

Generalized Autocontours:
Evaluation of Multivariate Density Models ¹

Gloria González-Rivera

Yingying Sun

Department of Economics

University of California, Riverside

and

Huazhong University of Science and Technology

March 2014

¹We are grateful to the participants at the International Symposium in Forecasting, Boston, 2012, the ASA Joint Statistical Meetings, San Diego, 2012, CSDA International Conference on Computational and Financial Econometrics, Oviedo, Spain, 2012, and ICSA 2012 Applied Statistics Symposium, Boston for useful comments. Special thanks to Cheng Hsiao, Tae-Hwy Lee, Andrew Patton, Hashem Pesaran, and Aman Ullah for very helpful insights, and other participants at seminars at USC, UCR Statistics, and Federal Reserve Bank of Cleveland. Gloria González-Rivera acknowledges the financial support of UC-Riverside Academic Senate grants.

Abstract

We propose a new tool, the Generalized Autocontour (G-ACR), as the basis for a battery of dynamic specification tests that are applicable (in-sample or out-of-sample) to univariate or multivariate random processes. We apply this methodology to the modeling of a multivariate system by specifying the dynamics of the marginal distributions of each process in the system and a copula that ties up the marginals to produce their multivariate distribution. We work with the probability integral transforms (PIT) of the system that, under correct specification of the conditional model, should be i.i.d. $U[0,1]$. The dimensionality of the system is not a constraint because the information contained in the vector of PITs is condensed into an indicator, which is the basis of the proposed tests. We construct hyper-cubes of different sizes within the maximum hyper-cube formed by a multidimensional uniform density $[0, 1]^n$, and we assess the location of the empirical PITs (duplex, triplex, n-plex of observations) within the corresponding population hyper-cubes. If the conditional model is correct, the volumes of the population hyper-cubes must be the same as those in their empirical counterparts. This approach allows the researcher to focus on different areas of the conditional density model to assess those regions of interest. We estimate a trivariate model for a very large number of trades on the stocks of three large U.S. banks and find that the contemporaneous dependence among institutions is asymmetric, which implies that when liquidity drains (lack of trading) in one institution, we should expect a concurrent effect among similar institutions. On the other hand, when liquidity is plenty (dense trading), the trades on the stocks of the institutions are not correlated. We assess the performance of the models by evaluating the one-step-ahead density forecasts of trades.

Key Words: Autocontour, Copula, Specification Testing, Multivariate Dynamic Models, Density Forecast.

JEL Classification: C01, C32, C34.

1 Introduction

The Generalized Autocontour (G-ACR) is a generalized version of the *autocontour* methodology proposed by González-Rivera *et al.* (2011) (GR2011) to detect misspecification in the dynamics of a time series model and departures from the assumed conditional density model. The G-ACR will overcome some limitations of the original methodology in GR2011. First, when the conditional density of interest departs from standard densities in financial econometrics, e.g. Normal, Student-t, Exponential, etc., the analytical expressions of the autocontours may be mathematically cumbersome to obtain and we need to resort to numerical methods to compute their density mass. The difficulty is compounded when the system is multivariate (González-Rivera and Yoldas, 2012). In contrast, the G-ACR is very easy to obtain for any density because it is based on the probability integral transforms (PIT) instead of standardized innovations, which are the basis of the original ACR. Second, GR2011 considers only univariate stochastic processes with dynamics restricted to the conditional mean and conditional variance, and a time-invariant functional form of the density of the standardized innovations of the model. The advantage of G-ACR is that it is applicable to univariate or multivariate random processes. In a multivariate framework, the dimensionality of the system is not a constraint because the information contained in the vector of PITs is condensed into an indicator, which constitutes the basis of the proposed tests. Furthermore, the components of the multivariate system may have different marginal densities, which could be individually tested, but more importantly, the multivariate density, obtained as a copula function linking the marginals, can also be jointly tested. As a result, our statistics based on G-ACR are also useful diagnostics for correct copula specification. G-ACR does not restrict the dynamics of the model to any particular moment(s) and it is also applicable to cases when the predictive density does not have a closed form solution, e.g. a multistep predictive densities in nonlinear models, and we have to resort to simulation or nonparametric methods, but yet we could obtain the PIT process from the simulated density. Third, the tests proposed in GR2011 have asymptotic variance-covariance matrices that do not all enjoy closed-form solutions, some combining parametric and nonparametric expressions. In contrast and because of the simplicity of G-ACR, the asymptotic variances of the tests have all closed formulations that depend on only

one parameter, the *a priori* specified probability level associated with the G-ACR.

As a brief introduction to G-ACR, explained in detail in the forthcoming sections, suffices to say that the basis of our testing techniques is the construction of hyper-cubes of different sizes within the maximum hyper-cube formed by a multidimensional uniform density $[0, 1]^n$. We assess the location of the empirical PITs (duplex, ..., n-plex of observations) within the corresponding population hyper-cubes. If the multivariate model is correct, the volumes of the population hyper-cubes must be the same as those of their empirical counterparts. Our tests evaluate these differences statistically to either reject or fail to reject the proposed density model. This approach also permits to focus on different areas of the conditional density to assess those regions of interest. There is also a graphical visualization aspect of our approach that is very helpful for guiding the modeling.

As an illustration of the proposed methodology, we will specify a multivariate model for the number of stock trades of three large U.S. banking institutions: Bank of America, JP Morgan Chase, and Wells Fargo. Though the number of trades is a discrete random variable, these three big banks enjoy an almost continuous trading so that, for a given interval of time, the number of trades is so large that the data can be considered to be almost continuous. For instance, at the 5-minute frequency (from January 3 to June 30, 2011), the median number of trades is 1,757 trades for Bank of America, 1,300 for JP Morgan, and 1,210 for Wells Fargo. Dynamic trading is important because it reflects arrival of news and it is intimately related to issues of liquidity risk and market microstructure, see O'Hara (1995) and Madhavan (2000) among others. We proceed by specifying an autoregressive system for the number of trades of each bank. We will entertain different distributional assumptions for the marginal densities of each component of the system but, most importantly, we are interested in the modeling of contemporaneous correlations of the trades as those may have implications for the risk that these large institutions pose to the banking system and beyond. We use a copula function to understand the contemporaneous correlation among the three banks. Heinen and Rengifo (2007) also implemented a copula approach but they restrict themselves to a normal copula where the dependence is contained in a correlation coefficient. As the recent crisis has shown, the correlation among institutions varies during episodes of low or high liquidity. We explore the possibility of asymmetric contemporaneous correlation such that the correlation may be different when the number of trades is large (the market is very active) or when

the number is small (the market is slow). We assess the model in an out-of-sample environment by evaluating the one-step-ahead density forecasts of the number of trades. This modeling and forecasting exercises will allow us to showcase the proposed testing methodology and, in particular, the use of visualization techniques to drive the specification exercise.

The paper is organized as follows. In section 2, we introduce G-ACR and present the testing methodology for univariate and multivariate models. In section 3, we offer extensive Monte Carlo simulations to assess the size and power of the tests within the context of multivariate processes with and without contemporaneous correlation. In section 4, we provide an empirical illustration dealing with a trivariate system for the number of trades for the three large banks mentioned above, and in section 5 we conclude. The appendix contains mathematical proofs.

2 Generalized Autocontour and Test Statistics

We introduce a device -the generalized autocontour- as the basis to construct statistical tests for the null hypothesis of a well-specified conditional density model either univariate or multivariate.

2.1 Generalized Autocontour: G-ACR

Following Diebold, Gunther, and Tay (1998) among others, if the proposed predictive density model for Y_t , i.e. $\{f_t^*(y_t|\Omega_{t-1})\}_{t=1}^T$ coincides with the true conditional density $\{f_t(y_t|\Omega_{t-1})\}_{t=1}^T$, then the sequence of probability integral transforms (PIT) of $\{Y_t\}_{t=1}^T$ w.r.t $\{f_t^*(y_t|\Omega_{t-1})\}_{t=1}^T$ i.e. $\{u_t\}_{t=1}^T$ must be *i.i.d* $U(0,1)$ where $u_t = \int_{-\infty}^{y_t} f_t^*(v_t|\Omega_{t-1})dv_t$. Thus, the null hypothesis $H_0 : f_t^*(y_t|\Omega_{t-1}) = f_t(y_t|\Omega_{t-1})$ is equivalent to the null hypothesis $H'_0 : \{u_t\}_{t=1}^T$ is *i.i.d* $U(0,1)$.

We construct the G-ACR under i.i.d. uniformity for univariate and multivariate predictive densities. We start with the univariate case. Within the process $\{u_t\}_{t=1}^T$, we choose any vector $(u_t, u_{t-k}) \subset \mathbb{R}^2$. Under $H'_0 : \{u_t\}_{t=1}^T$ *i.i.d* $U(0,1)$, the $G-ACR_{\alpha_i,k}$ is defined as the set of points in the plane (u_t, u_{t-k}) , i.e. $B(u_t, u_{t-k})$ such that the square with $\sqrt{\alpha_i}$ side contains $\alpha_i\%$ of observations, i.e.,

$$G-ACR_{\alpha_i,k} = \{B(u_t, u_{t-k}) \subset \mathbb{R}^2 \mid 0 \leq u_t \leq \sqrt{\alpha_i} \text{ and } 0 \leq u_{t-k} \leq \sqrt{\alpha_i}, \text{ s.t. } : u_t \times u_{t-k} \leq \alpha_i\}$$

We will call the proposed cutting of the domain ‘symmetric’ as we impose the same upper bound $\sqrt{\alpha_i}$ for each u in the vector (u_t, u_{t-k}) . However, it is possible to cut the domain in many different ways. The symmetric cutting will be very advantageous on computing the variance-covariance matrix of the test statistics that we present in the forthcoming sections, and it will also facilitate the construction of the uniform autocontours in the multivariate case, as we explain shortly.

The original autocontour proposed in González-Rivera *et. al.* (2011) was constructed for the process of (standardized) innovations of a dynamic model, say $\{\varepsilon_t\}_{t=1}^T$, with assumed conditional density $f(\cdot)$. Under the null hypothesis of i.i.d. random variables, the bivariate density function of any vector, say $(\varepsilon_t, \varepsilon_{t-k}) \subset R^2$ is $f(\varepsilon_t, \varepsilon_{t-k}) = f(\varepsilon_t)f(\varepsilon_{t-k})$, and the autocontour ACR_α was defined as the set of points in the plane $(\varepsilon_t, \varepsilon_{t-k}) \subset R^2$ that will contain $\alpha\%$ of observations when we *horizontally* slice the bivariate density function at a fixed value \bar{f}_α . Mathematically,

$$ACR_\alpha = \{B(\varepsilon_t, \varepsilon_{t-k}) \subset R^2 \mid \int_l^u \int_{l_k}^{u_k} f(\varepsilon_t)f(\varepsilon_{t-k})d\varepsilon_t d\varepsilon_{t-k} \leq \alpha\}$$

Observe that if we were to implement the same approach for the process $\{u_t\}_{t=1}^T$, under the i.i.d null hypothesis, the joint probability density is always a constant, i.e. $f(u_t, u_{t-k}) = f(u_t)f(u_{t-k}) = 1 \times 1 = 1$ and, unlike the original ACR, we cannot construct the autocontour by horizontally slicing the joint density. If we horizontally slice $f(u_t, u_{t-k})$ at any level, and project down the resulting segment on the plane (u_t, u_{t-k}) we always obtain a square of area one. For this reason, the proposed G-ACR ”cuts” the domain of the bivariate uniform density that is a square with unit side.

