensional t-copula and t-based new method do not pass the test on 90% confidence level for one conditional distribution (MSFT | GE, MCD, KO), whereas Clayton-based new method still passes the test on that level. Although, there is no explicit leader in this comparison, it should be stated here that the new method is still have the advantage: it is feasible for higher dimension, because the number of parameters in the optimization problem remains fixed (5 in the case of Clayton-based approach), whereas in standard single-copula based approach this number usually grows fast and optimization problem may become computationally infeasible. The following table demonstrates how the number of parameters in each optimization problem for standard and new approaches grows with the dimension of the whole joint distribution.

dimension	Standard t-copula	New t-based	New Clayton-based
3	6	6	5
4	9	6	5
5	13	6	5
6	18	6	5
7	24	6	5
8	31	6	5
9	39	6	5
10	48	6	5
11	58	6	5
12	69	6	5
13	81	6	5
14	94	6	5
15	108	6	5

Table 12: Growth of the number of parameters in each optimization problem for the standard t-copula approach and the new t-based sequential method

It seems that the simultaneous whole-distribution estimation is computationally infeasible for the dimensions of 10 and higher. Although, there are a lot optimization problems to solve in the new sequential approach, it proposes a great alternative to the standard approaches for modeling dynamic joint distributions in such high-dimensional

problems. For example, five dimensional problem requires solving 5 problems with 7 parameters on the first step, 20 problems with 6 parameters on the second step, 30 problems with 5 parameters on the third step, 20 problems with 5 parameters on the fourth step, and 5 problems with 5 parameters on the fifth step, whereas the standard t-copula based approach requires estimating of 5 problems with 7 parameters on the first step and 1 problem with 13 parameters on the second step. One can see that there are a huge number of the optimization problems in the new sequential approach, however, it takes only a few seconds to solve one problem for the Clayton-based compounding function, whereas the standard approach requires tens of minutes in order to be solved for such large number of parameters. This makes the whole estimation time for the new sequential approach even less than for the standard one. The time difference seems to become even greater in higher-dimensional case.

7 Further research

This paper has suggested the new sequential approach for dynamic modeling of joint distributions, and thus, outlined the directions of further research:

1. develop other goodness-of-fit tests that will be computationally feasible in higher dimensions (here only five stocks are considered, because in higher dimensions the conducted goodness-of-fit tests are computationally infeasible for t-based copulas as their c.d.f. is not a function that can be easily computed, it is an n-dimensional integral that can not be expressed in elementary functions);

2. examine the approaches of simplifications of the calculations by either using Maximum Spacings Estimator or considering random pairs, triples, etc. (note: the second simplification should be done with a great caution due to huge interdependence among steps of the new sequential approach);

3. use the approach for financial applications (like computing value at risk, expected shortfall, or other quantities that could be obtained from knowing the whole joint distribution of the assets) and conduct out-of-sample assessment of the effectiveness of the method.

8 Conclusion

This paper has developed the new sequential approach for dynamic joint distributions modeling based on combining small-dimensional distributions into higher-dimensional ones by compounding and aggregating functions. Additionally, it has demonstrated its implementation for the series of log-returns of five NYSE-traded stocks. (The author has considered only five stocks because in higher dimensions the conducted goodness-of-fit tests are computationally infeasible for t-based copulas as their c.d.f. is not a function that can be easily computed: it is an n-dimensional integral that can not be expressed in elementary functions.) First, for estimating the marginal distributions Skew-t-NAGARCH specification is used and it seems to be effective in modeling the log-returns, capturing heavy tails, skewness and leverage effect observed in the data. Then, the t-copula is applied for modeling pairwise bivariate distributions of the log-returns and it is proved to be rather good method for this purpose. Next, five-dimensional distributions are constructed using the new sequential approach that divides the huge problem into smaller ones, and hence, unlike to the standard sufficiently flexible single-copula based approaches, seems to be computationally feasible in very high dimensions. All goodness-of-fit tests are passed and indicate the usefulness of the new methodology, however they do not distinguish a single method as the best one.

The main advantage of the new sequential approach is that it makes the huge problem of dynamic modeling of the joint distributions of a number of stocks' log-returns feasible in practice. Although, there are a lot of optimization problems to solve while using the new method, each of them has fixed number of parameters across all dimensions, and hence, their estimation becomes computationally feasible for vast dimensional cases and makes it possible to use all advantages of modeling the whole stochastic behavior of the group of a number of stocks.

9 Acknowledgements

The author takes pleasure in thanking his research adviser Stanislav Anatolyev and Artem Prokhorov for useful directions.

10 References

1. Anatolyev, Stanislav and Grigory Kosenok (2005). An alternative to maximum likelihood based on spacings, *Econometric Theory*, Vol. 21, No. 2, pp. 472–476.

2. Ausin, M. Concepcion and Hedibert F. Lopes (2010). Time-varying joint distribution through copulas, *Computational Statistics and Data Analysis*, Vol. 54, pp. 2383–2399.

3. Azzalini, Adelchi and Antonella Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society*, Vol. 65, pp. 367–389.

4. Bollerslev, Tim (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics*, Vol. 69, pp. 542–547.

5. Breymann, Wolfgang, Alexandra Dias and Paul Embrechts (2003). Dependence structures for multivariate high–frequency data in finance, *Quantitative Finance*, Vol. 3, pp. 1–14.

6. Cheng, R.C.H., and N.A.K. Amin (1983). Estimating parameters in continuous univariate distributions with a shifted origin, *Journal of the Royal Statistical Society*, Vol. 45, No. 3, pp. 394–403.

7. Diebold, Francis X., Todd A. Gunther and Anthony S. Tay (1998). Evaluating density forecasts with applications to financial risk management, *International Economic Review*, Vol. 39, pp. 863–883.

8. Engle, Robert F. (2002). Dynamic Conditional Correlation – a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics*, Vol. 20, pp. 339–350.

9. Engle, Robert F. (2009). High dimension dynamic correlations, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry* edited by Jennifer Castle and Neil Shephard, OUP Oxford, pp. 122–149.

10. Engle, Robert F. and Victor K. Ng (1993). Measuring and testing the impact of news on volatility, *Journal of Finance*, Vol. 48, pp. 1749–1778.

11. Engle, Robert F., Neil Shephard and Kevin Sheppard (2008). Fitting vast dimensional time-varying covariance models, Discussion Paper, Department of Economics, University of Oxford, Oxford.

12. Fernández, Carmen and Mark F.J. Steel (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, Vol. 93, pp. 359–371.

13. Gallant, A. Ronald and George Tauchen, (1989). Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications, *Econometrica*, Vol. 57, pp. 1091–1120.

14. Hansen, Bruce E. (1994). Autoregressive conditional density estimation, *International Economic Review*, Vol. 35, pp. 705–730.

15. Khoudraji, Abdelhaq (1995). Contributions à l'étude des copules et à la modélisation des valeurs extrêmes bivariées, Ph.D. thesis, Université Laval, Sainte-Foy, Québec.

16. Nelsen, Roger B. (2006). *An introduction to copulas (second edition)*, Springer, New York.

17. Patton, Andrew (2006). Modelling asymmetric exchange rate dependence, *International Economic Review*, Vol. 47, pp. 527–556.

18. Ranneby, Bo (1984). The maximum spacing method. An estimation method related to the maximum likelihood method, *Scandinavian Journal of Statistics*, Vol. 11, No. 2, pp. 93–112.

19. Ranneby, Bo, S. Rao Jammalamadaka, and Alex Teterukovskiy (2004). The maximum spacing estimation for multivariate observations, Manuscript, Swedish University of Agricultural Sciences.

20. Rosenblatt, Murray (1952). Remarks on a multivariate transformation, *The Annals of Mathematical Statistics*, Vol. 23, pp. 470–472.

21. Rydberg, Tina Hviid (2000). Realistic statistical modelling of financial data, *International Statistical Review*, Vol. 68, pp. 233–258.

22. Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut Statistique de l'Université de Paris*, Vol. 8, pp. 229–231.

23. Theodossiou, Panayiotis (2000). Skewed Generalized Error Distribution of Financial Assets and Option Pricing, Working Paper, Cyprus University of Technology, Rutgers Business School, Camden.

