

Time-Varying Systemic Risk: Evidence from a Dynamic Copula Model of CDS Spreads*

Dong Hwan Oh[†]

Andrew J. Patton[‡]

Federal Reserve Board

Duke University

13 May 2015

Abstract

This paper proposes a new class of copula-based dynamic models for high-dimensional conditional distributions, facilitating the estimation of a wide variety of measures of systemic risk. Our proposed models draw on successful ideas from the literature on modeling high-dimensional covariance matrices and on recent work on models for general time-varying distributions. Our use of copula-based models enables the estimation of the joint model in stages, greatly reducing the computational burden. We use the proposed new models to study a collection of daily credit default swap (CDS) spreads on 100 U.S. firms over the period 2006 to 2012. We find that while the probability of distress for individual firms has greatly reduced since the financial crisis of 2008-09, the joint probability of distress (a measure of systemic risk) is substantially higher now than in the pre-crisis period.

Keywords: correlation, tail risk, financial crises, DCC.

J.E.L. codes: C32, C58, G01.

*We thank Christian Brownlees, Drew Creal, Jeff Russell, and seminar participants at Chicago, Duke, Monash, Montréal, NYU, Sveriges Riksbank, Toulouse School of Economics, UNC-Chapel Hill, UNSW and UTS for helpful comments. An appendix containing additional results and a MATLAB toolbox for this paper is available at <http://www.econ.duke.edu/~ap172/research.html>. The views in this paper are those of the authors and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff.

[†]Quantitative Risk Analysis Section, Federal Reserve Board, Washington DC 20551. Email: donghwan.oh@frb.gov

[‡]Department of Economics, Duke University, Box 90097, Durham NC 27708. Email: andrew.patton@duke.edu

1 Introduction

Systemic risk can be broadly defined as the risk of distress in a large number of firms or institutions. It represents an extreme event in two directions: a large loss (e.g., corresponding to a large left-tail realization for stock returns), across a large proportion of the firms. There are a variety of methods for studying risk and dependence for small collections of assets (see Patton (2013) for a recent review) but a paucity of methods for studying dependence between a large collection of assets, which is required for a general analysis of systemic risk.

Some existing methods for estimating systemic risk simplify the task by reducing the dimension of the problem to two: an individual firm and a market index. The “CoVaR” measure of Adrian and Brunnermeier (2009), for example, uses quantile regression to estimate a lower tail quantile (e.g., 0.05) of market returns conditional on a given firm having a returns equal to its lower tail quantile. The “marginal expected shortfall” proposed by Brownlees and Engle (2015) estimates the expected return on a firm conditional on the market return being below some low threshold. These methods have the clear benefit of being parsimonious, but by aggregating the “non firm i ” universe to a single market index, useful information about systemic risk may be missed. For example, firm i may impact a sizeable subset of the other firms, but fall short of precipitating a market-wide event. The objective of this paper is to provide models that can be used to handle large collections of variables, and facilitate the estimation of a wider variety of systemic risk measures.

We use Sklar’s theorem (see Nelsen, 2006), with an extension to conditional distributions from Patton (2006), to decompose the conditional joint distribution of a collection of N variables, $\mathbf{Y}_t = [Y_{1t}, \dots, Y_{Nt}]'$, into their marginal distributions and a conditional copula:

$$\mathbf{Y}_t | \mathcal{F}_{t-1} \sim \mathbf{F}_t = \mathbf{C}_t(F_{1t}, \dots, F_{Nt}) \tag{1}$$

We propose new models for the time-varying conditional copula, \mathbf{C}_t , that can be used to link models of the conditional marginal distributions (e.g., ARMA-GARCH models) to form a dynamic

conditional joint distribution. Of central relevance to this paper are cases where N is relatively large, around 50 to 250. In such cases, models that have been developed for low dimension problems (say, $N < 5$) are often not applicable, either because no generalization beyond the bivariate model exists, or because such generalizations are too restrictive (e.g., Archimedean copulas have just one or two free parameters regardless of N , which is clearly very restrictive in high dimensions), or because the obvious generalization of the bivariate case leads to a proliferation of parameters and unmanageable computational complexity. In high-dimensional applications, the challenge is to find a balance of flexibility and parsimony.

This paper makes two contributions. First, we propose a flexible and feasible model for capturing time-varying dependence in high dimensions. Our approach draws on successful ideas from the literature on dynamic modeling of high-dimensional covariance matrices and on recent work on models for general time-varying distributions. In particular, we combine the “GAS” model of Creal, *et al.* (2011, 2013) and the factor copula model of Oh and Patton (2012) to obtain a flexible yet parsimonious dynamic model for high-dimensional conditional distributions. This model is generally only estimable using numerical optimization techniques, and we propose a “variance targeting” (Engle and Mezrich (1996) and Engle (2002)) type estimator to dramatically reduce the dimension of the optimization problem. We show that this estimation method is a form of (nonlinear) GMM, and thus can be treated as part of a multi-stage GMM estimation of the entire model. A realistic simulation study confirms that our proposed models and estimation methods have satisfactory properties for relevant sample sizes.

Our second contribution is a detailed study of a collection of 100 daily credit default swap (CDS) spreads on U.S. firms. The CDS market has expanded enormously over the last decade, growing 40-fold from \$0.6 trillion of gross notional principal in 2001 to \$25.9 trillion at the end of 2011 according to the International Swaps and Derivatives Association (ISDA), yet it has received relatively little attention, compared with equity returns, in the econometrics literature. (Interest is growing, however, see Conrad, *et al.* (2011), Creal, *et al.* (2014), Christoffersen, *et al.* (2013),

Lucas, *et al.* (2014) and Creal and Tsay (2014) for recent work on CDS data.) We use our model of CDS spreads to provide insights into systemic risk, as CDS spreads are tightly linked to the health of the underlying firm. We find that systemic risk rose during the financial crisis, unsurprisingly. More interestingly, we also find that systemic risk remains high relative to the pre-crisis period, even though idiosyncratic risk has fallen.

The remainder of the paper is structured as follows. Section 2 presents a dynamic copula model for high-dimensional applications, and Section 3 presents a simulation study for the proposed model and estimation method. In Section 4 we present estimation results for various models of CDS spreads. Section 5 applies the model to estimate time-varying systemic risk, and Section 6 concludes. Technical details are presented in the appendix, and an internet appendix contains some additional results.