In the multivariate case, our interest is the modeling of an $m \times 1$ random vector $Y_t = (y_{1,t}, y_{2,t}, \dots, y_{m,t})'$, with a joint probability density $f_t(Y_t | \Omega_{t-1})$. Let us call $\{f_t^*(Y_t | \Omega_{t-1})\}_{t=1}^T$ the sequence of predicted densities. For each t , the joint density can be factorized as the product of the conditional densities and the marginal density, i.e.

$$f^*(Y_t | \Omega_{t-1}) = f^*(y_{1,t}, y_{2,t}, \dots, y_{m,t} | \Omega_{t-1}) = f^*(y_{m,t} | y_{1,t}, y_{2,t}, \dots, y_{m-1,t}) \times \dots \times f^*(y_{2,t} | y_{1,t}) \times f^*(y_{1,t} | \Omega_{t-1})$$

For each element in the factorization we obtain the sequence of PITs, $\{u_{1,t}, u_{2|1,t}, \dots, u_{m|1,2\dots m-1,t}\}_{t=1}^T$ such that $u_{1,t} = \int_{-\infty}^{y_{1,t}} f_t^*(v_t | \Omega_{t-1}) dv_t$, $u_{2|1,t} = \int_{-\infty}^{y_{2,t}} f_{y_2|y_1,t}^*(v_t | \Omega_{t-1}) dv_t$, , and $u_{m|1,2\dots m-1,t} =$

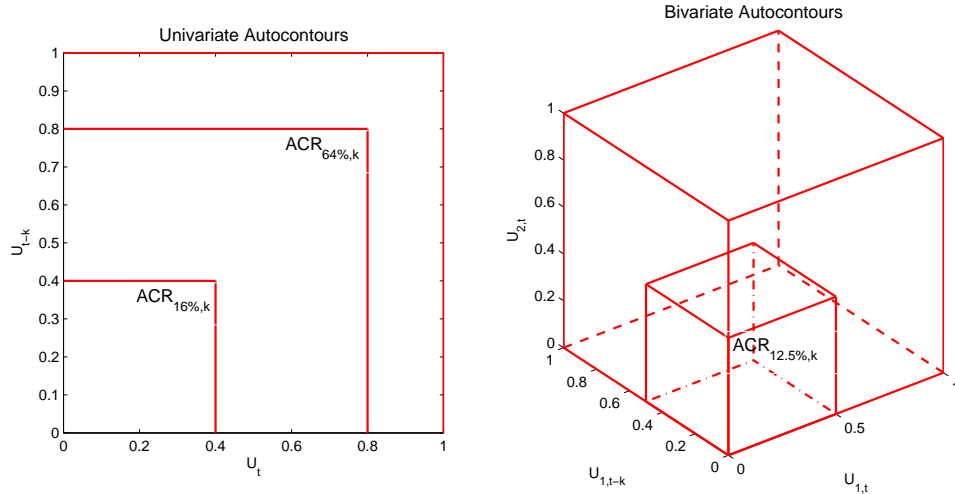
$$\int_{-\infty}^{y_{m,t}} f_{y_m|y_1,y_2,\dots,y_{m,t}}^*(v_t|\Omega_{t-1})dv_t.$$

Under the null hypothesis of a correct density model, the multivariate sequence $\{u_{1,t}, u_{2|1,t}, \dots, u_{m|1,2,\dots,m-1,t}\}_{t=1}^T$ is *i.i.d* $U(0, 1)$ (Diebold, Hahn, and Tay, 1999); thus the null hypothesis $H_0 : f_t^*(Y_t|\Omega_{t-1}) = f_t(Y_t|\Omega_{t-1})$ is equivalent to $H'_0 : \{u_{1,t}, u_{2|1,t}, \dots, u_{m|1,2,\dots,m-1,t}\}_{t=1}^T$ is *i.i.d* $U(0, 1)$.

Under this null, we construct the autocontour $G-ACR_{\alpha_i,k}$ by choosing any two m -dimensional vectors in the sequence of multivariate PITs that are k periods apart. Thus, $G-ACR_{\alpha_i,k}$ is now the set of points defined in R^{2m} such that the hyper-cube with $\alpha_i^{1/2m}$ side contains $\alpha_i\%$ of observations:

$$\begin{aligned} G-ACR_{\alpha_i,k} &= \{B(u_{1,t}, u_{2|1,t}, \dots, u_{m|1,2,\dots,m-1,t}; u_{1,t-k}, u_{2|1,t-k}, \dots, u_{m|1,2,\dots,m-1,t-k}) \subset \mathfrak{R}^{2m} \\ &\| 0 \leq u_{1,t} \leq \alpha_i^{1/2m}, 0 \leq u_{2|1,t} \leq \alpha_i^{1/2m}, \dots, 0 \leq u_{m|1,2,\dots,m-1,t} \leq \alpha_i^{1/2m} \\ &0 \leq u_{1,t-k} \leq \alpha_i^{1/2m}, 0 \leq u_{2|1,t-k} \leq \alpha_i^{1/2m}, \dots, 0 \leq u_{m|1,2,\dots,m-1,t-k} \leq \alpha_i^{1/2m} \\ &s.t. : u_{1,t} \times u_{2|1,t} \times \dots \times u_{m|1,2,\dots,m-1,t} \times u_{1,t-k} \times u_{2|1,t-k} \times \dots \times u_{m|1,2,\dots,m-1,t-k} \leq \alpha_i\} \end{aligned} \quad (2.1)$$

Figure 1: Autocontours for the Uniform Density



In Figure 1, we present 2-dimensional α_i -autocontours for the univariate case, $B(u_t, u_{t-k}) \subset \mathfrak{R}^2$ (left panel) and 3-dimensional autocontours for a bivariate case (right panel). In the latter, the

autocontour is a 4-dimensional hypercube $B(u_{1,t}, u_{2|1,t}, u_{1,t-k}, u_{2|1,t-k}) \subset \mathfrak{R}^4$, of which we can only plot obviously three dimensions.

2.2 Test Statistics

We start by defining an indicator function in the univariate as well as in the multivariate case.

In the univariate case, the indicator will take the value one whenever the observation (u_t, u_{t-k}) falls within the area defined by the $G-ACR_{\alpha_i, k}$ and zero otherwise, i.e.

$$I_t^{k, \alpha_i} = \mathbf{1}((u_t, u_{t-k}) \in G-ACR_{\alpha_i, k}) = \mathbf{1}(0 \leq u_t \leq \sqrt{\alpha_i}, 0 \leq u_{t-k} \leq \sqrt{\alpha_i})$$

In the multivariate case, the indicator will be defined similarly as follows

$$\begin{aligned} I_t^{k, \alpha_i} &= \mathbf{1}((u_{1,t}, u_{2|1,t}, \dots, u_{m|1,2\dots m-1,t}, u_{1,t-k}, u_{2|1,t-k}, \dots, u_{m|1,2\dots m-1,t-k}) \in G-ACR_{\alpha_i, k}) \\ &= \mathbf{1}(0 \leq u_{1,t} \leq \alpha_i^{1/2m}, 0 \leq u_{2|1,t} \leq \alpha_i^{1/2m}, \dots, 0 \leq u_{m|1,2\dots m-1,t} \leq \alpha_i^{1/2m}, \\ &0 \leq u_{1,t-k} \leq \alpha_i^{1/2m}, 0 \leq u_{2|1,t-k} \leq \alpha_i^{1/2m}, \dots, 0 \leq u_{m|1,2\dots m-1,t-k} \leq \alpha_i^{1/2m}) \end{aligned}$$

In both cases, I_t^{k, α_i} are Bernoulli random variables for which $\alpha_i\%$ of the observations will fall inside the $G-ACR_{\alpha_i, k}$ and $(1 - \alpha_i)\%$ observations will fall outside. Note that the indicator is also autocorrelated; it follows a MA process whose order depends on k . Thus, the moments of I_t^{k, α_i} are

$$\begin{aligned} E(I_t^{k, \alpha_i}) &= \alpha_i \\ Var(I_t^{k, \alpha_i}) &= \alpha_i(1 - \alpha_i) \\ r_h^{\alpha_i} &= cov(I_t^{k, \alpha_i}, I_{t-h}^{k, \alpha_i}) = \begin{cases} 0 & \text{if } h \neq k \\ \alpha_i^{3/2}(1 - \alpha_i^{1/2}) & \text{if } h = k \end{cases} \end{aligned}$$

The indicator I_t^{k, α_i} forms the basis of the following test statistics, which are applicable in the univariate as well as in the multivariate case. The next three propositions follow closely the tests derived in González-Rivera *et. al.* (2011). However, the forthcoming tests offer a great advantage over the previous ones because their asymptotic variance-covariance matrices have closed-form formulations which depend exclusively on the theoretical probability α_i of the autocontour specified

under the null. The proofs of the following propositions are all relegated to the Appendix.

Proposition 1

Let $\hat{\alpha}_i = \frac{\sum_{t=k+1}^T I_t^{k, \alpha_i}}{T-k}$ be the sample proportion. Under the null hypothesis of *i.i.d* $U(0, 1)$ PITs, the sample proportion $\hat{\alpha}_i$ is asymptotically normally distributed, i.e.,

$$\frac{\sqrt{T-k}(\hat{\alpha}_i - \alpha_i)}{\sigma_{k, \alpha_i}} \rightarrow N(0, 1)$$

where

$$\sigma_{k, \alpha_i}^2 = \alpha_i(1 - \alpha_i) + 2\alpha_i^{3/2}(1 - \alpha_i^{1/2})$$

In Proposition 1, the lag k and the autocontour α_i are fixed. By letting k and i run through many values, we can construct portmanteau statistics as those in the following propositions.

Proposition 2

For a given contour α_i , let $\ell_{k, \alpha_i} = \sqrt{T-k}(\hat{\alpha}_i - \alpha_i)$ and stack ℓ_{k, α_i} for $k = 1, \dots, K$. Let $L_{\alpha_i} = (\ell_{1, \alpha_i}, \dots, \ell_{K, \alpha_i})'$ be the $K \times 1$ stacked vector. Under the null hypothesis of *i.i.d* $U(0, 1)$ PITs, the asymptotic distribution of the vector L_{α_i} is multivariate normal, i.e., $L_{\alpha_i} \rightarrow N(0, \Lambda_{\alpha_i})$ and the following quadratic form follows asymptotically a chi-square with K degrees of freedom, i.e.,

$$L_{\alpha_i}' \Lambda_{\alpha_i}^{-1} L_{\alpha_i} \rightarrow \chi_K^2$$

where a typical element of the asymptotic covariance matrix Λ_{α_i} , say $\lambda_{j,k}$ is as follows

$$\lambda_{j,k} = \begin{cases} \alpha_i(1 - \alpha_i) + 2\alpha_i^{3/2}(1 - \alpha_i^{1/2}) & \text{if } j = k \\ 4\alpha_i^{3/2}(1 - \alpha_i^{1/2}) & \text{if } j \neq k \end{cases}$$

Proposition 3

For a given lag k , let $c_{k,i} = \sqrt{T-k}(\hat{\alpha}_i - \alpha_i)$ and stack $c_{k,i}$ for different contours levels $i = 1, 2, \dots, C$. Let $\mathbf{C}_k = (c_{k,1}, \dots, c_{k,C})'$ be the $C \times 1$ stacked vector. Under the null hypothesis of *i.i.d* $U(0, 1)$ PITs, the asymptotic distribution of the vector \mathbf{C}_k is multivariate normal, i.e., $\mathbf{C}_k \rightarrow$

$N(0, \Omega_k)$ and consequently, the following quadratic form asymptotically follows a chi-square with C degrees of freedom, i.e.,

$$\mathbf{C}'_k \Omega_k^{-1} \mathbf{C}_k \rightarrow \chi_C^2$$

where a typical element of the asymptotic covariance matrix Ω_k , say $\omega_{i,j}$ is as follows

$$\omega_{i,j} = \begin{cases} \alpha_i(1 - \alpha_i) + 2\alpha_i^{3/2}(1 - \alpha_i^{1/2}) & \text{if } i = j \\ \alpha_i(1 - \alpha_j) + 2\alpha_i\alpha_j^{1/2}(1 - \alpha_j^{1/2}) & \text{if } i < j \\ \alpha_j(1 - \alpha_i) + 2\alpha_j\alpha_i^{1/2}(1 - \alpha_i^{1/2}) & \text{if } i > j \end{cases}$$

2.3 Implementation of Test Statistics for Copula Models

Copula models are useful when we know the marginal distributions of the processes in a random vector but we do not know the multivariate density function. For instance, in a vector of durations, some process may be distributed exponential and some other Weibull but if we wish to model the bivariate system, we need to complete the model by assuming a bivariate distribution. In these instances, the choice of a copula function to tie up the assumed marginal densities will provide a full characterization of the model. Next, we explain how to implement the proposed G-ACR-based tests under Gaussian and Clayton copulas. In the empirical sections, we also entertain Gumbel.