11 Appendix

(see overleaf)

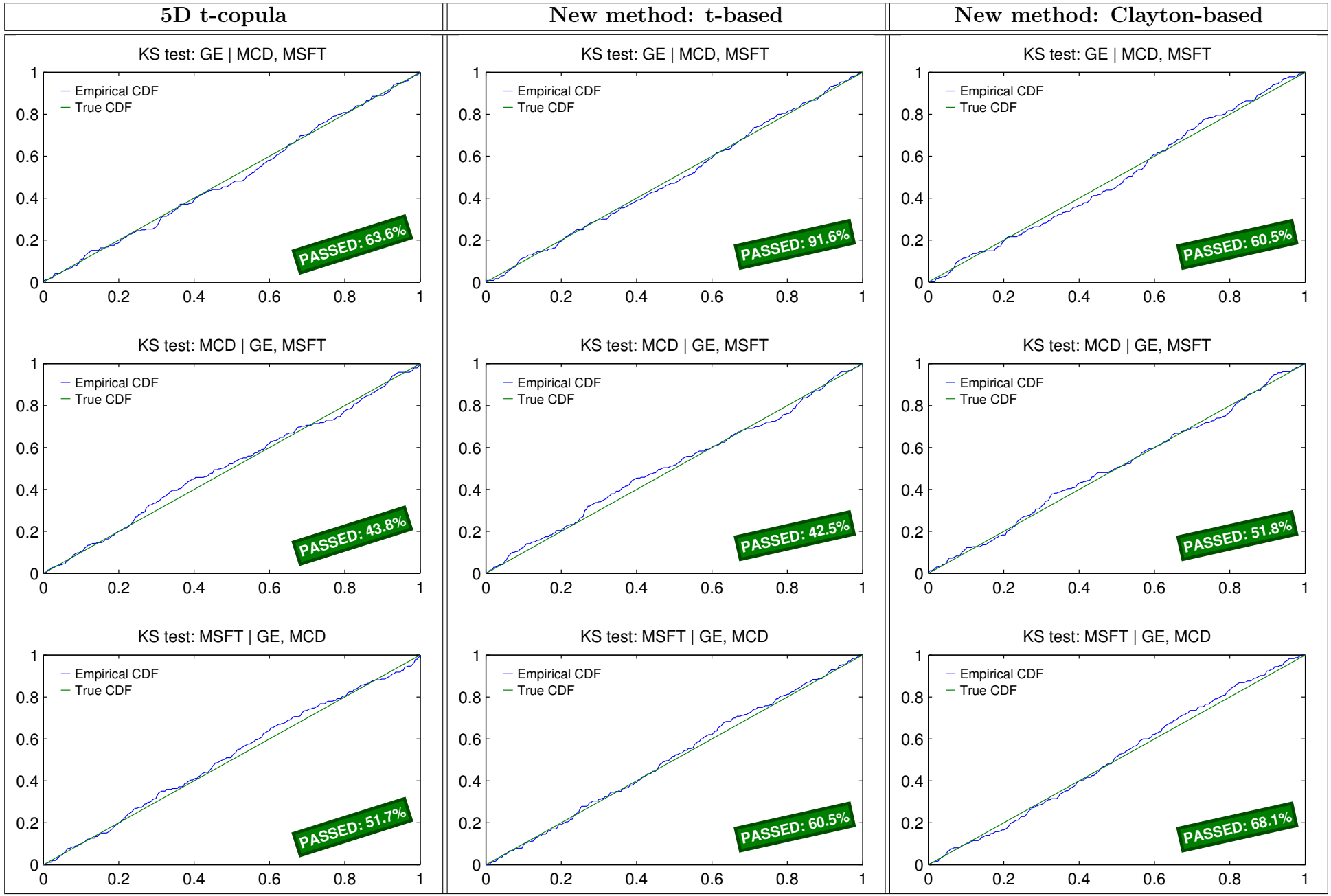

Table 13: Trivariate conditional distributions comparison: GE, MCD, MSFT

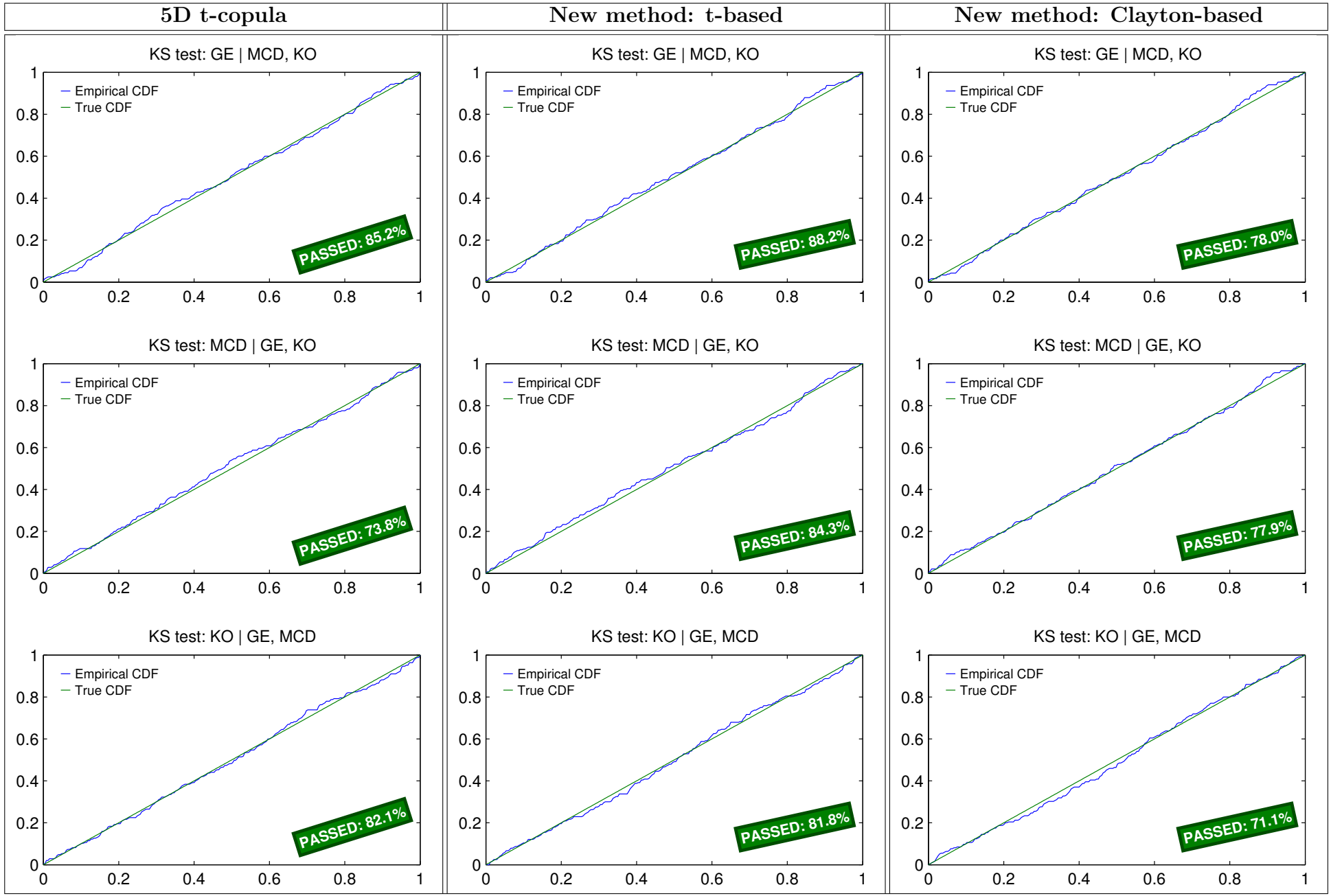

Table 14: Trivariate conditional distributions comparison: GE, MCD, KO