2 A dynamic copula model for high dimensions

In this section we describe our approach for capturing dynamics in the dependence between a relatively large number of variables. (A review of alternative methods from the small but growing literature on this topic is presented in Section 2.5.) We consider a class of data generating processes (DGPs) that allow for time-varying conditional marginal distributions, e.g., dynamic conditional means and variances, and also possibly time-varying higher-order moments:

$$\mathbf{Y}_t \equiv [Y_{1t}, \dots, Y_{Nt}]' \tag{2}$$

$$\text{where } Y_{it} = \mu_{it}(\phi_{i,0}) + \sigma_{it}(\phi_{i,0})\eta_{it}, \quad i = 1, 2, \dots, N$$

$$\eta_{it} | \mathcal{F}_{t-1} \sim F_{it}(\phi_{i,0})$$

where μ_{it} is the conditional mean of Y_{it} , σ_{it} is the conditional standard deviation, and $F_{it}(\phi_{i,0})$ is a parametric distribution with zero mean and unit variance. We will denote the parameters of the

marginal distributions as $\boldsymbol{\phi} \equiv [\phi'_1, \dots, \phi'_N]'$, the parameters of the copula as $\boldsymbol{\gamma}$, and the vector of all parameters as $\boldsymbol{\theta} \equiv [\boldsymbol{\phi}', \boldsymbol{\gamma}']'$. We assume that F_{it} is continuous and strictly increasing, which fits our empirical application, though this assumption can be relaxed. The information set is taken to be $\mathcal{F}_t = \sigma(\mathbf{Y}_t, \mathbf{Y}_{t-1}, \dots)$. Define the conditional probability integral transforms of the data as:

$$U_{it} \equiv F_{it} \left(\frac{Y_{it} - \mu_{it}(\boldsymbol{\phi}_{i,0})}{\sigma_{it}(\boldsymbol{\phi}_{i,0})}; \boldsymbol{\phi}_{i,0} \right), \quad i = 1, 2, \dots, N \quad (3)$$

Then the conditional copula of $\mathbf{Y}_t | \mathcal{F}_{t-1}$ is equal to the conditional distribution of $\mathbf{U}_t | \mathcal{F}_{t-1}$:

$$\mathbf{U}_t | \mathcal{F}_{t-1} \sim \mathbf{C}_t(\boldsymbol{\gamma}_0) \quad (4)$$

By allowing for a time-varying conditional copula, the class of DGPs characterized by equations (2) to (4) is a generalization of those considered by Chen and Fan (2006), for example, however the cost of this flexibility is the need to specify parametric marginal distributions. In contrast, Chen and Fan (2006), Rémillard (2010) and Oh and Patton (2013) allow for nonparametric estimation of the marginal distributions. The parametric margin requirement arises as the asymptotic distribution theory for a model with nonparametric margins and a time-varying copula is not yet available in the literature. We attempt to mitigate this requirement in our empirical work by using flexible models for the marginal distributions, which are allowed to differ for each variable, and conducting goodness-of-fit tests to verify that they provide a satisfactory fit to the data.

2.1 Factor copulas

In high-dimensional applications a critical aspect of any model is imposing some form of dimension reduction. A widely-used method to achieve this in economics and finance is to use a factor structure. Oh and Patton (2012) propose using a factor model with flexible distributions to obtain a flexible class of “factor copulas.” A one-factor version of their model is the copula for the latent

vector random variable $\mathbf{X}_t \equiv [X_{1t}, \dots, X_{Nt}]'$ implied by the following structure:

$$X_{it} = \lambda_{it}(\gamma_\lambda) Z_t + \varepsilon_{it}, \quad i = 1, 2, \dots, N \quad (5)$$

$$\text{where } Z_t \sim F_{z_t}(\gamma_z), \quad \varepsilon_{it} \sim iid F_{\varepsilon_t}(\gamma_\varepsilon), \quad Z_t \perp\!\!\!\perp \varepsilon_{it} \quad \forall i$$

where $F_{z_t}(\gamma_z)$ and $F_{\varepsilon_t}(\gamma_\varepsilon)$ are flexible parametric univariate distributions for the common factor and the idiosyncratic variables, and $\lambda_{it}(\gamma_\lambda)$ is a potentially time-varying weight on the common factor. The collection of all copula parameters is $\gamma \equiv [\gamma'_z, \gamma'_\varepsilon, \gamma'_\lambda]'$. The conditional joint distribution for \mathbf{X}_t can be decomposed into its conditional marginal distributions and its conditional copula via Sklar's theorem (see Nelsen (2006)) for conditional distributions, see Patton (2006):

$$\mathbf{X}_t \sim \mathbf{G}_t = \mathbf{C}_t(G_{1t}(\gamma), \dots, G_{Nt}(\gamma); \gamma) \quad (6)$$

Note that only the *copula* of these latent variables, denoted $\mathbf{C}_t(\gamma)$, is used as a model for the copula of the observable data \mathbf{Y}_t ; the marginal distributions of \mathbf{X}_t need *not* be the same as the marginal distributions of the observed data. (These marginals, G_{it} , are effectively discarded.) If we impose that the marginal distributions of the observable data are also driven by the factor structure in equation (5), then we would also use the implied marginal distributions, and this approach then becomes a standard factor model for a vector of variables. However, Oh and Patton (2012) propose imposing the factor structure only on the component of the multivariate model where dimension reduction is critical, namely the copula, and allow the marginal distributions to be modeled using a potentially different approach. In this case, the factor structure in equation (5) is used only for the copula that it implies, and this becomes a “factor copula” model.

The copula implied by equation (5) is known in closed form for only a few particular combinations of choices of F_z and F_ε (the most obvious example being where both of these distributions are Gaussian, in which case the implied copula is also Gaussian). For general choices of F_z and

F_ε the copula of \mathbf{X} will not be known in closed form, and thus the copula likelihood is not known in closed form. Numerical methods can be used to overcome this problem. Oh and Patton (2013) propose simulated method of moments-type estimation of the unknown parameters, however their approach is only applicable when the conditional copula is constant. A key objective of this paper is to allow the conditional copula to vary through time and so an alternate estimation approach is desirable. We use a simple numerical integration method, described in Appendix B, to overcome the lack of closed-form likelihood. This numerical integration exploits the fact that although the copula is N -dimensional, we need only integrate out the common factor, which is one-dimensional in the structure above.