2.3.1 Multivariate Distribution of a Vector with Gaussian Copula

Let (N_1, \dots, N_n) be a random vector, with marginal distributions $F_i(N_i)$. We specify the joint distribution $H(N_{1,t}, \dots, N_{n,t})$ of the vector by choosing a copula function C such that $H(N_{1,t}, \dots, N_{n,t}) = C(F_1(N_{1,t}), \dots, F_n(N_{n,t}))$. For a Gaussian copula, the multivariate distribution is

$$H(N_{1,t}, \dots, N_{n,t}) = \Phi^n(\Phi^{-1}(F_1(N_{1,t})), \dots, \Phi^{-1}(F_n(N_{n,t})); \Sigma)$$

with Φ^n as the n-dimensional multivariate standard normal with correlation matrix Σ .

Let us call $q_{i,t} = \Phi^{-1}(F_i(N_{i,t}))$, where Φ is the univariate standard normal distribution. Then $H(N_{1,t}, \dots, N_{n,t}) = \Phi^n(q_{1,t}, \dots, q_{n,t}; \Sigma)$ with corresponding density $f(q_{1,t}, \dots, q_{n,t}) = \phi^n(q_{1,t}, \dots, q_{n,t}; \Sigma)$, which is a multivariate standard normal density with correlation matrix Σ . Suppose that we wish

to evaluate a predictive density $\{h_t^*(N_t|\Omega_{t-1})\}_{t=1}^T$, generated by a Gaussian copula, for the vector $(N_{1,t}, \dots, N_{n,t})$. To test the null hypothesis that $\{h_t^*(N_t|\Omega_{t-1})\}_{t=1}^T$ is the correct density model is equivalent to test that the correct density for $(q_{1,t}, \dots, q_{n,t})$ is multivariate normal, i.e.,

$$H_0 : \{f_t^*(q_{1,t}, \dots, q_{n,t}|\Omega_{t-1}) = \phi^n(q_{1,t}, \dots, q_{n,t}|\Omega_{t-1}; \Sigma)\}_{t=1}^T$$

Therefore, we will proceed as follows,

1. Obtain the PIT for each element of the random vector $(N_{1,t}, \dots, N_{n,t})$, i.e., $u_{1,t} = F_1(N_{1,t}), \dots, u_{n,t} = F_n(N_{n,t})$, where F_i is the marginal CDF of the random variable $N_{i,t}$.
2. Obtain $q_{i,t}$ for each $N_{i,t}$ by transforming the PITs as $q_{i,t} = \Phi^{-1}(F_i(N_{i,t}))$ where Φ is the univariate standard normal distribution.
3. Factorize the multivariate density of the random vector $Q_t = (q_{1,t}, \dots, q_{n,t})$ as the product of conditional and marginal densities as

$$f(q_{1,t}, q_{2,t}, \dots, q_{n,t}|\Omega_{t-1}) = f(q_{n,t}|q_{1,t}, q_{2,t}, \dots, q_{n-1,t}) \times \dots \times f(q_{2,t}|q_{1,t}) \times f(q_{1,t}|\Omega_{t-1})$$

For a Gaussian copula, $f_t(q_{1,t}, \dots, q_{n,t}|\Omega_{t-1}) = \phi^n(q_{1,t}, \dots, q_{n,t}|\Omega_{t-1}; \Sigma)$, the marginal and conditional densities are also normal. Thus, we can now easily obtain the corresponding PITs, i.e., $u_{1,t} = \int_{-\infty}^{q_{1,t}} f_{q_{1,t}}(v_t|\Omega_{t-1})dv_t, \dots, u_{m|1,2,\dots,m-1,t} = \int_{-\infty}^{q_{m,t}} f_{q_m|q_1,\dots,q_{m-1,t}}(v_t|\Omega_{t-1})dv_t, \dots$, which under the null hypothesis of a correctly specified density model, must be i.i.d. $U[0,1]$ random variables.

2.3.2 Multivariate Distribution of a Vector with Clayton Copula

For the Clayton copula, the multivariate distribution of the vector (N_1, \dots, N_n) is given by

$$\begin{aligned} H(N_{1,t}, \dots, N_{n,t}) &= C(F_1(N_{1,t}), \dots, F_n(N_{n,t})) = (1 - n + \sum_{i=1}^n (F_i(N_{i,t}))^{-\theta})^{-1/\theta} \\ &= (1 - n + \sum_{i=1}^n (q_{i,t})^{-\theta})^{-1/\theta} \end{aligned}$$

where $q_{i,t} = F_i(N_{i,t})$, and the corresponding density is given by

$$c(q_{1,t}, \dots, q_{n,t}) = (1 - n + \sum_{i=1}^n q_{i,t}^{-\theta})^{-n-\frac{1}{\theta}} \prod_{i=1}^n (q_{i,t}^{-\theta-1} \{(i-1)\theta + 1\})$$

To obtain the PITs under the null hypothesis of a correct model under Clayton, proceed as follows

1. Obtain the PIT for each element of the random vector $(N_{1,t}, \dots, N_{n,t})$, i.e., $q_{1,t} = F_1(N_{1,t}), \dots, q_{n,t} = F_n(N_{n,t})$, where F_i is the marginal CDF of the random variable $N_{i,t}$.
2. Factorize the multivariate density of the random vector $Q_t = (q_{1,t}, \dots, q_{n,t})$ as the product of conditional and marginal densities as

$$c(q_{1,t}, q_{2,t}, \dots, q_{n,t} | \Omega_{t-1}) = c(q_{n,t} | q_{1,t}, q_{2,t}, \dots, q_{n-1,t}) \times \dots \times c(q_{2,t} | q_{1,t}) \times c(q_{1,t} | \Omega_{t-1})$$

In order to calculate the conditional densities, we exploit the property that all n-marginal distributions of a Clayton copula are identical, i.e., $C(F_1(N_{1,t}), \dots, F_n(N_{n-1,t}), 1) = (1 - (n-1) + \sum_{i=1}^{n-1} (q_{i,t})^{-\theta})^{-1/\theta}$. With the marginal and conditional distributions in place, we obtain the PITs as previously described, i.e., $u_{1,t} = \int_{-\infty}^{q_{1,t}} c_{q_{1,t}}(v_t | \Omega_{t-1}) dv_t, \dots, u_{m|1,2,\dots,m-1,t} = \int_{-\infty}^{q_{m,t}} c_{q_m | q_1, \dots, q_{m-1,t}}(v_t | \Omega_{t-1}) dv_t$, and proceed with the implementation of the autocontour-based tests.

3 Monte Carlo Simulations

We perform extensive Monte Carlo simulations to assess the finite sample properties of the tests in Propositions 1 to 3. We consider a bivariate vector $N_t = (N_{1,t}, N_{2,t})'$ under different multivariate distributions functions. Previous works dealing with financial data have considered multivariate normal and multivariate Student-t extensively. Here we depart from these popular distributions and simulate data from densities like Poisson and Negative Binomial that have not been considered as widely as the others. Though Poisson and Negative Binomial characterize discrete random variables, e.g. counts, we simulate data from processes that will generate a large number of counts

so that the data can be considered continuous and the proposed tests are readily applicable. ¹

3.1 Size of the tests

We consider five models for which the conditional mean of counts, $\mu_{i,t} \equiv E[N_{i,t}|\Omega_{t-1}]$ for $i = 1, 2$, obeys dynamics of order one, i.e.,

$$E(N_t|\Omega_{t-1}) = W + A \times N_{t-1} + B \times \mu_{t-1} \quad (3.1)$$

where Ω_{t-1} is the information set, and W is a vector and A and B are matrices of parameters with the following values: $W = \begin{pmatrix} 35 \\ 35 \end{pmatrix}$ for Model S1; $W = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$ for Models S2 to S5; $A = \begin{pmatrix} 0.4 & 0 \\ 0.15 & 0.45 \end{pmatrix}$, $B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.45 \end{pmatrix}$.

The difference among the following five models lies on the assumed marginal densities and whether the elements of the vector are contemporaneously correlated or not. We entertain the following density specifications:

Model S1: Bivariate Conditional Autoregressive Poisson (P) Model ²

$$\begin{aligned} N_{i,t}|\Omega_{t-1} &\sim P(\mu_{i,t}), \text{ for } i = 1, 2 \\ f(n_{t,i}|\Omega_{t-1}) &= \frac{e^{-\mu_{i,t}} \mu_{i,t}^{n_{t,i}}}{n_{t,i}!} \\ E[N_{i,t}|\Omega_{t-1}] &= \mu_{i,t} \\ Var[N_{i,t}|\Omega_{t-1}] &= \mu_{i,t} \end{aligned}$$

In empirical financial applications, we observe overdispersion in the data very frequently. A Poisson model does not allow for overdispersion because the mean and the variance are the same. For this reason, we consider next Negative Binomial marginal densities.

¹When the number of counts is very large, a Poisson density converges to a Normal density. The Poisson density is also a limiting case of the Negative Binomial.

²This DGP generates an average (over the time series) number of counts of 349 for N_{1t} and of 874 for N_{2t} . These numbers are large enough to consider the data continuous.

Model S2: Bivariate Conditional Autoregressive Negative Binomial (NB) Model ($\kappa_i = 4$)³

$$\begin{aligned}
 N_{i,t}|\Omega_{t-1} &\sim NB(\mu_{i,t}), \text{ for } i = 1, 2 \\
 f(n_{t,i}|\Omega_{t-1}) &= \frac{\Gamma(n_{t,i} + \kappa_i)}{\Gamma(\kappa_i)\Gamma(n_{t,i} + 1)} \left(\frac{\kappa_i}{\kappa_i + \mu_{i,t}}\right)^{\kappa_i} \left(\frac{\mu_{i,t}}{\mu_{i,t} + \kappa_i}\right)^{n_{t,i}} \\
 E[N_{i,t}|\Omega_{t-1}] &= \mu_{i,t} \\
 Var[N_{i,t}|\Omega_{t-1}] &= \mu_{i,t}\left(1 + \frac{\mu_{i,t}}{\kappa_i}\right)
 \end{aligned}$$

Because $\kappa_i > 0$ and $\mu_{i,t} > 0$, the conditional variance is greater than the conditional mean, thus the model generates overdispersion in the data (Cameron and Trivedi, 2005). When $1/\kappa_i \rightarrow 0$, the Negative Binomial converges to the Poisson density.