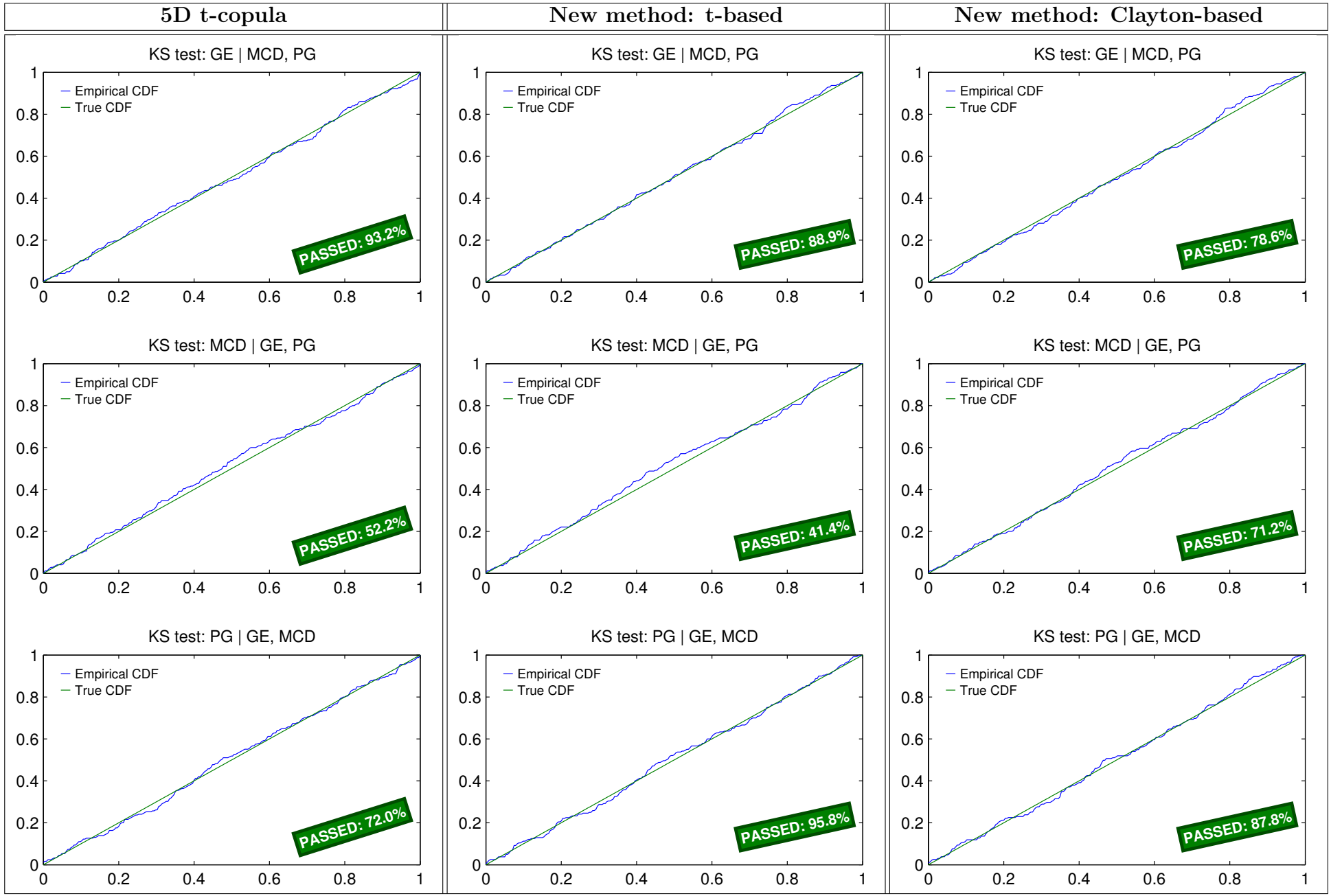

Table 15: Trivariate conditional distributions comparison: GE, MCD, PG

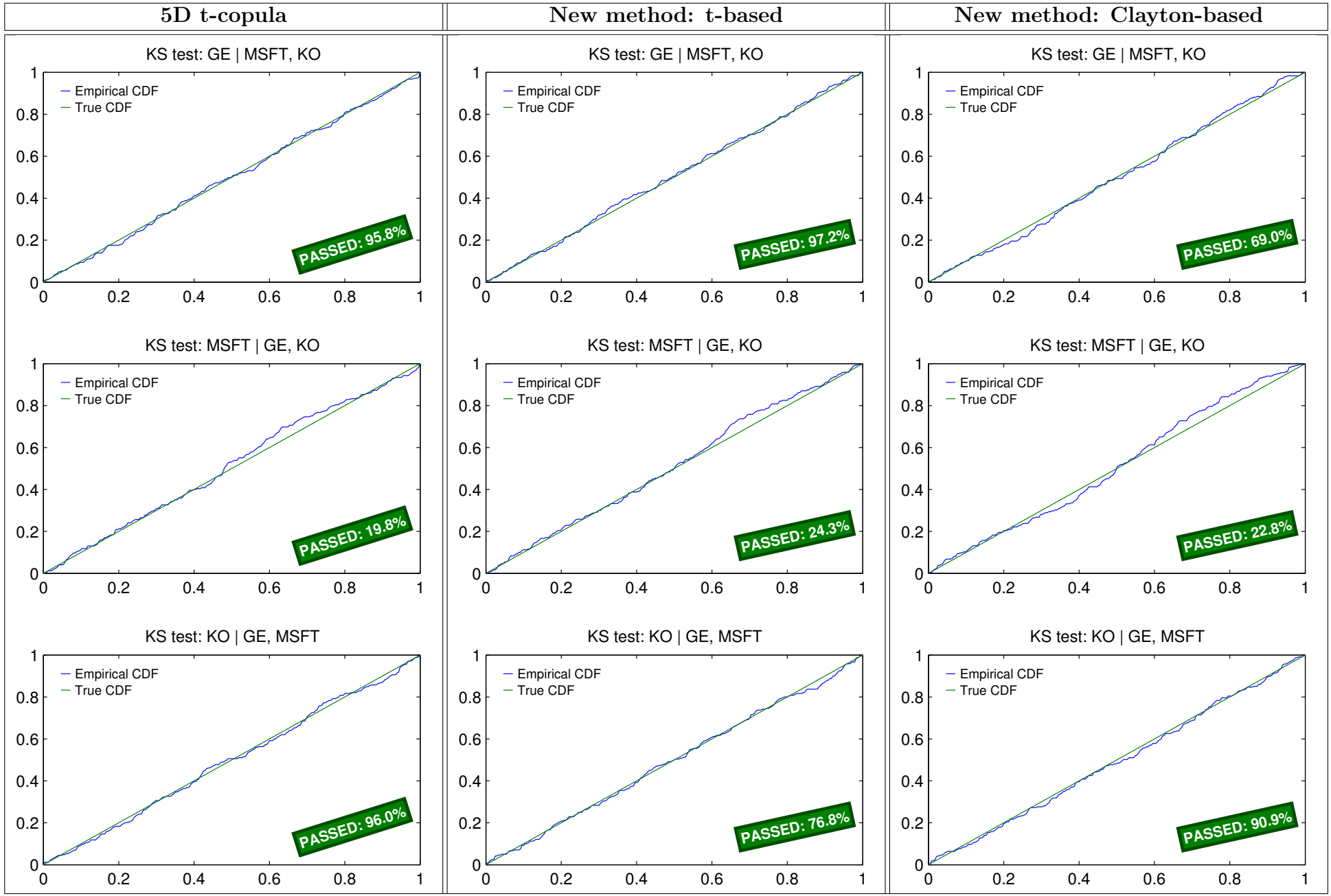

Table 16: Trivariate conditional distributions comparison: GE, MSFT, KO

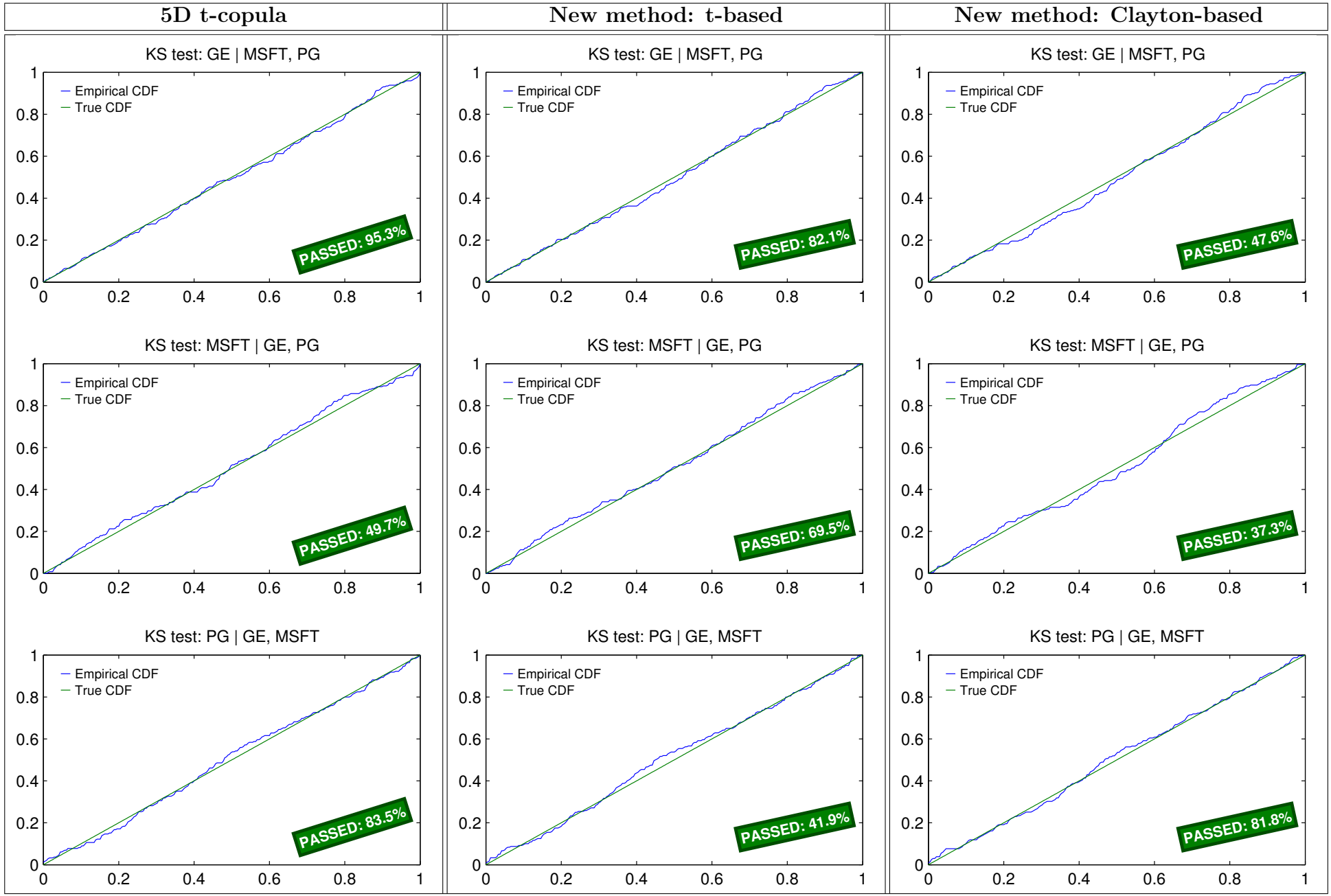