Also, analogous to this issue in other latent factor models, the copula implied by equation (5) is only identified up to the sign of the loadings (λ_{it}) and the sign of the common factor (Z_t). That is, for a given set of loadings and common factor distribution, an identical copula can be obtained by switching the signs of all loadings and by flipping the distribution of Z_t around zero. In some applications, this issue can complicate interpreting the parameters of the factor model. In our application, we exploit the fact that all of the variables are positively related, and we thus impose that the factor loadings are all positive.

Dynamics in the factor copula model in equation (5) arise by allowing the loadings on the common factor, λ_{it} , to vary through time, and/or by allowing the distributions of the common factor and the idiosyncratic variables to change through time. For example, holding F_{zt} and $F_{\varepsilon t}$ fixed, an increase in the factor loadings corresponds to an increase in the level of overall dependence (e.g., rank correlation) between the variables. Holding the factor loadings fixed, an increase in the thickness of the tails of the distribution of the common factor increases the degree of tail dependence. In the next section we describe how we model these dynamics.

2.2 Generalized autoregressive score dynamics

An important feature of any dynamic model is the specification for how the parameters evolve through time. Some specifications, such as stochastic volatility models (see Shephard (2005) for example) and related stochastic copula models (see Hafner and Manner (2012) and Manner and Segers (2011)) allow the varying parameters to evolve as a latent time series process. Others, such as ARCH-type models for volatility (see Engle, 1982) and related models for copulas (see Patton (2006), Jondeau and Rockinger (2006), and Creal, *et al.* (2013) for example) model the varying parameters as some function of lagged observables. An advantage of the latter approach over the former, especially for high-dimensional applications, is that it avoids the need to “integrate out” the innovation terms driving the latent time series processes.

Within the class of ARCH-type models (“observation driven”, in the terminology of Creal, *et al.* (2013)), the question of *which* function of lagged observables to use as a forcing variable in the evolution equation for the varying parameter arises. For models of the conditional variance, an obvious choice is the lagged squared residual, as in the ARCH model, but for models with parameters that lack an obvious interpretation the choice is less clear. We adopt the generalized autoregressive score (GAS) model of Creal, *et al.* (2013) to overcome this problem. (Harvey (2013) and Harvey and Sucarrat (2014) propose a similar method for modeling time-varying parameters, which they call a “dynamic conditional score,” or “DCS,” model.) These authors propose using the lagged score of the density model (copula model, in our application) as the forcing variable. Specifically, for a copula with time-varying parameter $\boldsymbol{\delta}_t$, governed by fixed parameter $\boldsymbol{\gamma}$, we have:

$$\begin{aligned} \text{Let } \mathbf{U}_t | \mathcal{F}_{t-1} &\sim \mathbf{C}(\boldsymbol{\delta}_t(\boldsymbol{\gamma})) \\ \text{then } \boldsymbol{\delta}_t &= \boldsymbol{\omega} + B\boldsymbol{\delta}_{t-1} + A\mathbf{s}_{t-1} \\ \text{where } \mathbf{s}_{t-1} &= S_{t-1} \boldsymbol{\nabla}_{t-1} \\ \boldsymbol{\nabla}_{t-1} &= \frac{\partial \log \mathbf{c}(\mathbf{u}_{t-1}; \boldsymbol{\delta}_{t-1})}{\partial \boldsymbol{\delta}_{t-1}} \end{aligned} \tag{7}$$

and S_t is a scaling matrix, e.g., the inverse Hessian or its square root. (The specification for the dynamics of the parameter δ_t can also be generalized to include observable variables, similar to the GARCH-X (Brenner *et al.*, 1996) extension of a GARCH model, see De Lira Salvatierra and Patton (2014), although we do not pursue that here.) If the copula likelihood is not available in closed form, as is generally the case with factor copulas, then the vector ∇_t can be obtained via numerical differentiation.

While this specification for the evolution of a time-varying parameter is clearly somewhat arbitrary, Creal, *et al.* (2013) motivate it by showing that it nests a variety of popular and successful existing models: GARCH (Bollerslev (1986)) for conditional variance; ACD (Engle and Russell (1998)) for models of trade durations (the time between consecutive high frequency observations); Davis, *et al.*'s (2003) model for Poisson counts. In recent work, Blasques, *et al.* (2014a) show that the GAS model can be motivated information-theoretically, by showing that this approach minimizes the local Kullback-Liebler divergence between the true conditional density and the model-implied density, and Harvey (2013) motivates this specification as an approximation to a filter for a model driven by a stochastic latent parameter, or an “unobserved components” model. Finally, the recursion above can also be interpreted as the steepest ascent direction for improving the model’s fit, in terms of the likelihood, given the current value of the model parameter δ_t , similar to numerical optimization algorithms such as the Gauss-Newton algorithm.

2.3 High-dimensional factor copulas with GAS dynamics

We employ the GAS model to allow for time variation in the factor loadings in the factor copula implied by equation (5), but to keep the model parsimonious we impose that the parameters governing the “shape” of the common and idiosyncratic variables (γ_z and γ_ε) are constant. We use the skewed t distribution of Hansen (1994) as the model for F_z . This distribution has two shape parameters, a degrees of freedom parameter ($\nu_z \in (2, \infty]$) and an asymmetry parameter ($\psi_z \in (-1, 1)$), and it simplifies to the standardized t distribution when $\psi = 0$. We use the standardized t distribution as

the model for F_ε for simplicity.

The general GAS framework in equation (7) applied to the factor loadings (λ_{it}) in equation (5), would imply $N + 2N^2$ parameters governing their evolution, which represents an infeasibly large number for even moderate values of N . To keep the model parsimonious, we adopt a simplification from the DCC model of Engle (2002), and impose that the coefficient matrices (B and A) are diagonal with a scalar parameter (β and α respectively) on the diagonal. To avoid the estimation of an $N \times N$ scaling matrix we set $S_t = I$. This simplifies our model to be (in logs):

$$\log \lambda_{it} = \omega_i + \beta \log \lambda_{i,t-1} + \alpha s_{i,t-1}, \quad i = 1, 2, \dots, N \quad (8)$$

where $s_{it} \equiv \partial \log \mathbf{c}(\mathbf{u}_t; \boldsymbol{\lambda}_t, \nu_z, \psi_z, \nu_\varepsilon) / \partial \log \lambda_{it}$ and $\boldsymbol{\lambda}_t \equiv [\lambda_{1t}, \dots, \lambda_{Nt}]'$. In our empirical application below, we found slightly better results from modeling $\log \lambda_{it}$, and thus imposing positive weights, than from modeling λ_{it} directly. The use of the logarithm here is not essential to our empirical approach.