The three following models allow for contemporaneous dependence among the elements of the vector. The dependence is modeled by a copula function. We consider three one-parameter Archimedean copulas (Nelsen, 2005): a Normal copula for which the dependence is summarized by the correlation coefficient, and the Clayton and Gumbel copulas that allow for asymmetric dependence among the elements of the vector.

Model S3: Bivariate Conditional Autoregressive Negative Binomial with Gaussian Copula Model (correlation coefficient $\rho=0.8$).

This model has the same marginal distributions as in Model S2 but the contemporaneous cross-correlation between the vector components is generated by using the Gaussian copula, which functional form is described in section 2.3.1.

Model S4: Bivariate Conditional Autoregressive Negative Binomial with Clayton Copula Model (copula parameter $\theta = 2$).

This model has the same marginal distributions as in Model S2 but the contemporaneous cross-correlation between the vector components is generated by using the Clayton copula, which allows for stronger correlation at low values than at high values of the data. The functional form of the Clayton copula is described in section 2.3.2.

Model S5: Bivariate Conditional Autoregressive Negative Binomial with Gumbel Copula

³Models S2 to S5 generate an average (over the time series) number of counts of 50 for N_{1t} and of 123 for N_{2t} , which are large enough for the data to be considered continuous.

Model (copula parameter $\theta = 2$).

This model has the same marginal distributions as in Model S2 but the contemporaneous cross-correlation between the vector components is generated by the Gumbel copula, which allows for stronger correlation at high values than at low values of the data. Gumbel functional form is

$$C(u_1, u_2, \dots, u_n) = \exp(- [(-\ln u_1)^\theta + (-\ln u_2)^\theta + \dots + (-\ln u_n)^\theta]^{1/\theta}), \theta > 0$$

We have estimated the five models S1 to S5 by maximum likelihood. The log-likelihood function is constructed according to the distributional assumptions specified in each model. For models with a copula function we follow a two-stage estimation procedure as in Patton (2006) by first estimating the parameters in the marginal model, and secondly using these estimates to estimate the copula parameter. Since parameter uncertainty is more important in-sample testing, we conduct the experiments in sample. We obtain the PITs associated with the one-step-ahead conditional expectation of the count vector following the procedures explained in section 2.3, and proceed to implement the tests described in Propositions 1 to 3. We have implemented a parametric bootstrap procedure to approximate the asymptotic variance of the tests.⁴ We obtain bootstrap samples for each model by replacing the true value θ_0 with the the estimate $\hat{\theta}_T$, computing the conditional mean $\mu_{i,t}$, and making draws $N_{i,t}$ from the specified parametric distributions. This is a standard procedure to overcome the difficulties of estimating asymptotic variances when parameter uncertainty is relevant. The following experiments will show that bootstrapping the variance of the tests and using standard asymptotic critical values provides statistics with the right size.

In Tables 1 and 2, we show the size of the tests for different sample sizes $T = 250, 500$ and 1000 observations. The overall performance of the tests is very good. Across models and across sample sizes, the average empirical size is 5%. We do not observe any instance in which the tests are grossly over- or undersized. In Table 1, we also include the size of the test for model S1 without bootstrapping the variance of the test. As we expected, the size is very distorted and the tests

⁴When testing in-sample specification, ignoring parameter uncertainty may cause severe distortions in the size of the tests. When testing out-of-sample specification, the importance of parameter uncertainty will depend on the forecasting scheme and on the size of the estimation sample (R) relative to the prediction sample (P). Under the assumption of \sqrt{R} -consistent estimators, if $R \rightarrow \infty$, $P \rightarrow \infty$, and $P/R \rightarrow 0$ as $T \rightarrow \infty$, parameter uncertainty is asymptotically negligible and no adjustment is needed in the tests provided in Propositions 1 to 3.

are all uniformly under-sized so that the tests do not reject the null hypothesis as much as they should. These results support the practice of implementing a bootstrap procedure when parameter uncertainty is a concern.

Table 1: **Size of t_{k,α_i} -statistics for 13 autocontours (Nominal size 5% and $k = 1$)**

	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
T	Model S1												
250	0.031	0.033	0.042	0.067	0.066	0.065	0.067	0.063	0.043	0.052	0.063	0.043	0.068
500	0.035	0.038	0.048	0.056	0.065	0.063	0.061	0.061	0.057	0.053	0.057	0.057	0.061
1000	0.042	0.059	0.049	0.054	0.06	0.06	0.059	0.055	0.057	0.047	0.047	0.055	0.049
T	Model S1 (no bootstrapped variance)												
250	0.034	0.021	0.012	0.006	0.006	0.014	0.017	0.015	0.02	0.023	0.021	0.034	0.032
500	0.037	0.023	0.012	0.008	0.007	0.018	0.022	0.015	0.024	0.026	0.024	0.035	0.034
1000	0.033	0.024	0.018	0.007	0.009	0.016	0.019	0.018	0.023	0.029	0.023	0.04	0.035
T	Model S2												
250	0.036	0.037	0.065	0.063	0.043	0.064	0.047	0.059	0.067	0.047	0.067	0.037	0.068
500	0.041	0.045	0.042	0.049	0.058	0.061	0.049	0.062	0.063	0.056	0.065	0.054	0.043
1000	0.046	0.052	0.051	0.05	0.051	0.052	0.054	0.05	0.06	0.055	0.056	0.052	0.053
T	Model S3												
250	0.06	0.047	0.045	0.045	0.036	0.045	0.04	0.05	0.05	0.059	0.063	0.037	0.069
500	0.053	0.046	0.055	0.045	0.047	0.042	0.043	0.042	0.047	0.056	0.059	0.053	0.059
1000	0.052	0.045	0.049	0.045	0.048	0.049	0.046	0.046	0.043	0.043	0.042	0.047	0.049
T	Model S4												
250	0.037	0.037	0.044	0.049	0.062	0.066	0.043	0.04	0.046	0.043	0.066	0.039	0.045
500	0.056	0.04	0.04	0.041	0.05	0.053	0.047	0.046	0.044	0.045	0.049	0.04	0.056
1000	0.052	0.047	0.042	0.043	0.053	0.045	0.049	0.042	0.056	0.053	0.054	0.055	0.047
T	Model S5												
250	0.048	0.038	0.044	0.046	0.066	0.06	0.044	0.056	0.042	0.061	0.055	0.034	0.064
500	0.043	0.04	0.052	0.049	0.043	0.046	0.045	0.053	0.058	0.058	0.056	0.047	0.044
1000	0.045	0.059	0.054	0.053	0.051	0.057	0.052	0.05	0.051	0.054	0.047	0.052	0.045

Notes: The 13 autocontours are $C = [0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99]$. 1000 Monte Carlo replications and 500 bootstrap samples.

Table 2: **Size of t_{k,α_i} , L_{α_i} , C_k statistics (Nominal size 5%)**

	$t_{1,7}$	$t_{2,7}$	$t_{3,7}$	$t_{4,7}$	$t_{5,7}$	$L_{2,7}$	$L_{3,7}$	$L_{4,7}$	$L_{5,7}$	$C_{1,13}$
T	Model S1									
250	0.067	0.047	0.042	0.046	0.049	0.053	0.058	0.058	0.067	0.064
500	0.061	0.056	0.058	0.061	0.056	0.059	0.062	0.056	0.058	0.049
1000	0.059	0.046	0.055	0.054	0.052	0.051	0.055	0.055	0.058	0.055
T	Model S2									
250	0.047	0.041	0.038	0.043	0.041	0.055	0.048	0.05	0.052	0.054
500	0.049	0.062	0.059	0.058	0.063	0.052	0.063	0.054	0.053	0.05
1000	0.054	0.057	0.051	0.055	0.057	0.05	0.053	0.052	0.056	0.045
T	Model S3									
250	0.04	0.039	0.043	0.045	0.049	0.056	0.055	0.054	0.054	0.048
500	0.043	0.047	0.047	0.05	0.049	0.046	0.053	0.044	0.051	0.054
1000	0.046	0.047	0.046	0.043	0.053	0.056	0.041	0.05	0.048	0.047
T	Model S4									
250	0.043	0.049	0.046	0.048	0.049	0.051	0.054	0.052	0.058	0.059
500	0.047	0.04	0.047	0.047	0.05	0.049	0.046	0.047	0.049	0.059
1000	0.049	0.041	0.044	0.041	0.041	0.04	0.046	0.046	0.045	0.053
T	Model S5									
250	0.044	0.046	0.049	0.052	0.045	0.043	0.063	0.058	0.058	0.046
500	0.045	0.053	0.056	0.053	0.05	0.054	0.063	0.056	0.052	0.05
1000	0.052	0.052	0.055	0.049	0.053	0.052	0.055	0.053	0.051	0.048

Notes: $t_{k,7}$ for $k = 1, 2, \dots, 5$, and 7 refers to the 50% autocontour.

$L_{k,7}$ for $k = 2, \dots, 5$ stacking lags up to k and considering the 50% autocontour.

$C_{1,13}$ stacking all 13 autocontours for lag $k = 1$.

1000 Monte Carlo replications and 500 bootstrap samples.

3.2 Power of the tests

To study the power properties of the tests we consider as a null hypothesis Model S3: a bivariate autoregressive model of order one as in (3.1) for the vector of conditional means, with negative binomial marginal densities ($\kappa_i = 4$) and a normal copula function with contemporaneous correlation $\rho = 0.8$. We consider four data generating mechanisms:

Model P1: Conditional Autoregressive Poisson Model with Normal Copula ($\rho = 0.8$). In this case, we maintain the same dynamic structure with parameters $W = \begin{pmatrix} 35 \\ 35 \end{pmatrix}$, $A = \begin{pmatrix} 0.4 & 0 \\ 0.15 & 0.45 \end{pmatrix}$,

$B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.45 \end{pmatrix}$, and the same copula characteristics. We study departures from the hypothesized marginal densities.

Model P2: Conditional Autoregressive Negative Binomial Model with a Clayton copula ($\theta = 2$). We maintain the same dynamics and marginal densities and study departures from the hypothesized copula, and in particular, detecting asymmetric contemporaneous dependence.

Model P3: Conditional Autoregressive Negative Binomial Model with a Gumbel copula ($\theta = 2$). This case is similar to the previous one but the contemporaneous asymmetric dependence in Gumbel runs in opposite direction to that of Clayton.

Model P4: Conditional Autoregressive Negative Model with Normal Copula ($\rho = 0.8$) with high order dynamics in the conditional means, i.e.,

$$E(N_t|\Omega_{t-1}) = D + A_1 \times N_{t-1} + A_2 \times N_{t-2} + A_3 \times N_{t-3} + A_4 \times N_{t-4} + B \times \mu_{t-1}$$

where $D = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$, $A_1 = \begin{pmatrix} 0.45 & 0 \\ 0.25 & 0.35 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0.09 & 0 \\ 0.045 & 0.06 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0.03 & 0 \\ 0.01 & 0.02 \end{pmatrix}$, $A_4 = \begin{pmatrix} 0.005 & 0 \\ 0.006 & 0.003 \end{pmatrix}$, and $B = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.38 \end{pmatrix}$. We study departures from the hypothesized dynamics and maintain the distributional assumptions on the marginal densities and the copula.