Table 17: Trivariate conditional distributions comparison: GE, MSFT, PG

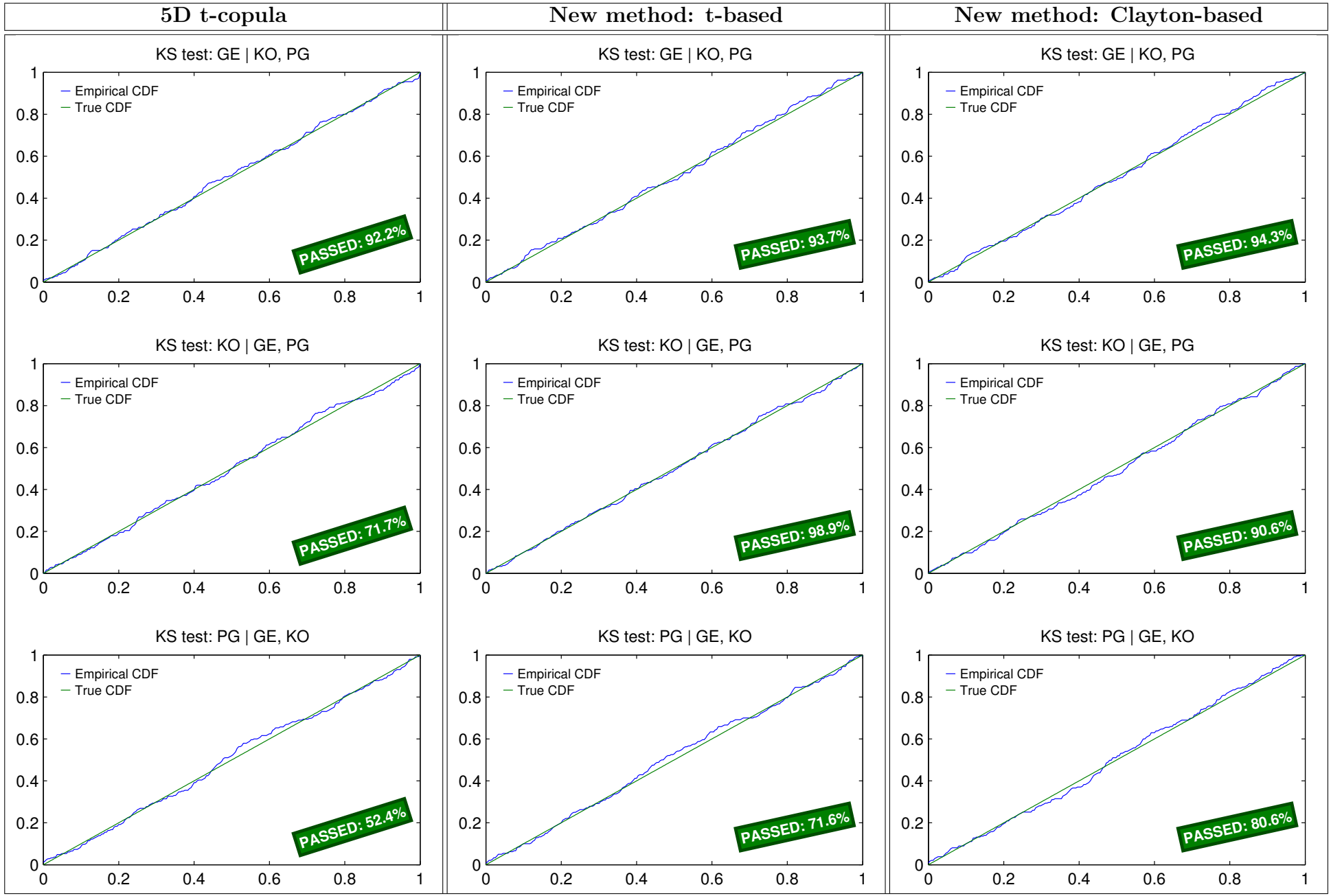

Table 18: Trivariate conditional distributions comparison: GE, KO, PG

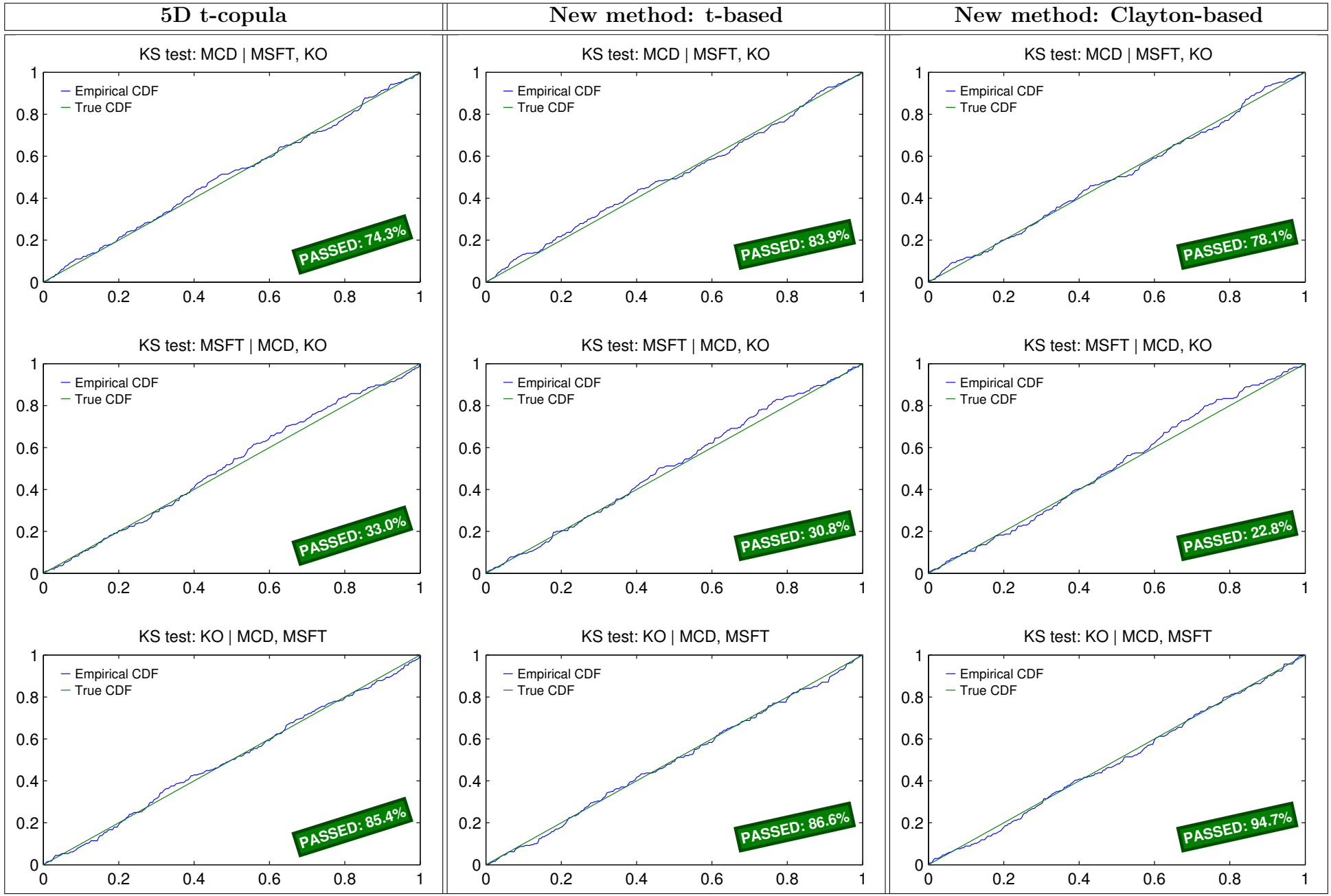