The dynamic copula model implied by equations (5) and (8) contains $N + 2$ parameters for the GAS dynamics, 3 parameters for the shape of the common and idiosyncratic variables, for a total of $N + 5$ parameters. Numerically optimizing this model when $N = 50$ or 100 represents quite a computational challenge. We propose a method to overcome this challenge by adapting an idea from the DCC model of Engle (2002), known as “variance targeting.” The nature of our GAS specification means that the variance targeting approach needs to be modified for use here.

The evolution equation for λ_{it} in equation (8) can be re-written as

$$\log \lambda_{it} = E[\log \lambda_{it}] (1 - \beta) + \beta \log \lambda_{i,t-1} + \alpha s_{i,t-1} \quad (9)$$

using the result from Creal, *et al.* (2013) that $E_{t-1}[s_{it}] = 0$, and so $E[\log \lambda_{it}] = \omega_i / (1 - \beta)$. The proposition below provides a method for using sample rank correlations to obtain an estimate of

$E[\log \lambda_{it}]$, which eliminates the need to numerically optimize over the intercept parameters, ω_j .

The proposition is based on the following assumption.

Assumption 1 (a) *The conditional copula of $\mathbf{Y}_t|\mathcal{F}_{t-1}$ is the time-varying factor copula given in equations (5) and (8).*

(b) *The process $\{\boldsymbol{\lambda}_t\}$ generated by equation (8) is strictly stationary.*

(c) *Let $\boldsymbol{\rho}_{t,X} \equiv \text{vech}(\text{RankCorr}_{t-1}[\mathbf{X}_t])$. Then $\log \boldsymbol{\lambda}_t$ is a linear function of $\boldsymbol{\rho}_{t,X}$.*

(d) *Let $\rho_{ij,X} \equiv \text{RankCorr}[X_i, X_j]$ and $\rho_{ij,X}^L \equiv \text{Corr}[X_i, X_j]$. Then, for fixed values of $(\gamma_z, \gamma_\varepsilon)$, the mapping $\rho_{ij,X} = \varphi(\rho_{ij,X}^L)$ is strictly increasing.*

Part (a) of this assumption makes explicit that the copula of the data is the GAS-factor copula model, and so the conditional copula of $\mathbf{Y}_t|\mathcal{F}_{t-1}$ is the same as that of $\mathbf{X}_t|\mathcal{F}_{t-1}$. Blasques, *et al.* (2014b) provide conditions under which univariate GAS models satisfy stationarity conditions, and Blasques, *et al.* (2014a) provide results for GAS models with a vector dependent variable but a scalar time-varying parameter; corresponding theoretical results for our multivariate case are not yet available in the literature, and in part (b) we simply assume that stationarity holds. Part (c) formalizes the applicability of a Taylor series expansion of the function mapping $\boldsymbol{\rho}_t$ to $\boldsymbol{\lambda}_t$. In practice this assumption will hold only approximately, and its applicability needs to be verified via simulation, which we discuss further in Section 3 and Appendix B. Part (d) enables us to map rank correlations to linear correlations. Note that we can take $(\gamma_z, \gamma_\varepsilon)$ as fixed, as we call this mapping for each evaluation of the log-likelihood. Importantly, this mapping can be computed *prior* to estimation, and then just called during estimation, rather than re-computed each time the likelihood function is evaluated, which greatly speeds up estimation.

Proposition 1 *Let Assumption 1 hold, and denote the vech of the rank correlation matrix of the standardized residuals, $\boldsymbol{\eta}_t$, as $\bar{\boldsymbol{\rho}}_\eta$ and its sample analog as $\hat{\boldsymbol{\rho}}_\eta$. Then:*

(i) $E[\log \boldsymbol{\lambda}_t] = H(\bar{\boldsymbol{\rho}}_\eta)$, where H is defined in equation (26).

(ii) $\widehat{\log \boldsymbol{\lambda}} = H(\hat{\boldsymbol{\rho}}_\eta)$ is a GMM estimator of $E[\log \boldsymbol{\lambda}_t]$.

Part (i) of the above proposition provides the mapping from the population rank correlation of the standardized residuals to the mean of the (log) factor loadings, which is the basis for considering a variance-targeting type estimator. Part (ii) shows that the sample analog of this mapping can be interpreted as a standard GMM estimator. This is useful as it enables us to treat the estimation of the complete vector of distribution parameters as multi-stage GMM. Under standard regularity conditions, multi-stage GMM estimation (which nests multi-stage MLE) can be shown to yield consistent and asymptotically normal parameter estimates (see White (1994), Engle and Sheppard (2001), and in particular Newey and McFadden, 1994, Theorem 6.1). That is, if we let $\hat{\boldsymbol{\theta}}_{MS-GMM} \equiv [\hat{\boldsymbol{\phi}}'_{MS-GMM}, \hat{\boldsymbol{\gamma}}'_{MS-GMM}]'$ denote the collection of all estimated parameters, then

$$\sqrt{T} \left(\hat{\boldsymbol{\theta}}_{MS-GMM} - \boldsymbol{\theta}^* \right) \xrightarrow{d} N(0, V_{MS-GMM}^*) \text{ as } T \rightarrow \infty \quad (10)$$

Consistent estimation of V_{MS-GMM}^* is theoretically possible, however in high dimensions it is not computationally feasible. Gonçalves *et al.* (2013) provide conditions under which a block bootstrap may be used to obtain valid standard errors on parameters estimated via multi-stage GMM. The resulting standard errors are not higher-order efficient, like some bootstrap inference methods, but they do enable us to avoid having to handle Hessian matrices of size on the order of $N \times N$. Note that sample rank correlations cannot in general be considered as moment-based estimators, as they depend on the sample ranks of the observed data, and studying their estimation properties requires alternative techniques. However, we exploit the fact that the marginal distributions of the data are known up to an unknown parameter vector, and thus rank correlation can be obtained as a sample moment of a nonlinear function of the data.

2.4 Equidependence vs. heterogeneous dependence

To investigate whether we can further reduce the number of free parameters in this model we consider two restrictions of the model in equation (8), motivated by the “dynamic equicorrelation”

model of Engle and Kelly (2012). If we impose that $\omega_i = \omega \forall i$, then the pair-wise dependence between each of the variables will be identical, leading to a “dynamic equidependence” model. (The copula implied by this specification is “exchangeable” in the terminology of the copula literature.) In this case we have only six parameters to estimate, independent of the number of variables N , vastly reducing the estimation burden, but imposing a lot of homogeneity on the model.