We present the power results in Tables 3 and 4. In Table 3, we analyze the behavior of the t-statistics, t_{k,α_i} for a fixed lag $k = 1$ and 13 autocontours that span the entire uniform density. The Poisson case, model P1, is very easy to detect because of the property of overdispersion generated by the Negative Binomial densities. The power is one regardless of sample size not only for the t-statistics but also for the portmanteau tests L_{α_i} and C_k reported in Table 4. In general, the tests are more powerful at detecting departures from distributional assumptions than at detecting misspecified dynamics. In models P2 and P3, we find very high power even with small sample sizes. In model P4 we need a large sample (above 500 observations) to find power above 50%. At the lowest autocontours, i.e. 1, 5, or 10 %, and the highest autocontour 99%, and mainly for small samples ($T = 250$), the power is lower because there are only a few (or a

lot) observations within each autocontour and, consequently there is not much variability in the indicator function, which is at the core of the tests. The power for model P2 (Clayton) is lower than that for model P3 (Gumbel) for the intermediate autocontours around the 10- 40% level, which is expected as the correlation of the observations in this range for a Clayton copula is closer to that of the hypothesized Normal than the correlation of those observations generated by Gumbel copula. The opposite happens when we examine the high autocontours 50-99% precisely because of the opposite reason, the Gumbel correlation for observations in the upper contours is closer to the Normal correlation than that generated by the Clayton.

Table 3: **Power of t_{k,α_i} -statistics for 13 autocontours (Nominal size 5% and $k = 1$)**

	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
T	Model P1 (Poisson/Normal)												
250	0.53	1	0.344	1	1	1	1	1	1	1	1	1	1
500	0.999	1	0.44	1	1	1	1	1	1	1	1	1	1
1000	1	1	0.62	1	1	1	1	1	1	1	1	1	1
T	Model P2 (Negative Bin/Clayton)												
250	0.07	0.075	0.046	0.143	0.186	0.289	0.455	0.515	0.713	0.84	0.86	0.857	0.869
500	0.095	0.086	0.065	0.178	0.204	0.474	0.686	0.87	0.925	0.951	0.952	0.984	0.982
1000	0.179	0.283	0.218	0.385	0.548	0.661	0.922	0.99	0.992	0.99	0.997	1	1
T	Model P3 (Negative Bin/Gumbel)												
250	0.021	0.17	0.188	0.245	0.247	0.364	0.421	0.461	0.478	0.521	0.395	0.444	0.366
500	0.045	0.211	0.239	0.381	0.451	0.593	0.502	0.61	0.634	0.649	0.733	0.632	0.594
1000	0.085	0.342	0.416	0.517	0.627	0.787	0.692	0.805	0.827	0.859	0.924	0.918	0.854
T	Model P4 (Dynamic Misspecification)												
250	0.045	0.056	0.075	0.073	0.134	0.172	0.152	0.252	0.35	0.375	0.35	0.33	0.39
500	0.053	0.06	0.083	0.105	0.241	0.39	0.33	0.424	0.418	0.421	0.43	0.37	0.42
1000	0.063	0.072	0.169	0.225	0.39	0.465	0.42	0.53	0.516	0.524	0.531	0.543	0.495

Notes: The 13 autocontours are $C = [0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99]$. 1000 Monte Carlo replications and 500 bootstrap samples.

The null hypothesis is a bivariate conditional autoregressive model with Negative Binomial marginal densities and a Normal copula function.

In Table 4, we present the power of the t-statistics, $t_{k,7}$, and of the portmanteau statistics L_7^k and C_k^{13} for several values of k . In general, the power is excellent across models and sample sizes. The overall findings are similar to those from Table 3. The tests C_k^{13} are obviously the

most powerful as they collect information for all the autocontours. However, if the researcher is interested in partial aspects of the densities, such as a quantile or a collection of quantiles, it would be more informative examining the $L_{\alpha_i}^k$ statistics, where we fix the level of the quantile α_i , and the individual t-statistics that provide information about the quantile desired.

Table 4: **Power of t_{k,α_i} , L_{α_i} , C_k statistics (Nominal size 5%)**

	$t_{2,7}$	$t_{3,7}$	$t_{4,7}$	$t_{5,7}$	L_2^7	L_3^7	L_4^7	L_5^7	C_1^{13}	C_2^{13}	C_3^{13}	C_4^{13}	C_5^{13}
T	Model P1 (Poisson/Normal)												
250	1	1	1	1	1	1	1	1	1	1	1	1	1
500	1	1	1	1	1	1	1	1	1	1	1	1	1
1000	1	1	1	1	1	1	1	1	1	1	1	1	1
T	Model P2 (Negative Bin/Clayton)												
250	0.485	0.477	0.468	0.472	0.497	0.405	0.477	0.374	0.875	0.967	0.916	0.868	0.921
500	0.698	0.696	0.689	0.692	0.752	0.748	0.769	0.852	0.969	0.974	0.962	0.964	0.973
1000	0.905	0.898	0.875	0.894	0.883	0.849	0.813	0.789	1	0.999	0.999	1	1
T	Model P3 (Negative Bin/Gumbel)												
250	0.435	0.449	0.431	0.442	0.331	0.34	0.313	0.331	0.736	0.728	0.727	0.723	0.705
500	0.505	0.514	0.607	0.606	0.545	0.531	0.517	0.509	0.912	0.887	0.892	0.889	0.898
1000	0.717	0.693	0.7	0.697	0.679	0.685	0.695	0.692	1	1	0.99	0.99	0.994
T	Model P4 (Dynamic Misspecification)												
250	0.156	0.149	0.146	0.15	0.161	0.159	0.159	0.156	0.36	0.323	0.32	0.345	0.365
500	0.291	0.3	0.286	0.31	0.324	0.34	0.34	0.354	0.56	0.554	0.561	0.563	0.57
1000	0.4	0.42	0.432	0.412	0.451	0.429	0.445	0.467	0.71	0.692	0.692	0.72	0.734

Notes: $t_{k,7}$ for $k = 2, \dots, 5$, and 7 refers to the 50% autocontour.

L_7^k for $k = 2, \dots, 5$ stacking lags up to lag k and considering the 50% autocontour.

C_k^{13} stacking all 13 autocontours for lags $k = 1, \dots, 5$.

1000 Monte Carlo replications and 500 bootstrap samples.

The null hypothesis is a bivariate conditional autoregressive model with Negative Binomial marginal densities and a Normal copula function.

4 Empirical Illustration: Trading Activity in Large Banks

We illustrate our testing methodology and the use of visualization techniques that drives the specification and forecasting exercises. We estimate a multivariate model for the trades of three large U.S. banks: Bank of America, JP Morgan Chase, and Wells Fargo. We are interested in exploring models that produce asymmetric contemporaneous correlation among banks in times of

intense versus low trading activity. In an out-of-sample environment, we calculate the multivariate one-step-ahead density forecasts and evaluate the performance of the various models.

4.1 Data Description

We collect transaction data from the TAQ database for three U.S. large commercial banks: Bank of America (BOA), JP Morgan Chase (JPM), and Wells Fargo (WF), trading in the New York Stock Exchange, from January 3rd to June 30th, 2011 for a total of 125 trading days. We record the number of trades at the 5-minute frequency and remove any trades before 9:30 am (opening time) and after 4:00 pm (closing time) for a total of 9,750 observations per bank. In Table 5, we present the descriptive statistics of the three series. The three banks exhibit high level of trading activity with BAC leading to WFC and JPM. For each bank, the average number of trades is very large (above 1,500 trades) so we can treat them as a continuous variable. The overdispersion is evident as the variance is substantially larger than the mean. In Figure 2, we plot the histograms of the number of trades; they show evidence for overdispersion as there is a large tail to the right with most of the observations concentrated around the mean. The range in the number of trades is very wide, from minima in the 100s trades to maxima in the 20,000s. The pairwise contemporaneous correlation is high with coefficients of around 0.60, and there is substantial autocorrelation in the series as the Q-statistics show. In order to evaluate the one-step-ahead density forecasts, we split the total sample by choosing the first 8034 observations (103 trading days) as the estimation sample and the last 1716 observations (22 trading days) as the prediction sample. The evaluation of the one-step-density forecast is performed under a fixed scheme where the models are estimated only once; for every period in the prediction sample, the one-step forecast is calculated based on an information set that expands one observation at the time until the prediction sample is exhausted.

It is very common to find intra-day seasonality in high frequency data. Trading activity is intense at the the beginning and towards the end of the trading day and substantially lighter during the mid-day hours, which gives rise to a U-shaped curve of diurnal effects. We take care of these effects by estimating a set of half-hour dummies.

Figure 2: Histograms of the data

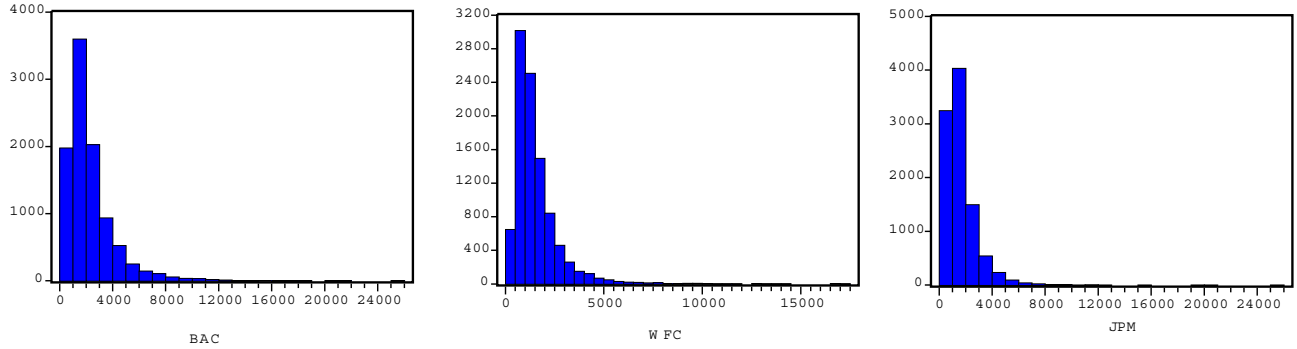


Table 5: Descriptive Statistics

	BAC	WFC	JPM
Mean	2280.03	1527.99	1636.72
Median	1757.50	1210.00	1300.00
Maximum	25409	17314	25877
Minimum	167	103	175
Std.Dev	1858.67	1204.04	1250.72
Skewness	2.82	3.74	4.12
Kurtosis	16.79	29.30	43.76
No. of trades	22230258	14897909	15957994
Q(10)	25947	27007	23489
Q(20)	29629	30799	26916
Contemporaneous correlation			
BAC	1.00	0.58	0.65
WFC	0.58	1.00	0.68
JPM	0.65	0.68	1.00

4.2 Dynamic Models and Forecast Evaluation

We estimate several dynamic models for the conditional mean $\mu_{i,t} \equiv E[N_{i,t}|\Omega_{t-1}]$ for $i = 1(BAC)$, $2(WFC)$, $3(JPM)$ under different distributional assumptions. To illustrate how our methodology works, we will present our results in a sequential fashion. First, we will show the results of the Poisson model and the Negative Binomial Model ignoring contemporaneous correlation. Though

a priori we know that the Poisson model will not fit our data because of overdispersion, it is interesting to contrast the estimation results with those of a Negative Binomial model and observe the behavior of the corresponding PITs and the generalized autocontour-based specification tests. Secondly, we will introduce contemporaneous correlation by estimating the dependence parameters under Gaussian, Clayton, and Gumbel copulas, and we will show the specification improvements by analyzing the autocontour-based specification tests and the responses of the corresponding PITs.