Table 19: Trivariate conditional distributions comparison: MCD, MSFT, KO

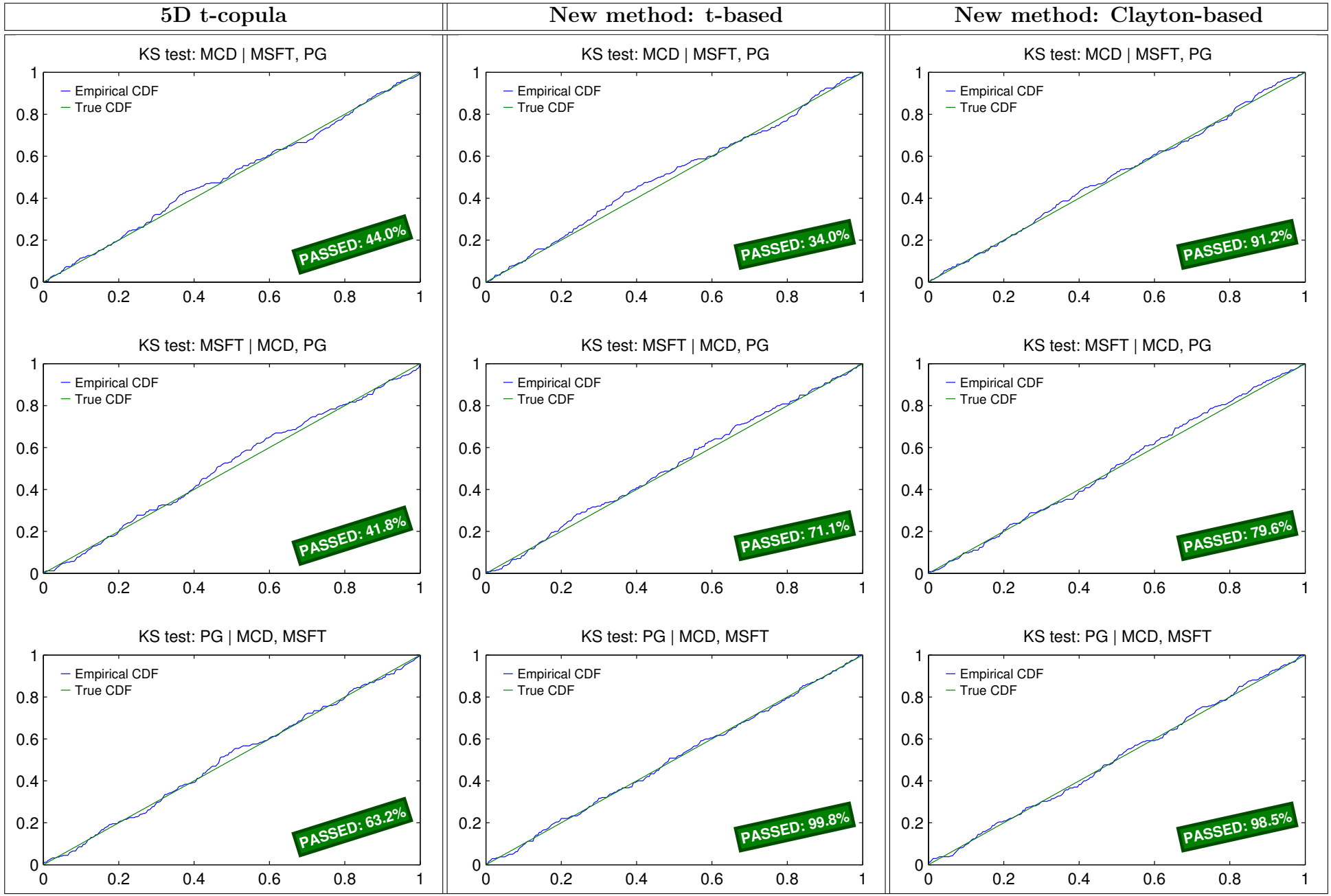

Table 20: Trivariate conditional distributions comparison: MCD, MSFT, PG

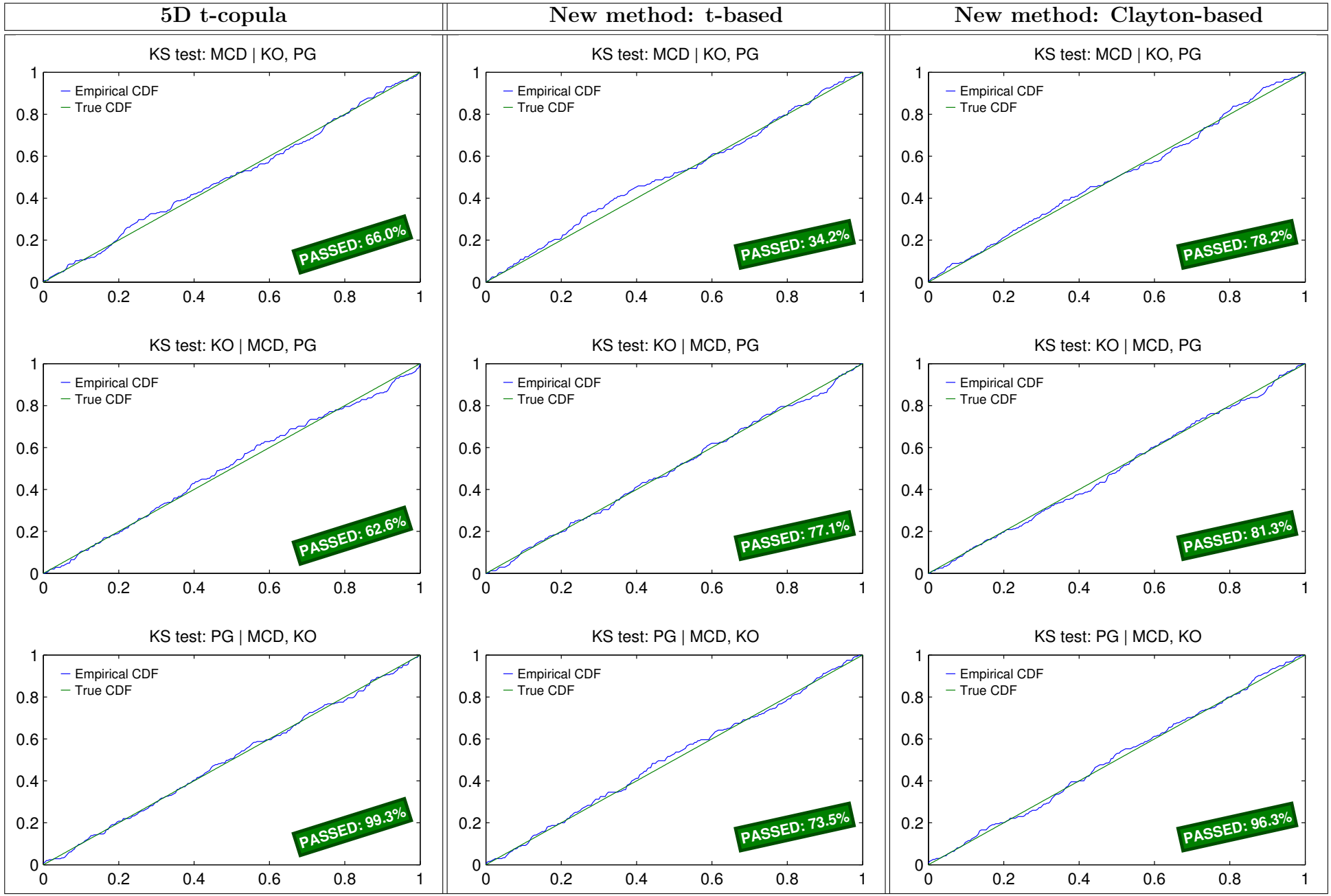

Table 21: Trivariate conditional distributions comparison: MCD, KO, PG