An intermediate step between the fully flexible model and the equidependence model is to group the assets using some *ex ante* information (e.g., by industry) and impose homogeneity only within groups. This leads to a “block equidependence” copula model, with

$$\begin{aligned} X_{it} &= \lambda_{g(i),t} Z_t + \varepsilon_{it}, \quad i = 1, 2, \dots, N \\ \log \lambda_{g,t} &= \omega_g + \beta \log \lambda_{g,t-1} + \alpha s_{g,t-1}, \quad g = 1, 2, \dots, G \end{aligned} \tag{11}$$

where $g(i)$ is the group to which variable i belongs, and G is the number of groups. In this case the number of parameters to estimate in the copula model is $G + 5$. In our empirical application we have $N = 100$ and we consider grouping variables into $G = 5$ industries, meaning this model has 10 parameters to estimate rather than 105. In our empirical analysis below, we compare these two restricted models ($G = 1$ and $G = 5$) with the “heterogeneous dependence” model which allows a different factor for each variable (and so $G = N$).

2.5 Other models for dynamic, high-dimensional copulas

As noted above, the literature contains relatively few models for dynamic, high-dimensional, copulas. Exceptions to this are discussed here. Lucas, *et al.* (2014) combine GAS dynamics with a skewed t copula to model ten sovereign CDS spreads. A similar model, though with an alternative skew t specification and with Engle’s (2002) DCC dynamics, is used by Christoffersen, *et al.* (2012, 2013). The former of these two papers analyzes equity returns on up to 33 national stock indices, while the latter studies weekly equity returns and CDS spreads on 233 North American firms (and

is the largest time-varying copula model in the extant literature). Creal and Tsay (2014) propose a stochastic copula model based on a factor structure, and use Bayesian estimation methods to apply it to an unbalanced panel of CDS spreads and equity returns on 100 firms. Almeida *et al.* (2012) use “vine” copulas to model the dependence between 30 German stock return series, with dynamics captured through a stochastic volatility-type equation for the parameters of the copula. Stöber and Czado (2012) also use vine copulas, combined with a regime-switching model for dynamics, to model dependence between ten German stock returns.

3 Simulation study

This section presents an analysis of the finite sample properties of maximum likelihood estimation for factor copulas with GAS dynamics. Factor copulas do not have a closed-form likelihood, and we approximate the likelihood using standard numerical integration methods, details of which can be found in Appendix B.

We consider three different copula models for the Monte Carlo simulation: a dynamic equidependence model ($G = 1$), a dynamic block equidependence model ($G = 10$), and a dynamic heterogeneous dependence model ($G = N$), all of them governed by:

$$\begin{aligned}
 X_{it} &= \lambda_{g(i),t} Z_t + \varepsilon_{it}, \quad i = 1, 2, \dots, N & (12) \\
 \log \lambda_{g,t} &= \omega_g + \beta \log \lambda_{g,t-1} + \alpha s_{g,t-1}, \quad g = 1, 2, \dots, G \\
 Z &\sim \text{Skew } t(\nu_z, \psi_z) \\
 \varepsilon_i &\sim \text{iid } t(\nu_\varepsilon), \quad \text{and } \varepsilon_i \perp\!\!\!\perp Z \quad \forall i
 \end{aligned}$$

We set $N = 100$ to match the number of series in our empirical application below. For simplicity, we impose that $\nu_z = \nu_\varepsilon$, and we estimate ν^{-1} rather than ν , so that Normality is nested at $\nu^{-1} = 0$ rather than $\nu \rightarrow \infty$. Broadly matching the parameter estimates we obtain in our empirical

application, we set $\omega = 0$, $\beta = 0.98$, $\alpha = 0.05$, $\nu = 5$, and $\psi_z = 0.1$ for the equidependence model. The block equidependence model uses the same parameters but sets $\omega_1 = -0.03$ and $\omega_{10} = 0.03$, and with ω_2 to ω_9 evenly spaced between these two bounds, and the heterogeneous dependence model similarly uses $\omega_1 = -0.03$ and $\omega_{100} = 0.03$, with ω_2 to ω_{99} evenly spaced between these two bounds. Rank correlations implied by these values range from 0.1 to 0.7. With these choices of parameter values and dependence designs, various dynamic dependence structures are covered, and asymmetric tail dependence, which is a common feature of financial data, is also allowed. We use a sample size of $T = 500$ and we repeat each simulation 100 times.

The results for the equidependence model presented in Panel A of Table 1 reveal that the average estimated bias for all parameters is small, and the estimates are centered on true values. The results for the block equidependence model, presented in Panel B, are also satisfactory, and, as expected, the estimation error in the parameters is generally slightly higher for this more complicated model.

The heterogeneous dependence model is estimated using the variance targeting-type approach for the intercepts, ω_i , described in Section 2.3, combined with numerical optimization for the remaining parameters. Appendix B presents reports simulation results that verify the applicability of Assumption 1 for this model, and the results presented in Panel C confirm that the approach leads to estimators with satisfactory finite-sample properties. (Panel C reports only every fifth intercept parameter, in the interests of space. The complete set of results is available in the internet appendix.) The standard errors on the estimated intercept parameters are approximately twice as large, on average, as in the block equidependence case, however this model has seven times as many parameters as the block equidependence (104 vs. 14) and so some loss in accuracy is inevitable. Importantly, all estimated parameters are approximately centered on their true values, confirming that the assumptions underlying Proposition 1 are applicable for this model.

[INSERT TABLE 1 ABOUT HERE]

4 Data description and estimation results

4.1 CDS spreads

We apply the new dynamic copula model proposed in the previous section to daily credit default swap (CDS) spreads, obtained from Markit. In brief, a CDS is a contract in which the seller provides insurance to the buyer against any losses resulting from a default by the “reference entity” within some horizon. We focus on North American corporate CDS contracts, and the reference entities are thus North American firms. The CDS spread, usually measured in basis points and payable quarterly by the buyer to the seller, is the cost of this insurance. See Duffie and Singleton (2003) and Hull (2012) for more detailed discussions of CDS contracts, and see Barclays “CDS Handbook” (2010) for institutional details.