4.2.1 Models with no contemporaneous correlation

In Table 6, we present maximum likelihood estimates for the parameters of the best dynamic model for the conditional mean of trade counts with Poisson and Negative Binomial marginal densities and without contemporaneous correlation. We have experimented with different lag structures and, through standard specification tests on the Pearson residuals, we finally settled in a low order model such as $\mu_t \equiv E(N_t|\Omega_{t-1}) = w + A \times N_{t-1} + B \times \mu_{t-1}$ where N_t is a 3×1 vector of trades, A is a 3×3 matrix, not necessarily symmetric, with typical element $\{a_{i,j}\}$, and B is 3×3 diagonal matrix with typical element $\{b_{i,i}\}$.

For each equation in the system, the trading dynamics for each bank are mainly driven by the most recent activity of the bank itself as we can see in the magnitude of the estimates of $a_{i,i}$ and $b_{i,i}$. The effect of trading in the other institutions, i.e. estimates of $a_{i,j}$, are much smaller in magnitude although statistically significant at the conventional levels. The overall system is stationary as the eigenvalues of the matrix $A + B$ are inside the unit circle. The overdispersion parameter $1/\kappa_i$ is evidently different from zero, as we expected. The last panel of Table 6 provides some descriptive statistics of the Pearson residuals, i.e. $(N_{i,t} - \mu_{i,t})/\sigma_{i,t}$. If the model is well specified, the residuals should have mean zero and standard deviation one. This is not the case for the Poisson model because of the overdispersion in the data. The Q-statistics are now substantially smaller than the raw Q-statistics, indicating that the estimated dynamics are adequate. As expected, we find enough evidence to reject the Poisson model in favor of the Negative Binomial.

Table 6: Estimation Results. Maximum Likelihood Estimates

Parameter	Poisson Model			Negative Binomial Model		
	BAC	WFC	JPM	BAC	WFC	JPM
w	157.79 (29.69)	135.52 (23.05)	168.56 (22.32)	262.46 (25.83)	179.92 (22.74)	36.38 (8.99)
$a_{1,1}$	0.418 (71.67)			0.425 (38.92)		
$a_{1,2}$	0.0232 (4.85)			0.006 (2.00)		
$a_{1,3}$	0.078 (3.68)			0.07 (5.69)		
$a_{2,1}$		0.006 (2.13)			0.0099 (2.14)	
$a_{2,2}$		0.480 (13.37)			0.4561 (36.31)	
$a_{2,3}$		0.0329 (3.78)			0.021 (2.57)	
$a_{3,1}$			0.0273 (8.04)			0.031 (6.25)
$a_{3,2}$			0.0329 (3.05)			0.0466 (6.70)
$a_{3,3}$			0.463 (13.14)			0.431 (34.52)
$b_{1,1}$	0.431 (32.76)			0.381 (31.13)		
$b_{2,2}$		0.379 (17.78)			0.363 (22.14)	
$b_{3,3}$			0.35 (9.61)			0.456 (28.46)
Dispersion						
$1/\kappa_1$	0			0.192 (38.46)		
$1/\kappa_2$		0			0.159 (40.66)	
$1/\kappa_3$			0			0.138 (46.70)
Pearson Residuals						
mean	0.0607	0.1197	-0.075	0.014	-0.0003	0.011
std.deviation	21.02	16.12	16.91	1.48	1.20	1.34
Q(10)	10.56	68.73	137.60	5.26	29.26	17.70
Q(20)	56.85	107.50	176.19	17.11	63.91	21.49

Note: t -statistics in parenthesis. The models also include five dummies to take care of diurnal effects. The Pearson residual is defined as $(N_{i,t} - \mu_{i,t})/\sigma_{i,t}$.

Table 7: Autocontour-based Tests on One-step-ahead Density Forecasts:

Negative Binomial Model with no Contemporaneous Correlation

	t-tests					L_{α_i}			
α_i	lag 1	lag 2	lag 3	lag 4	lag 5	lag 2	lag 3	lag 4	lag 5
1%	11.87	12.87	13.04	14.70	10.89	231.24	311.43	391.74	396.22
5%	12.03	11.55	11.84	13.93	11.99	185.04	207.91	252.90	262.57
10%	12.81	12.48	11.83	12.34	12.08	195.17	204.19	215.78	218.75
20%	10.80	10.79	10.08	10.72	10.10	134.68	135.64	139.51	139.51
30%	8.28	8.59	7.72	8.20	7.29	78.58	78.59	80.13	81.08
40%	7.48	7.63	6.79	7.037	6.47	61.42	61.89	61.89	63.36
50%	5.39	5.38	4.59	5.09	4.54	30.43	32.71	32.87	34.53
60%	3.90	3.84	3.41	3.94	3.67	15.48	16.47	17.58	17.59
70%	2.37	2.58	2.07	2.44	2.44	7.73	8.47	8.76	9.05
80%	0.76	0.62	0.25	0.60	0.41	5.82	5.16	5.27	5.73
90%	-3.03	-3.51	-3.60	-3.58	-3.39	29.11	25.26	27.55	28.00
95%	-4.76	-5.03	-5.20	-5.06	-4.85	38.87	39.78	40.26	43.58
99%	-8.56	-8.72	-8.73	-8.69	-8.69	80.79	85.20	85.20	85.21
	C_k								
	339.37	333.01	330.57	383.99	318.77				

Notes: t -tests critical values: ± 1.96 (5% level), ± 2.58 (1%), ± 3.5 (0.05%).

L_{α_i} stacks lags up to 2, 3,4,5 for α_i autocontour. Critical values at 5% level:

5.99 (2 lags), 7.81 (3 lags), 9.49 (4 lags) and 11.1 (5 lags).

C_k stacks all 13 autocontours for lag k .

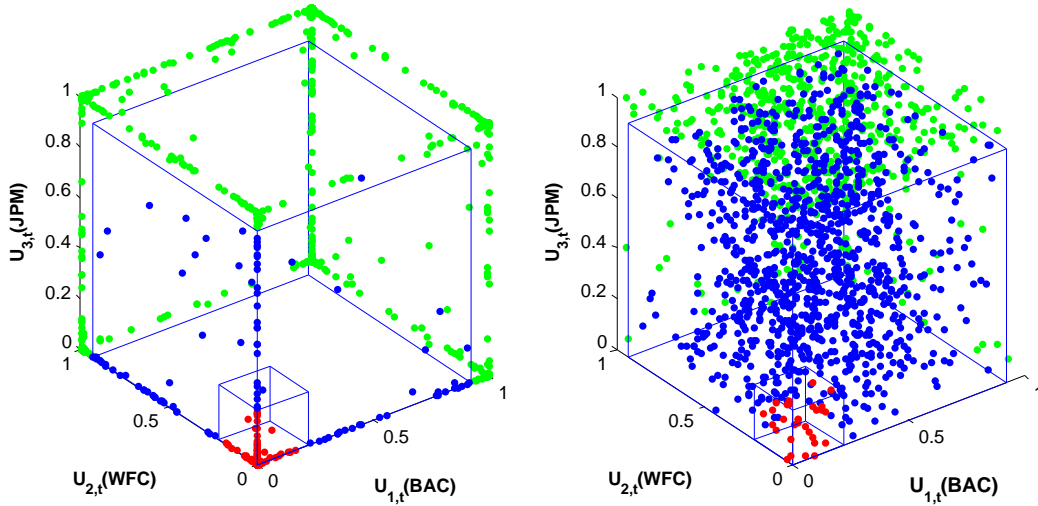
Critical values: 22.4 (5% level), 27.7 (1%), 34.5 (0.1%).

Based on these estimation results, we proceed to construct the one-step-ahead density forecasts of trades and to evaluate them with the generalized autocontour-based specification tests proposed in section 2 for the Negative Binomial model with no contemporaneous correlation. In Table 7, we present the values of the t -tests for 5 lags and 13 autocontours, the L_{α_i} tests stacking up to 2, 3, 4, and 5 lags for a given autocontour α_i , and the C_k tests stacking the 13 autocontours for a given lag $k = 1, \dots, 5$. The t -tests reject very strongly the model at the lower and middle (1 to 60%) and upper (95 and 99%) autocontour levels. At the lower levels, the test indicate that the number of observations within the autocontours is much larger than what we should expect given the assumed dynamics and contemporaneously independent Negative Binomial marginals. At the upper levels, the opposite happens, there are fewer observations than expected. However, since the t -tests show similar values across lags, the rejection of the model is not due to misspecified

dynamics but to the assumed distributional assumptions. The umbrella tests L_{α_i} and C_k , which aggregate over lags or over contours, strongly confirm the message delivered by the t -tests.

Figure 3: Contemporaneous PITs:

Poisson model (left), Negative Binomial with No Contemporaneous Correlation (right)



The graphs in Figure 3 reveal why the rejection of the Negative Binomial model with no contemporaneous correlation is so strong. We plot the 3-dimensional vector of contemporaneous PITs $\{u_{t,i}\}$ (from the prediction sample) for the Poisson model (left panel) and for the Negative Binomial model (right panel) for three autocontours, the 10% level (red dots), the 10-80% levels (blue dots), and the 80-100% levels (green dots). The rejection of Poisson is very evident because most of the points are distributed along the boundaries of the contour levels, whose location is very far from where we expect the i.i.d. uniform observations to fall. The unmodeled overdispersion drives the location of the observations towards the boundaries of the cube. The Negative Binomial model corrects this problem as the observations are more evenly distributed within the cubes associated with the levels considered. Nevertheless, we observe an elliptic concentration of the observations towards the diagonals of the cubes, which is the reason why the autocontour-based

tests reject this model and, more importantly, it points out the need to model the contemporaneous dependence among the components of the vector.

4.2.2 Models with contemporaneous correlation

We model the contemporaneous dependence of trades by using copula functions, which offer great flexibility to account for the dependence structure among several series regardless of the specification of their marginal distributions. We have selected three copula functions: Gaussian, Gumbel, and Clayton. These functions summarize the dependence of the data with one parameter. The Gaussian copula characterizes dependence through the correlation coefficients and, as such, they provide a global measure of dependence for the entire collection of observations. On the contrary, the dependence parameters in the Gumbel and Clayton copulas allow for stronger or weaker dependence in different regions of the observations. Clayton (Gumbel) produces stronger (weaker) dependence for low values and weaker (stronger) dependence for large values of the observations. We also consider the inverse Gumbel copula (also known as the survival copula), which by inverting its shape, delivers an asymmetric dependence similar to that of Clayton.⁵

The maximum likelihood estimates of the correlation matrix for the Gaussian copula are presented in Table 8. The pairwise contemporaneous correlation is significant and it runs between 0.34 and 0.43. The maximum likelihood estimates of the dependence parameter for the inverse Gumbel copula is $\theta = 1.38$ with a t-ratio of 161.78, and for the Clayton copula $\theta = 0.54$ with a t-ratio of 43.75. These values suggest very strongly that the dependence is asymmetric.