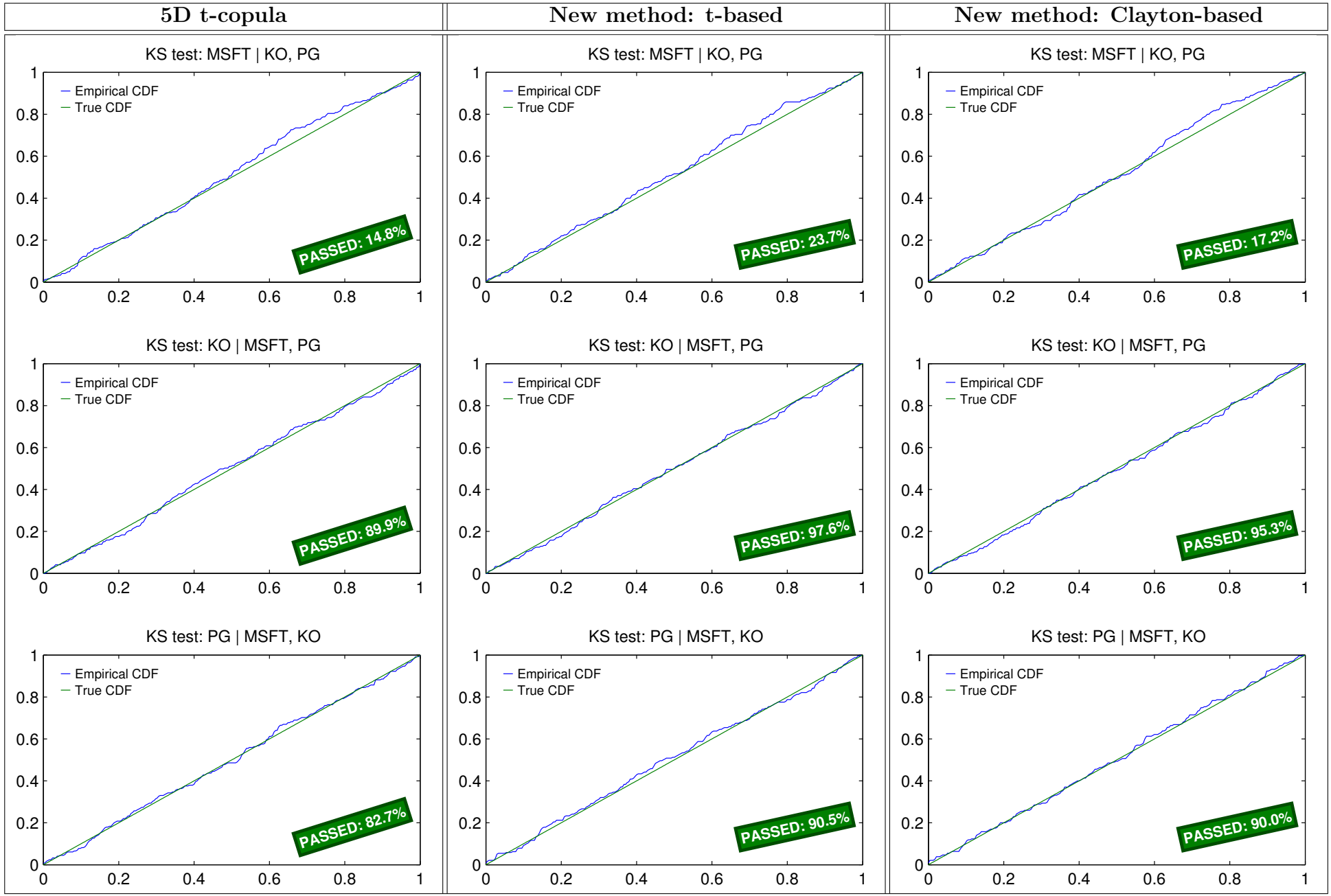

Table 22: Trivariate conditional distributions comparison: MSFT, KO, PG

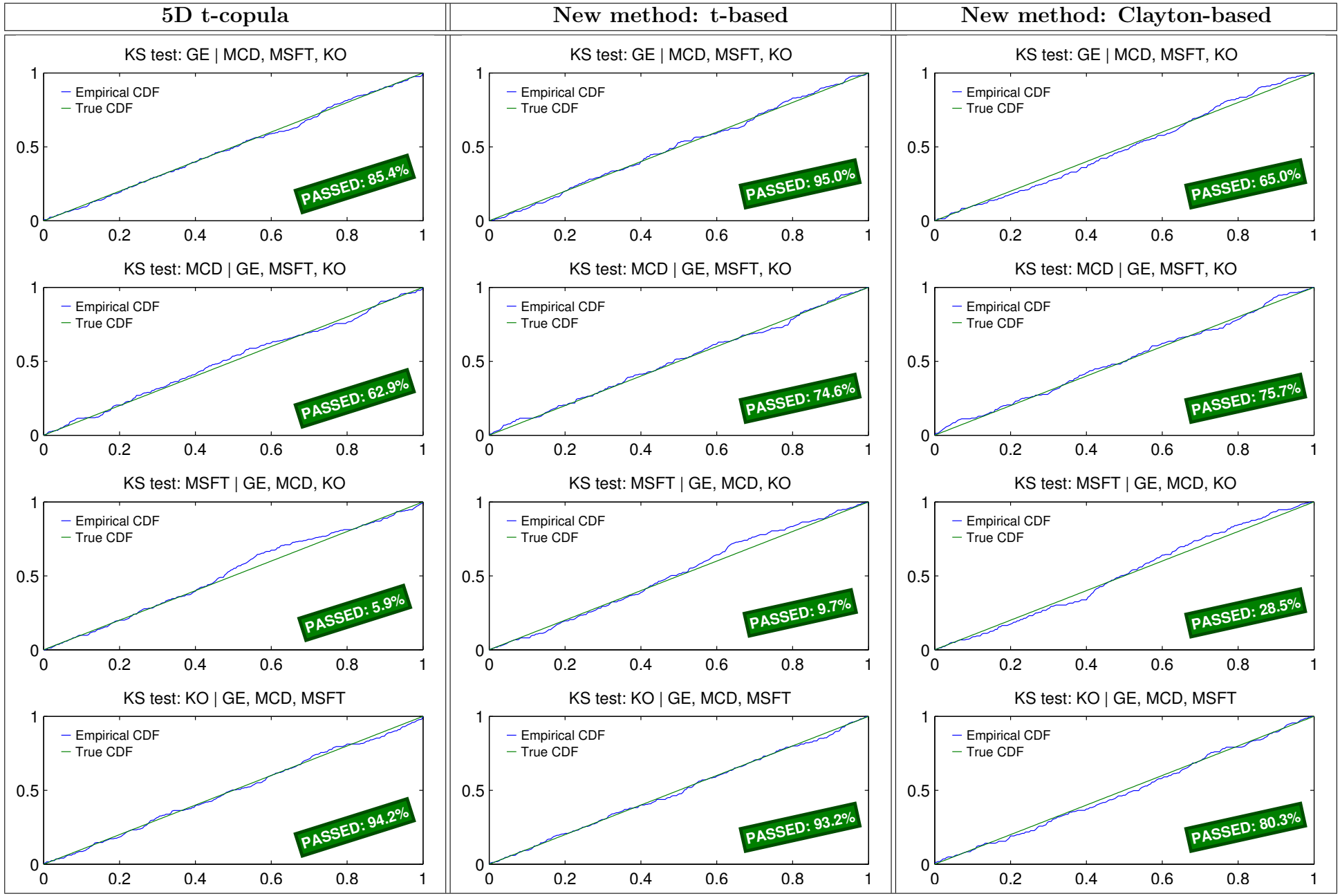

Table 23: Four-dimensional conditional distributions comparison: GE, MCD, MSFT, KO