A key reason for interest in CDS contracts is the sensitivity of CDS spreads to changes in market perceptions of the probability of default, see Conrad, *et al.* (2011), Creal, *et al.* (2014), Christoffersen, *et al.* (2013) and Liu (2014) for recent empirical studies of implied default probabilities. Under some simplifying assumptions (such as a constant risk free rate and default hazard rate) see Carr and Wu (2011) for example, it can be shown that the CDS spread in basis points is:

$$S_{it} = 100^2 P_{it}^{\mathbb{Q}} L_{it} \quad (13)$$

where L_{it} is the loss given default (sometimes shortened to “LGD,” and often assumed to equal 0.6 for U.S. firms) and $P_{it}^{\mathbb{Q}}$ is the implied probability of default. The same formula can also be obtained as a first-order approximation at $P_{it}^{\mathbb{Q}} \approx 0$ for other more complicated pricing equations. This expression can be written in terms of the objective probability of default, $P_{it}^{\mathbb{P}}$:

$$S_{it} = 100^2 P_{it}^{\mathbb{P}} \mathcal{M}_{it} L_{it} \quad (14)$$

where \mathcal{M}_{it} is the market price of risk (stochastic discount factor). An increase in a CDS spread can

be driven by an increase in the LGD, an increase in the market price of default risk for this firm, or an increase in the objective probability of default. Any one of these three effects is indicative of a worsening of the health of the underlying firm.

In the analysis below we work with the log-difference of CDS spreads, to mitigate their autoregressive persistence, and under this transformation we obtain:

$$Y_{it} \equiv \Delta \log S_{it} = \Delta \log P_{it}^{\mathbb{P}} + \Delta \log \mathcal{M}_{it} + \Delta \log L_{it} \quad (15)$$

If the loss given default is constant then the third term above vanishes, and if we assume that the market price of risk is constant (as in traditional asset pricing models) or evolves slowly (for example, with a business cycle-type frequency) then daily changes in CDS spreads can be attributed primarily to changes in the objective probability of default. We will use this to guide our interpretation of the empirical results below, but we emphasize here that an increase in any of these three terms represents “bad news” for firm i , and so the isolation of the objective probability of default is not required for our interpretations to follow.

4.2 Summary statistics

Our sample period spans January 2006 to April 2012, a total of 1644 days. We study the 5-year CDS contract, which is the most liquid horizon (see Barclays (2010)), and we use “XR” (“no restructuring”) CDS contracts, which became the convention for North America following the CDS market standardization in 2009 (the so-called “Big Bang”). To obtain a set of active, economically interesting, CDS data, we took all 125 individual firms in the CDS index covering our sample period (CDX NA IG Series 17). Of these, 90 firms had data that covered our entire sample period, and ten firms had no more than three missing observations. We use these 100 firms for our analysis. (Of the remaining 25 firms, six are not U.S.-based firms and one firm stopped trading because of a firm split. None of the firms defaulted over this sample period.) A plot of these CDS spreads

is presented in Figure 1, which reveals that the average CDS spread was around 100 basis points (bps), and it varied from a low (averaged across firms) of 24 bps on February 22, 2007, to a high of 304 bps on March 9, 2009.

[INSERT FIGURE 1 ABOUT HERE]

The levels of our CDS spread data are suggestive of a large autoregressive root, with the median first-order autocorrelation across all 100 series being 0.996 (the minimum is 0.990). Further, augmented Dickey-Fuller tests reject the null hypothesis of a unit root at the 0.05 level for only 12 series. Like interest rate time series, these series are unlikely to literally obey a random walk, as they are bounded below, however we model all series in log differences to avoid the need to consider these series as near unit root processes.

Table 2 presents summary statistics on our data. Of particular note is the positive skewness of the log-differences in CDS spreads (average skewness is 1.087, and skewness is positive for 89 out of 100 series) and the excess kurtosis (25.531 on average, and greater than 3 for all 100 firms). Ljung-Box tests for autocorrelation at up to the tenth lag find significant (at the 0.05 level) autocorrelation in 98 out of 100 of the log-differenced CDS spreads, and for 89 series significant autocorrelation is found in the squared log-differences. This motivates specifying models for the conditional mean and variance to capture this predictability.

[INSERT TABLE 2 ABOUT HERE]

4.3 Conditional mean and variance models

Daily log-differences of CDS spreads have more autocorrelation than is commonly found for daily stock returns (e.g., the average first-order autocorrelation is 0.161) and so the model for the conditional mean of our data needs more structure than the commonly-used constant model for daily stock returns. We use an AR(5) model, and augment it with a lag of the market variable (an

equal-weighted average of all 100 series) to capture any dynamic spillover effects. We show below that this model passes standard specification tests.

$$Y_{it} = \phi_{0i} + \sum_{j=1}^5 \phi_{ji} Y_{i,t-j} + \phi_{mi} Y_{m,t-1} + e_{it} \quad (16)$$

For the market return we use the same model (omitting, of course, a repeat of the first lag of the market return). We need a model for the market return as we use the residuals from the market return model in our conditional variance specification.

Our model for the conditional variance is the asymmetric volatility model of Glosten, *et al.* (1993), the “GJR-GARCH” model. The motivation for the asymmetry in this model is that “bad news” about a firm increases its future volatility more than good news. For stock returns, bad news comes in the form of a negative residual. For CDS spreads, on the other hand, bad news is a *positive* residual, and so we reverse the direction of the indicator variable in the GJR-GARCH model to reflect this. In addition to the standard GJR-GARCH terms, we also include terms relating to the lagged market residual:

$$\begin{aligned} V_{t-1}[e_{it}] &\equiv \sigma_{it}^2 = \omega_i + \beta_i \sigma_{i,t-1}^2 + \alpha_i e_{i,t-1}^2 + \delta_i e_{i,t-1}^2 \mathbf{1}\{e_{i,t-1} > 0\} \\ &\quad + \alpha_{im} e_{m,t-1}^2 + \delta_{im} e_{m,t-1}^2 \mathbf{1}\{e_{m,t-1} > 0\} \end{aligned} \quad (17)$$

Finally, we specify a model for the marginal distribution of the standardized residuals, η_{it} . We use the skewed t distribution of Hansen (1994), which allows for non-zero skewness and excess kurtosis:

$$\eta_{it} \equiv \frac{e_{it}}{\sigma_{it}} \sim iid \text{ Skew } t(\nu_i, \psi_i) \quad (18)$$

Table 3 summarizes the results of estimating the above models on the 100 time series. For the conditional mean model, we find strong significance of the first three AR lags, as well as the lagged market return. The conditional variance models reveal only mild statistical evidence of

asymmetry in volatility, however the point estimates suggest that “bad news” (a positive residual) increases future volatility about 50% more than good news. The average estimated degrees of freedom parameter is 3.6, suggestive of fat tails, and the estimated skewness parameter is positive for 94 firms, and significantly different from zero for 41 of these, indicating positive skewness.