Table 8: Copula-based Correlation Matrix

	BAC	WFC	JPM
BAC	1.00	0.35	0.34
WFC	0.35	1.00	0.43
JPM	0.34	0.43	1.00

With the estimated copula and the estimated model, we calculate the one-step-ahead density forecasts, and as before, we evaluate them with the autocontour tests. In Table 9, we present

⁵We are grateful to Andrew Patton for suggesting the inverse Gumbel copula to us. If $\{u_t\}$ are the PITs from Gumbel, $\{1 - u_t\}$ will be the corresponding PITs in the inverse Gumbel copula.

the results of the autocontour-based t -tests for the Negative Binomial model with contemporaneous correlation modeled by Gaussian, inverse Gumbel, Clayton, and a mixture of Inverse Gumbel/Clayton copulas. By comparing these results with those in Table 7, we observe the substantial overall reduction in the values of the t -tests, which indicates the need to model the contemporaneous dependence among the series. Focusing on the results from the Gaussian copula, the tests fail to reject the density model in the lowest (1% and 5%) and central autocontours (50% to 80% levels). At the upper autocontours 90% to 99%, however, there is a strong rejection as the model imposes more dependence between large trades than what is granted in the data. For the upper autocontours, the Clayton copula is more responsive to the needs of the data and the values of the tests are significantly lower than those in the Gaussian case. However, the Clayton copula imposes too much dependence on the observations in the lower autocontours so that the model is also rejected. The Inverse Gumbel copula offers a much better fit than the Gaussian or the Clayton copulas; from the 1% to 90% autocontours we fail to reject the model at the conventional levels, it is only for the 90%, 95% and 99% autocontours that the t -tests from Inverse Gumbel are larger than those from Clayton.

Table 9: Autocontour-based t -tests on One-step-ahead Density Forecasts:

Negative Binomial Model with Contemporaneous Correlation

α_i	Gaussian Copula			Inv. Gumbel Copula			Clayton Copula			Inv.Gumbel/Clayton		
	lag 1	lag 2	lag 3	lag 1	lag 2	lag 3	lag 1	lag 2	lag 3	lag 1	lag 2	lag 3
1%	0.61	0.61	0.74	-0.74	-0.69	-0.75	1.91	1.85	1.86	0.43	0.49	0.48
5%	2.22	2.19	1.99	1.32	1.58	1.21	5.51	4.90	4.49	1.24	1.31	1.26
10%	3.62	3.63	3.46	2.12	2.01	1.92	6.43	6.16	5.92	2.03	2.19	2.02
20%	4.48	5.14	4.57	2.03	2.23	2.02	7.38	7.26	6.45	2.21	2.19	1.98
30%	3.89	3.61	3.76	2.24	2.12	1.85	6.31	6.24	5.67	2.44	2.38	1.93
40%	3.46	3.57	3.56	1.96	2.17	1.82	5.07	5.57	4.93	2.04	2.28	1.91
50%	2.81	2.70	2.26	1.97	2.06	1.92	4.54	4.67	4.22	2.13	2.20	2.02
60%	1.72	1.57	1.55	1.09	1.04	1.06	4.40	4.30	4.07	1.23	1.16	1.17
70%	0.38	-0.32	0.32	0.29	0.24	0.22	2.53	2.59	2.09	0.60	0.72	0.57
80%	-1.12	-1.14	-1.10	-0.34	-0.44	-0.66	1.21	1.29	1.15	0.31	0.21	-0.11
90%	-3.08	-3.59	-3.70	-2.58	-2.87	-2.75	-1.42	-1.62	-1.64	-1.55	-1.71	-1.75
95%	-5.71	-5.71	-5.67	-5.49	-5.58	-5.36	-3.79	-3.70	-4.05	-3.84	-3.92	-4.17
99%	-10.28	-10.47	-10.71	-7.10	-7.51	-7.22	-7.05	-7.09	-7.10	-6.89	-7.04	-6.98

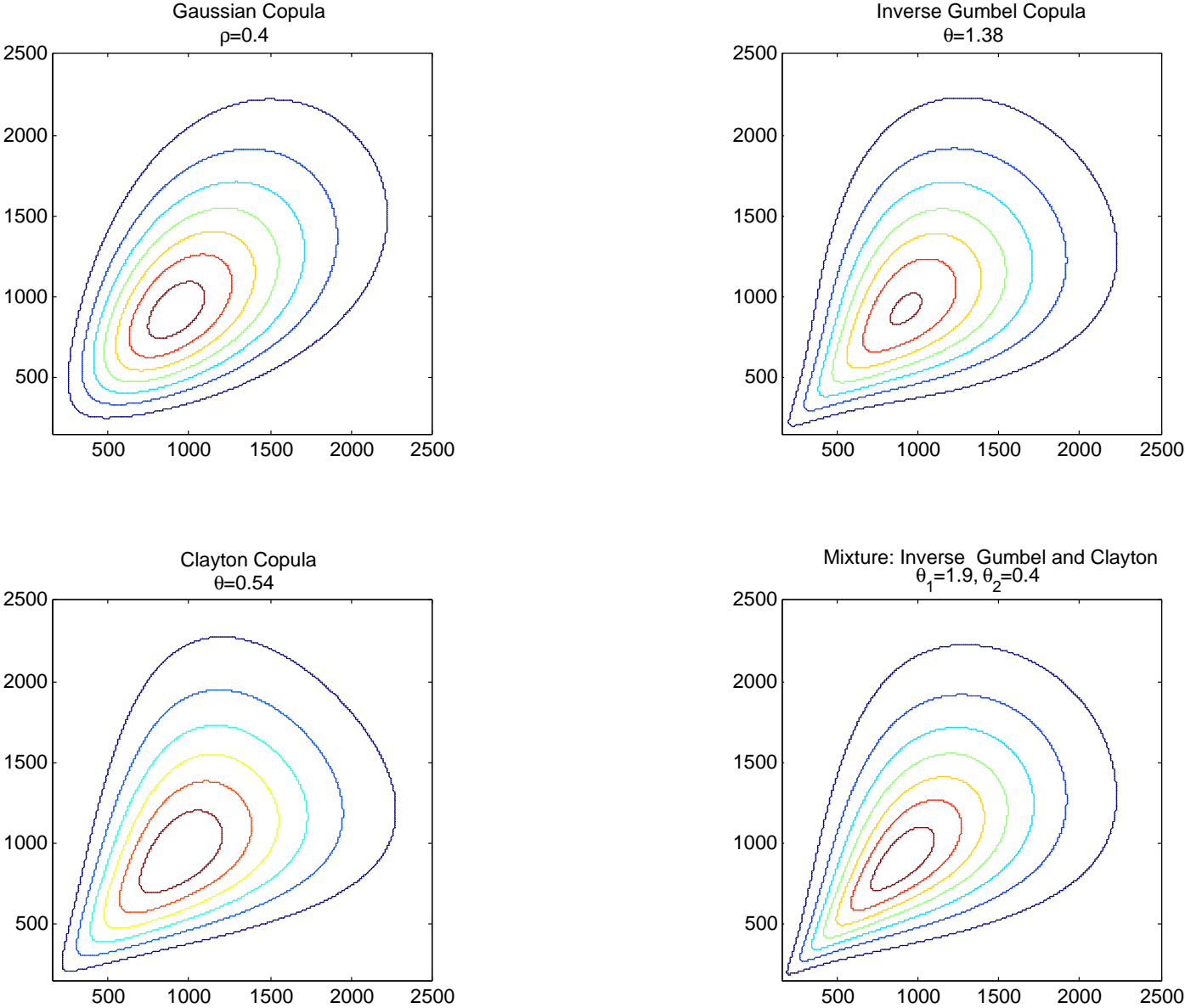
Notes: t -tests critical values: ± 1.96 (5% level), ± 2.58 (1%), ± 3.5 (0.05%).

In Figure 4, we plot 2-dimensional contours of several copula functions with Negative Binomial marginal densities ($1/\kappa_i = 0.2$) and dependence parameters similar to the estimates obtained for this data. Observe the characteristics that prompt the rejection discussed above. The contours of the Gaussian copula exhibit the most elliptical shapes compared to those from the remaining copulas. The rejection of this copula is due to too much correlation imposed by the model on the observations in the upper north-east corner; the inverse Gumbel and Clayton relax the dependence in this area at the expense of imposing stronger dependence in the lower south-west region.

Given the results in Table 9, it seems sensible to think about a model that combines the properties of different copulas. The most natural mixture is Inverse Gumbel with Clayton because it will combine the low t -statistics in autocontours 1% to 80% from the Inverse Gumbel and the lower t -statistics in autocontours 90% to 99%. In the last three columns of Table 9, we report the results for the t -tests for the Inverse Gumbel/Clayton mixture. We estimated by maximum likelihood the weights of the mixture and the estimate is 0.34 on the Inverse Gumbel and 0.66 on Clayton. In addition, the estimate of the Inverse Gumbel dependence is larger ($\theta = 1.9$) than that from the pure Inverse Gumbel model, and the estimate of the Clayton dependence parameter ($\theta = 0.43$) is smaller than that from the pure Clayton model. This mixture provides an improvement in the t -tests for all autocontours from 1% to 90% with values that fail to reject the model at the conventional significance levels, while for the high autocontours 95% and 99%, the t -tests are lower than those from Inverse Gumbel with a mild rejection of the model in the 95% autocontour. In summary, only in the 99% autocontour we find a rejection of the model indicating that in periods with a huge number of trades the dependence in the trading activities of the three banks is smaller than that predicted by the model.

In Figure 4, we see why the mixture model performs better than the other models. The contours of the mixture Inverse Gumbel/Clayton show that the correlation in the observations in the south-west corner is larger than in the Gaussian or in the Clayton copulas, and that the observations in the north-east corner are less correlated than those in the Gaussian copula; nevertheless we still need less dependence among the observations in the most upper north-east corner.