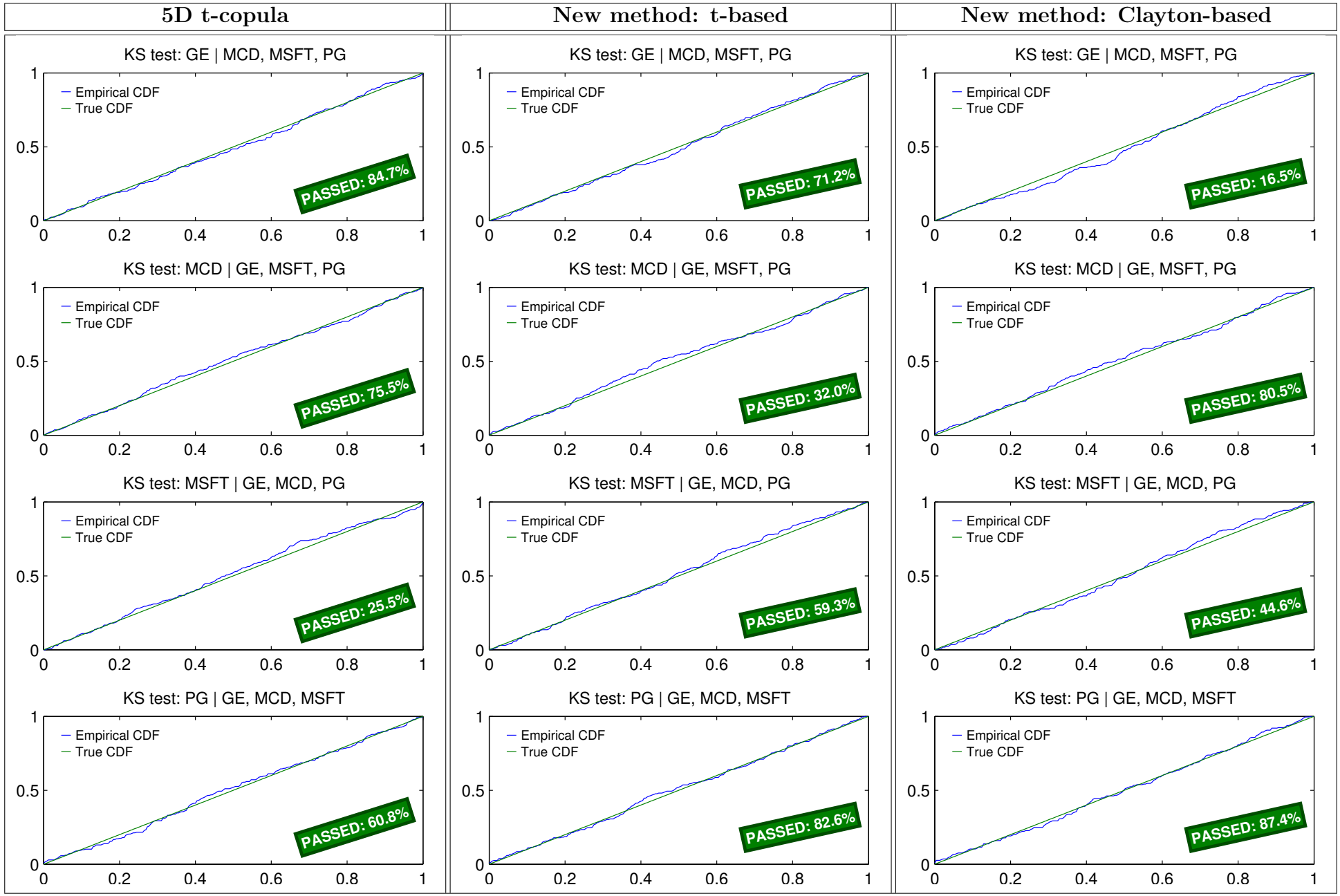

Table 24: Four-dimensional conditional distributions comparison: GE, MCD, MSFT, PG

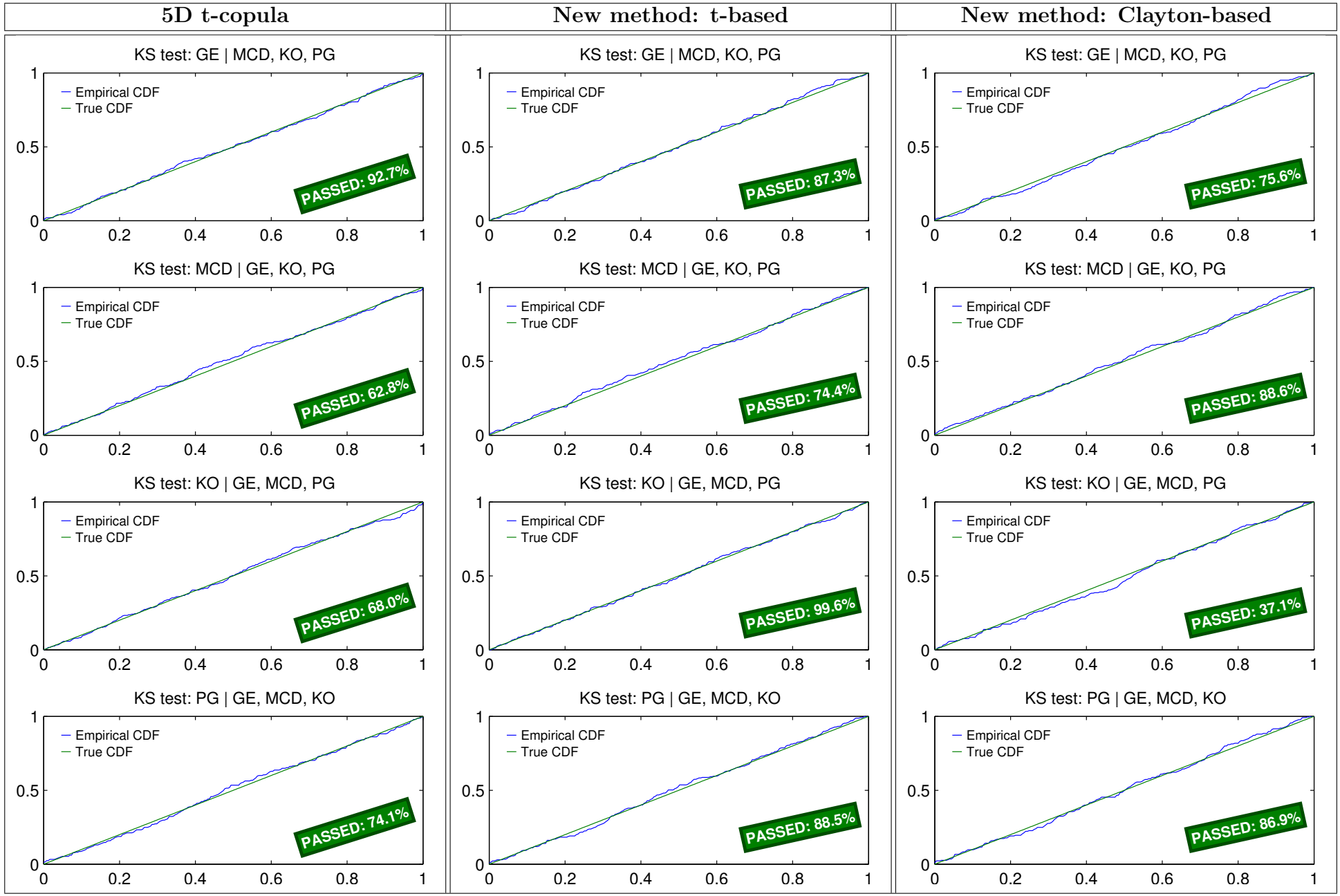

Table 25: Four-dimensional conditional distributions comparison: GE, MCD, KO, PG

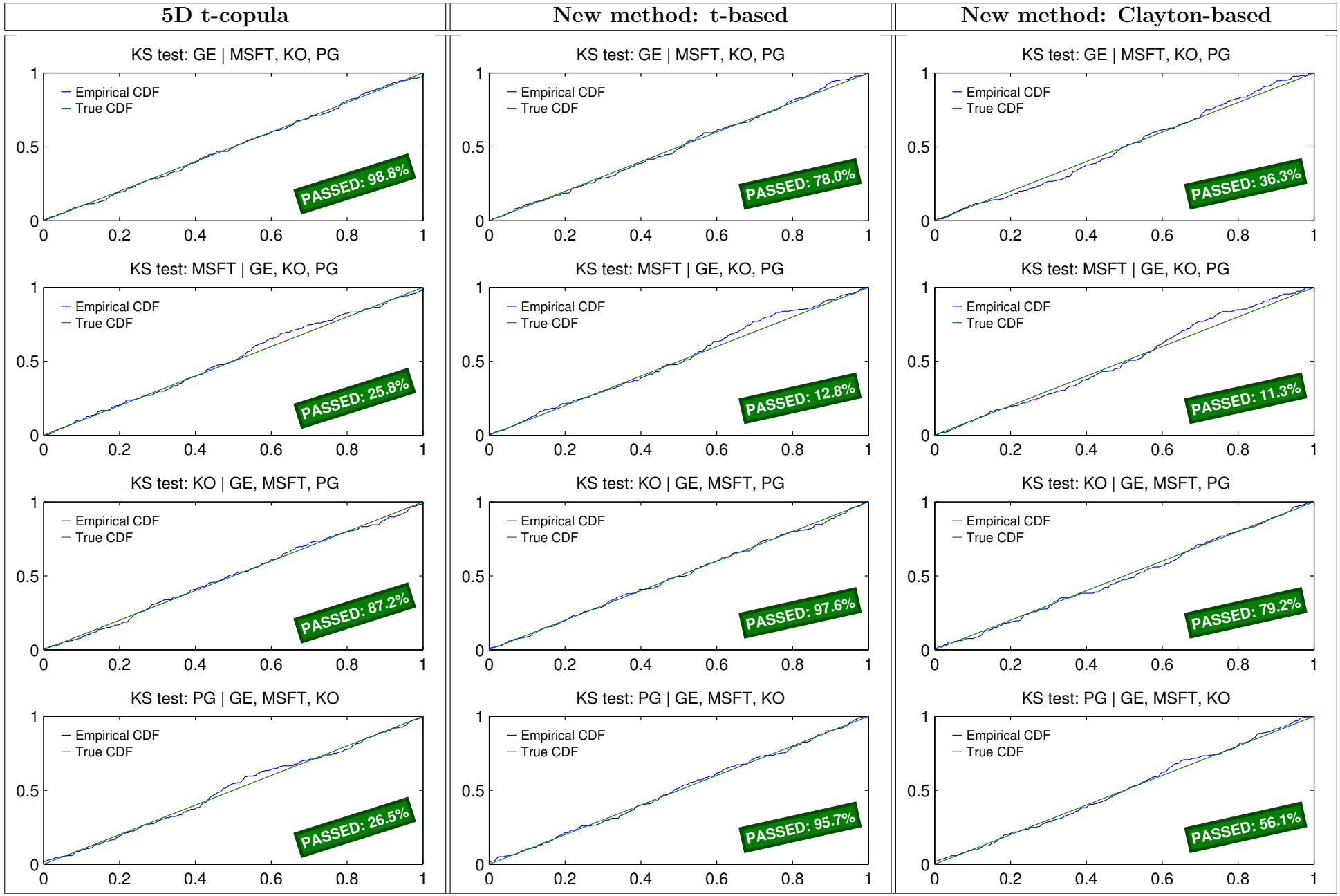

Table 26: Four-dimensional conditional distributions comparison: GE, MSFT, KO, PG