We now discuss goodness-of-fit tests for the marginal distribution specifications. We firstly use the Ljung-Box test to check the adequacy of these models for the conditional mean and variance, and we are able to reject the null of zero autocorrelation up to the tenth lag for only nine of the residual series, and only two of the squared standardized residual series. We conclude that these models provide a satisfactory fit to the conditional means and variances of these series. Next, we use the Kolmogorov-Smirnov test to investigate the fit of the skewed t distribution for the standardized residuals, using 100 simulations to obtain critical values that capture the parameter estimation error, and we reject the null of correct specification for eleven of the 100 firms. This is slightly higher than the level of the test (0.05), but we do not pursue the use of a more complicated marginal distribution model for those eleven firms in the interests of parsimony and comparability.

[INSERT TABLE 3 ABOUT HERE]

4.4 The CDS “Big Bang”

On April 8, 2009, the CDS market underwent changes driven by a move towards more standardized CDS contracts. Details of these changes are described in Markit (2009). It is plausible that the changes to the CDS market around this so-called “Big Bang” changed the dynamics and distributional features of CDS time series, and we test for that possibility here. We do so by allowing the parameters of the mean, variance, and marginal distribution models to change on the date of the Big Bang, and we test the significance of these changes. We have 591 pre-break observations and 1053 post-break observations. (It is worth noting that the turmoil in financial markets in 2007-09 makes identifying any single event as a source of a structural break almost

impossible. We have good reason to suspect that the Big Bang led to a structural break, but it is possible that a better break date, in terms of model fit, is earlier or later than the date we use.)

We find that the conditional mean parameters changed significantly (at the 0.05 level) for 39 firms, and the conditional variance and marginal density shape parameters changed significantly for 66 firms. In what follows, the results we report are based on models that allow for a structural break in the mean, variance and distribution parameters. Given the prevalence of these changes, all of the copula models we consider allow for a break at the date of the Big Bang.

4.5 Comparing models for the conditional copula

The class of high-dimensional dynamic copula models described in Section 2 includes a variety of possible specifications: static vs. GAS dynamics; normal vs. skew t - t factor copulas; equidependence vs. block equidependence vs. heterogeneous dependence.

Table 4 presents results for six different dynamic models (a corresponding table for the six static copula models is available in the internet appendix). Bootstrap standard errors are presented in parentheses below the estimated parameters. (We use the stationary block bootstrap of Politis and Romano (1994) with an average block length of 120 days, applied to the log-difference of the CDS spreads, and we use 100 bootstrap replications.) Similar to other applications of GAS models (see, Creal *et al.* (2011, 2013)) we find strong persistence, with the β parameter ranging from 0.85 to 0.99. (Note that the β parameter in GAS models plays the same role as $\alpha + \beta$ in a GARCH(1,1) model, see Example 1 in Creal, *et al.* (2013)). We also find that the inverse degrees of freedom parameters are greater than zero (i.e., the factor copula is not Normal), which we test formally below. For the two more parsimonious models, we do not find the common factor to be significantly asymmetrically distributed, but for our preferred model based on heterogeneous dependence, we find that the asymmetry parameter for the common factor is positive, and significantly so in the post-break period, indicating greater dependence for joint upward moves in CDS spreads. This is consistent with financial variables being more correlated during bad times: for stock returns

bad times correspond to joint downward moves, which have been shown in past work to be more correlated than joint upward moves, while for CDS spreads bad times correspond to joint upward moves.

[INSERT TABLE 4 ABOUT HERE]

Table 4 shows that the estimated degrees of freedom parameter for the common factor is larger than that for the idiosyncratic term. Oh and Patton (2012) show that when these two parameters differ the tail dependence implied by this factor copula is on the boundary: either zero ($\nu_z > \nu_\varepsilon$) or one ($\nu_z < \nu_\varepsilon$); only when these parameters are equal can tail dependence lie inside $(0, 1)$. We test the significance of the difference between these two parameters by estimating a model with them imposed to be equal and then conducting a likelihood ratio test, the log-likelihoods from these two models are reported in Table 5. The results strongly suggest that $\nu_z > \nu_\varepsilon$, and thus that extreme movements in CDS spreads are uncorrelated. The average gain in the log likelihood from estimating just this one extra parameter is around 200 points. This does not mean, of course, that “near extreme” movements must be uncorrelated, only that they are uncorrelated in the limit.

Table 5 also presents a comparison of the Skew t - t factor copula with the Normal copula, which is obtained by using a Normal distribution for both the common factor and the idiosyncratic factor. We see very clearly that the Normal copula performs worse than the Skew t - t factor copula, with the average gain in the log likelihood of the more flexible model being over 2000 points. This represents yet more evidence against the Normal copula model for financial time series; the Normal copula is simply too restrictive.

[INSERT TABLE 5 ABOUT HERE]

Finally, Table 5 can be used to compare the results from models with three different degrees of heterogeneity: equidependence vs. block equidependence vs. heterogeneous dependence. We see that the data support the more flexible models, with the block equidependence model improving

on the equidependence model by around 200 points, and the heterogeneous model improving on the block equidependence model by around 800 points. It should be noted that our use of industry membership to form the “blocks” is just one method, and alternative grouping schemes may lead to better results. We do not pursue this possibility here.

Given the results in Table 5, our preferred model for the dependence structure of these 100 CDS spread series is a skew t - t factor copula, with separate degrees of freedom for the common and idiosyncratic variables, allowing for a separate loading on the common factor for each series (the “heterogeneous dependence” model) and allowing for dynamics using the GAS structure described in the previous section. Figure 2 presents the time-varying factor loadings implied by this model, and Figure 3 presents time-varying rank correlations. To summarize these results, Figure 2 averages the loadings across all firms in the same industry, and Figure 3 averages all pair-wise correlations between firms in the same pairs of industries. (Thus the plotted factor loadings and rank correlations are smoother than any individual rank correlation plot.) Also presented in Figure 3 are 60-day rolling window rank correlations, again averaged across pairs of the firms in the same pair of industries. This figure reveals a secular increase in the correlation between CDS spreads, rising from around 0.1 in 2006 to around 0.5 in 2013. Interestingly, rank correlations do *not* appear to spike during the financial crisis, unlike individual volatilities and probabilities of default; rather they continue a mostly steady rise through the sample period.