Figure 4: Contours of Copulas with Negative Binomial Marginal Densities

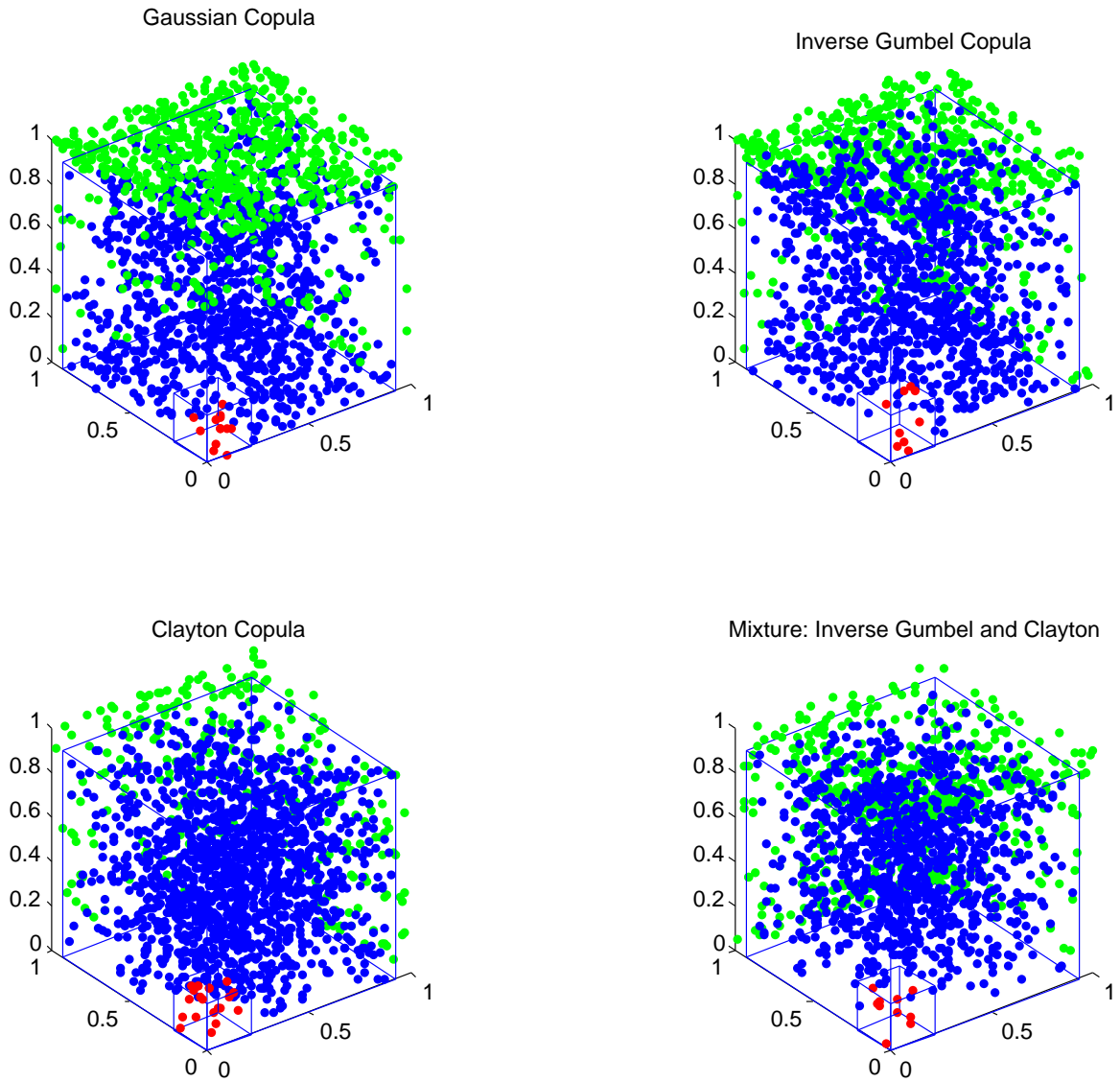


In summary, the combination of statistical tests and graphical devices offer plenty of evidence on the highly asymmetric dependence among trades of these three banks. When the number of trades is low, the dependence is high, and when the number of trades is very large, the dependence is low or non-existent. This finding is analogous to the Epps (1979) effect in high frequency returns, which

shows that the sample correlation among stock returns goes to zero as the sampling interval shrinks. If we translate the number of trades into durations, our findings can be interpreted as an Epps-like effect: when the durations become shorter (large number of trades), the correlation among trades goes to zero, and viceversa, for long durations (low number of trades), the correlation is strong. This empirical effect may be relevant for measurements of liquidity because when liquidity drains (lack of trading) , we should expect a concurrent effect among similar institutions.

In Figure 5 and for the prediction sample, we plot the 3-dimensional vector of contemporaneous PITs for the Negative Binomial model with Gaussian, inverse Gumbel, Clayton, and mixture Inverse Gumbel/Clayton. It is evident that the four models have captured the contemporaneous dependence among the three series observed in Figure 3 (right panel). Now the observations are more evenly distributed within the cubes considered, as it is expected from i.i.d. data. With the Gaussian copula, observe the uneven location of the observations in the contours 80%-99% (green dots), which is the reason why the tests strongly reject this model for the higher contours. The Clayton model corrects substantially the distribution of the observations in the upper contours but at the expense of misplacing observations in the 10%-80% levels (blue dots). The Inverse Gumbel model provides a very good fit at all contours except for the very top 90-99% levels. The mixture Inverse Gumbel/Clayton model offers the best fit as it is able to produce the right distribution of observations in the 10%-90% levels with no so large distortions in the very upper 95%-99% levels.

Figure 5: Contemporaneous PITs: Negative Binomial Model with Copula Dependence



5 Conclusions

We have proposed a new tool, the Generalized Autocontour (G-ACR) as the basis for a battery of dynamic specification tests. G-ACR overcomes some important limitations of the autocontour methodology proposed in González-Rivera *et. al.* (2011), and in doing so, G-ACR tests are useful

diagnostics to assess the conditional density model of random univariate or multivariate processes, either in-sample or out-of-sample. The tests enjoy standard distributions and have outstanding finite sample properties, they are correctly sized and are very powerful to detect departures from the assumed conditional density. In addition, the G-ACR methodology brings a visual aspect to the modeling exercise that is helpful on driving the specification and forecasting exercises. To illustrate the usefulness of our approach we have focused on the modeling of a trivariate system of trades. We have analyzed a high frequency trivariate system of the number of trades in three US large banks. We have shown the need to specify not only the correct dynamics and the individual marginal densities of trades but also the contemporaneous dependence of the three banks. We have modeled such dependence with copula functions and we have found that there is a highly asymmetric response depending on whether trading activity is dense or thin. When the number of trades is low, the contemporaneous dependence is stronger than when the number is high, so that when liquidity dries out in one institution, we should expect similar behavior in the rest.

References

- [1] Cameron, C., and Trivedi, K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge.
- [2] Diebold, F., Gunther, A., and Tay, A. (1998), “Evaluating Density Forecasts with Applications to Financial Risk Management”, *International Economic Review*, vol. 39, 4, pp. 863-883.
- [3] Diebold, F., Hahn, J. and Tay, A. (1999), “Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange”, *The Review of Economics and Statistics*. vol. 81, 4, pp. 661-673.
- [4] Epps, T. (1979), “Comovements in Stock Prices in the Very Short Run ”, *Journal of American Statistical Association*, vol. 74, 336, pp. 291-298.
- [5] González-Rivera, and Yoldas, E. (2012), “Autocontour-based Evaluation of Multivariate Predictive Densities”, *International Journal of Forecasting*, vol. 28, 2, pp. 328-342.
- [6] González-Rivera, G., Senyuz,Z., and Yoldas, E. (2011), “Autocontours: Dynamic Specification Testing”, *Journal of Business and Economic Statistics*, vol. 29, 1, pp. 186-200.
- [7] Heinen, A., and Rengifo, E. (2007), “Multivariate Autoregressive Modeling of Time Series Count Data Using Copulas”, *Journal of Empirical Finance*, Vol. 14, pp. 564-583.
- [8] Madhavan, A. (2000), “Market Microstructure: A Survey”, *Journal of Financial Markets*, vol. 3, pp. 205-258.
- [9] Nelsen, R. (2005). *An Introduction to Copulas*. Second Edition, Springer.
- [10] O’Hara, M. (1995). *Market Microstructure Theory*. Blackswell.
- [11] Patton, A.J. (2006), “Estimation of Multivariate Models for Time Series of Possibly Different Lengths”, *Journal of Applied Econometrics*, vol. 21, 4, pp. 147-173.

Appendix: Mathematical Proofs

At the core of all proofs is the indicator function I_t^{k,α_i} , which is a Bernoulli random variable with the following moments: $E(I_t^{k,\alpha_i}) = \alpha_i$, $Var(I_t^{k,\alpha_i}) = \alpha_i(1 - \alpha_i)$ and covariance

$$r_h^{\alpha_i} \equiv cov(I_t^{k,\alpha_i}, I_{t-h}^{k,\alpha_i}) = \begin{cases} 0 & \text{if } h \neq k \\ \alpha_i^{3/2}(1 - \alpha_i^{1/2}) & \text{if } h = k \end{cases}$$

When $h \neq k$ there is not common information between the indicators I_t^{k,α_i} and I_{t-h}^{k,α_i} , and since u_t 's are i.i.d., their covariance is zero. When $h = k$, $cov(I_t^{k,\alpha_i}, I_{t-k}^{k,\alpha_i}) = E(I_t^{k,\alpha_i} \times I_{t-k}^{k,\alpha_i}) - \alpha_i^2 = \alpha_i^{3/2} - \alpha_i^2$.

Proof of Proposition 1 Asymptotic normality of the test follows as in González-Rivera *et al.* (2011). The asymptotic variance of the test follows directly from the first and second moments of the indicator function stated above.

Proof of Proposition 2 Asymptotic chi-square distribution of the test follows from the proof in González-Rivera *et al.* (2011). The elements $\lambda_{j,k}$ of the asymptotic variance-covariance Λ_{α_i} are obtained as follows. When $j = k$ and by invoking Proposition 1, we have

$$\lambda_{j,k} = cov(\ell_{j,\alpha_i}, \ell_{k,\alpha_i}) = var(\sqrt{T-k}(\hat{\alpha}_i - \alpha_i)) = \alpha_i(1 - \alpha_i) + 2\alpha_i^{3/2}(1 - \alpha_i^{1/2})$$

When $j > k$ (and similarly for $j < k$),

$$cov(\ell_{k,\alpha_i}, \ell_{j,\alpha_i}) = cov(I_t^{k,\alpha_i}, I_t^{j,\alpha_i}) + cov(I_t^{k,\alpha_i}, I_{t-k}^{j,\alpha_i}) + cov(I_{t-j}^{k,\alpha_i}, I_t^{j,\alpha_i}) + cov(I_t^{j,\alpha_i}, I_{t-j+k}^{k,\alpha_i}) + o(1)$$

from which each covariance term is

$$cov(I_t^{k,\alpha_i}, I_t^{j,\alpha_i}) = E(I_t^{k,\alpha_i} \times I_t^{j,\alpha_i}) - \alpha_i^2 = \alpha_i^{3/2} - \alpha_i^2 = \alpha_i^{3/2}(1 - \alpha_i^{1/2})$$

and taking into account that $cov(I_t^{k,\alpha_i}, I_t^{j,\alpha_i}) = cov(I_t^{k,\alpha_i}, I_{t-k}^{j,\alpha_i}) = cov(I_{t-j}^{k,\alpha_i}, I_t^{j,\alpha_i}) = cov(I_{t-j+k}^{k,\alpha_i}, I_t^{j,\alpha_i})$, the result directly follows.

Proof of Proposition 3 Asymptotic chi-square distribution of the test follows from the proof in González-Rivera *et al.* (2011). The elements $\omega_{i,j}$ of the asymptotic variance-covariance Ω_k are

obtained as follows. We need to calculate

$$\text{cov}(c_{k,i}, c_{k,j}) = \text{cov}(I_t^{k,\alpha_i}, I_t^{k,\alpha_j}) + \text{cov}(I_t^{k,\alpha_i}, I_{t-k}^{k,\alpha_j}) + \text{cov}(I_{t-k}^{k,\alpha_i}, I_t^{k,\alpha_j}) + o(1)$$

If $i = j$, by Proposition 1, $\omega_{i,i} = \text{var}(\sqrt{T-k}(\hat{\alpha}_i - \alpha_i)) = \alpha_i(1 - \alpha_i) + 2\alpha_i^{3/2}(1 - \alpha_i^{1/2})$. If $i < j$ (and similarly for $i > j$), $\alpha_i \subset \alpha_j$. Then, we have that

$$\begin{aligned} \text{cov}(I_t^{k,\alpha_i}, I_t^{k,\alpha_j}) &= E(I_t^{k,\alpha_i} \times I_t^{k,\alpha_j}) - \alpha_i \times \alpha_j = \alpha_i(1 - \alpha_j) \\ \text{cov}(I_t^{k,\alpha_i}, I_{t-k}^{k,\alpha_j}) &= E(I_t^{k,\alpha_i} \times I_{t-k}^{k,\alpha_j}) - \alpha_i \times \alpha_j = \alpha_i \times \alpha_j^{1/2} - \alpha_i \times \alpha_j \\ \text{cov}(I_{t-k}^{k,\alpha_i}, I_t^{k,\alpha_j}) &= E(I_{t-k}^{k,\alpha_i} \times I_t^{k,\alpha_j}) - \alpha_i \times \alpha_j = \alpha_i \times \alpha_j^{1/2} - \alpha_i \times \alpha_j \end{aligned}$$

and the result follows.