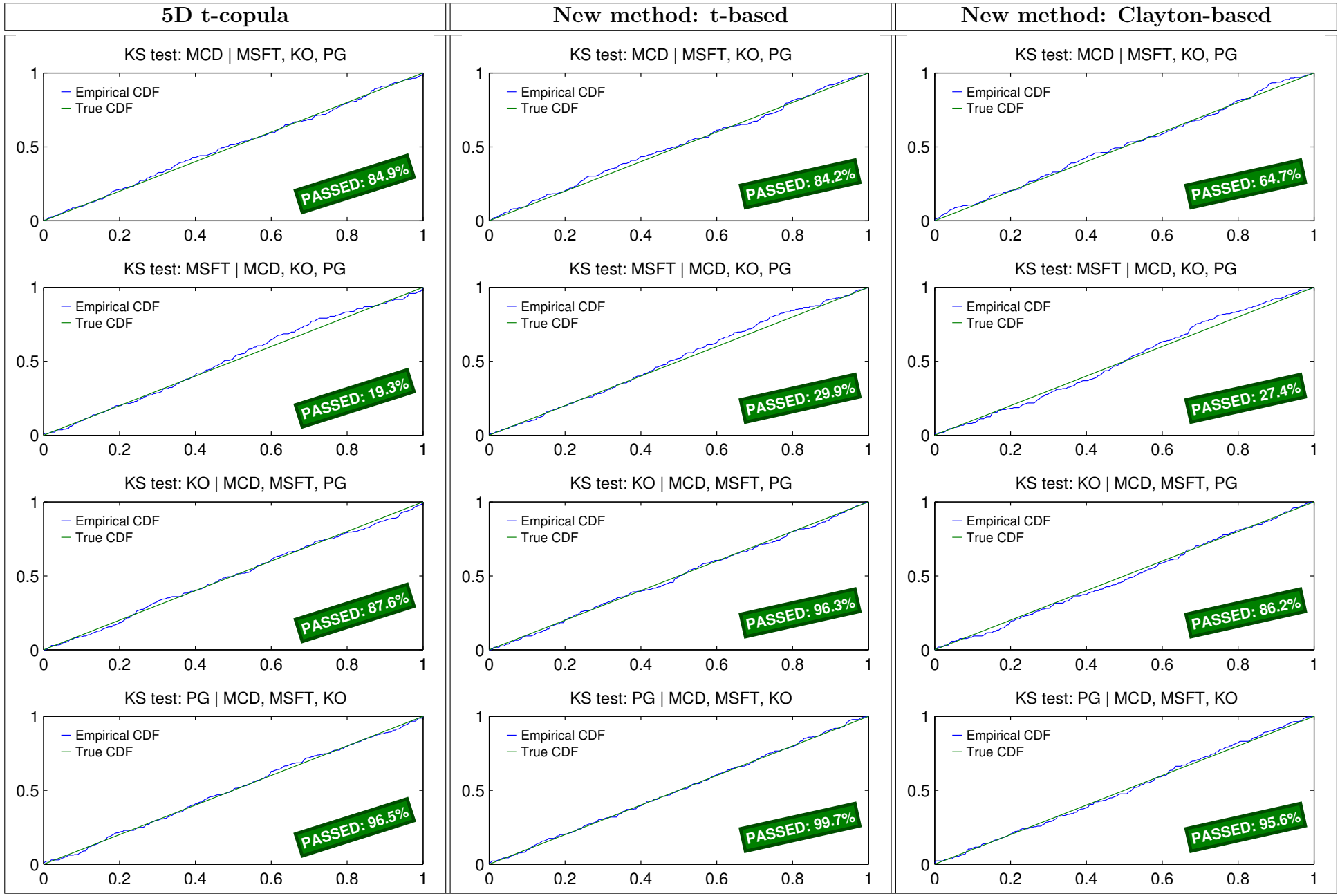

Table 27: Four-dimensional conditional distributions comparison: MCD, MSFT, KO, PG

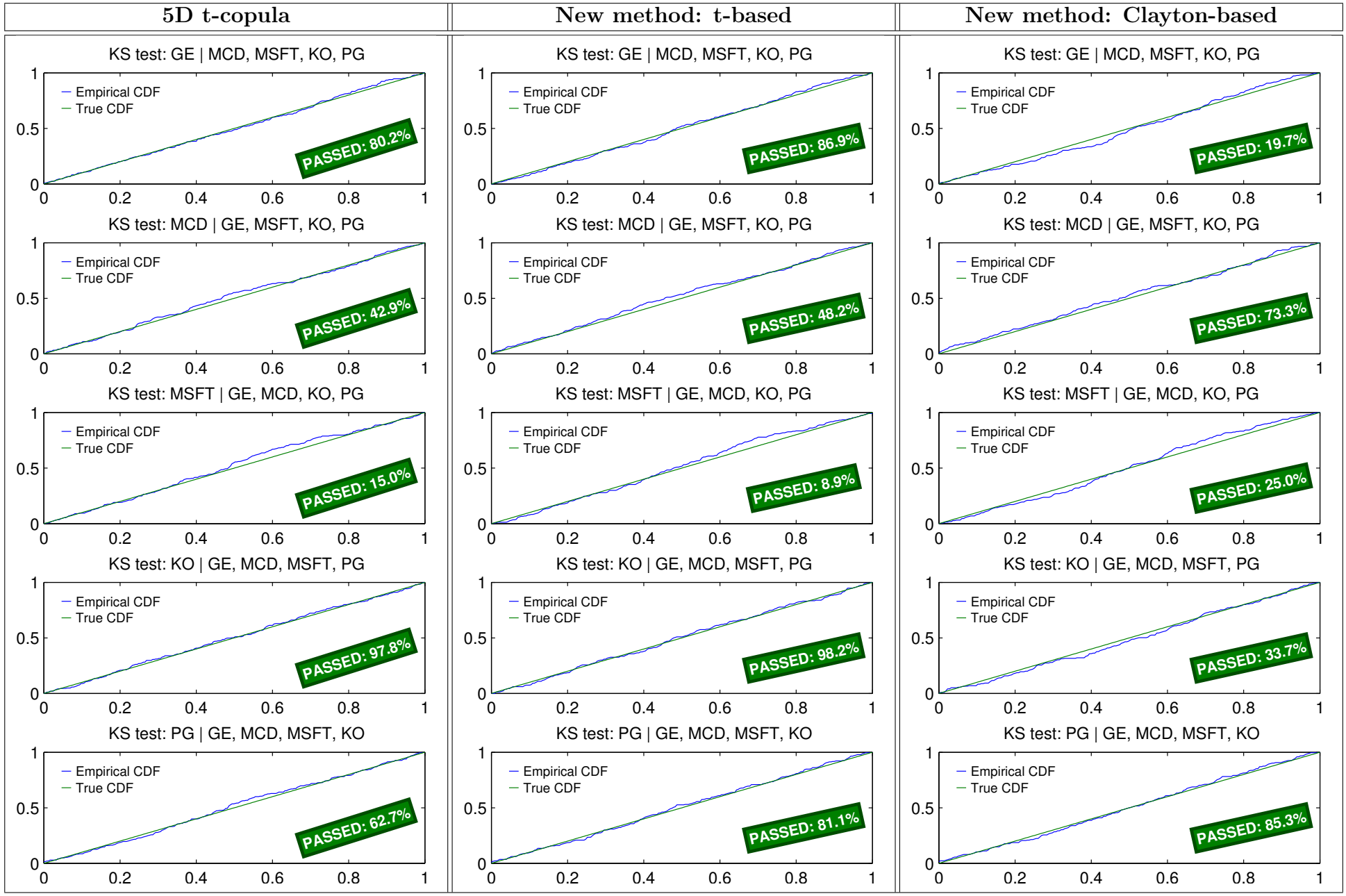

Table 28: Five-dimensional conditional distributions comparison: GE, MCD, MSFT, KO, PG