[INSERT FIGURES 2 AND 3 ABOUT HERE]

5 Time-varying systemic risk

In this section we use the dynamic multivariate model presented above to obtain estimates of measures of systemic risk. A variety of measures of systemic risk have been proposed in the literature to date. One influential measure is “CoVaR,” proposed by Adrian and Brunnermeier (2009), which uses quantile regression to estimate the lower tail (e.g., 0.05) quantile of market

returns conditional on a given firm having a return equal to its lower tail quantile. This measure provides an estimate of how firm-level stress spills over to the market index. An alternative measure is “marginal expected shortfall” (MES) proposed by Brownlees and Engle (2015), which estimates the expected return on a firm conditional on the market return being below some low threshold. Segoviano and Goodhart (2009) and Giesecke and Kim (2009) propose measuring systemic risk via the probability that a “large” number of firms are in distress, Lucas, *et al.* (2014) use the same measure applied to European sovereign debt. Hartmann, *et al.* (2006) consider a measure based on the probability of a large number of firms being in distress *conditional* on a subset of firms being in distress. Huang, *et al.* (2009) suggest using the price of a hypothetical contract insuring against system-wide distress, valued using a mix of CDS and equity data, as a measure of systemic risk. Schwaab (2010) and Bisias *et al.* (2012) present reviews of these and other measures of systemic risk. Giglio *et al.* (2015) consider a large collection of systemic risk measures, and show that an index based on these measures can help to predict lower quantiles of shocks to economic growth.

We consider two different estimates of systemic risk, defined in detail in the following two subsections. In all cases we use the dynamic copula model that performed best in the previous section, namely the heterogeneous dependence factor copula model.

5.1 Joint probability of distress

The first measure of systemic risk we implement is an estimate of the probability that a large number of firms will be in distress, similar to the measure considered by Segoviano and Goodhart (2009), Giesecke and Kim (2009) and Lucas, *et al.* (2014). We define distress as a firm’s one-year-ahead CDS spread lying above some threshold:

$$D_{i,t+250} \equiv \mathbf{1} \{ S_{i,t+250} > c_{i,t+250}^* \} \tag{19}$$

We choose the threshold as the 99% conditional quantiles of the individual CDS spreads. In our sample, the average 99% threshold corresponds to a CDS spread of 339 basis points. Using equation (13) above, this threshold yields an implied probability of default (assuming LGD is 0.6) of 5.7%. (The average CDS spread across all firms is 97 basis points, yielding an implied probability of default of 1.6%.) We also considered a threshold quantile of 0.95, corresponding to an average CDS spread of 245 basis points, and the results are qualitatively similar.

We use the probability of a large proportion of firms being in distress as a measure of systemic risk. We define the “joint probability of distress” as:

$$JPD_{t,k} \equiv \Pr_t \left[\left(\frac{1}{N} \sum_{i=1}^N D_{i,t+250} \right) \geq \frac{k}{N} \right] \quad (20)$$

where k is a user-chosen threshold for what constitutes a “large” proportion of the N firms. We use $k = 30$, and the results corresponding to $k = 20$ and $k = 40$ are qualitatively similar. The JPD is only estimable via simulations from our model, and we obtain these using 10,000 simulations. Given the computational burden, we compute estimates only every 20 trading days (approximately once per month).

The estimated joint probability of distress is presented in Figure 4. For comparison, we also include the JPD implied by a “baseline” copula model, the constant Normal copula. (This copula allows for a break at the Big Bang, and so is constant only within each sub-period.) We observe that the JPD implied by the factor copula with GAS dynamics spikes up in late 2007, roughly doubling in value. It rises even further in early 2008, reaching a peak of around 1%. This implies that at the height of the financial crisis, there was around a 1% chance that at least 30 of these 100 firms would be simultaneously in distress, defined as a CDS spread lying in its upper 1% tail. The joint probability of distress declines in mid 2009, though remains somewhat higher than before the crisis. This is consistent with results for systemic sovereign default risk in the U.S. and Europe reported in Ang and Longstaff (2013) and Lucas, *et al.* (2014).

Comparing the estimates of JPD from the two copula models, we see that the constant normal copula estimate is almost uniformly lower than that from the factor copula in the pre-crisis period, and substantially lower from mid 2007 until the Big Bang, when the copula parameter was allowed to change. In the second sub-sample the two estimates are closer, although the JPD implied by the normal copula is generally lower than from the factor copula. These two estimates of JPD reveal the large variation in estimates of systemic risk that can be obtained depending on the assumed shape of the copula, and the assumed form of the dynamics of the copula.

[INSERT FIGURE 4 ABOUT HERE]

5.2 Expected proportion in distress

Our second measure of systemic risk more fully exploits the ability of our dynamic copula model to capture heterogeneous dependence between individual CDS spread changes. For each firm i , we compute the expected proportion of firms in distress conditional on firm i being in distress:

$$EPD_{i,t} \equiv E_t \left[\frac{1}{N} \sum_{j=1}^N D_{j,t+250} \middle| D_{i,t+250} = 1 \right] \quad (21)$$

The minimum value this can take is $1/N$, as we include firm i in the sum, and the maximum is one. A version of this measure was proposed in Hartmann, *et al.* (2006). We use the same indicator for distress as in the previous section (equation (19)). This measure of systemic risk is similar in spirit to the CoVaR measure proposed by Adrian and Brunnermeier (2009), in that it looks at distress “spillovers” from a single firm to the market as a whole. (Note that the summation sign in equation (21) weights all firms equally; a simple extension would allow firms to be weighted by their market capitalization, debt-to-equity ratio, or any other measurable characteristic of a firm.)

In the upper panel of Figure 5 we summarize the results from the EPD estimates, and present the average, and 20% and 80% quantiles of this measure across the 100 firms in our sample. We observe that the average EPD is around 30% in the pre-crisis period, rising to almost 60% in late

2008, and returning to around 40% in the last year of our sample. Thus this figure, like the JPD plot in Figure 4, is also suggestive of a large increase in systemic risk around the financial crisis, and higher level of systemic risk in the current period than in the pre-crisis period.

[INSERT FIGURE 5 ABOUT HERE]

The expected proportion in distress measure enables us to identify firms that are more strongly correlated with market-wide distress than others. When the EPD is low for a given firm, it reveals that distress for that firm is not a signal of widespread distress, i.e., firm i is more idiosyncratic. Conversely, when the EPD is high, it reveals that distress for this firm is a sign of widespread distress, and so this firm is a “bellwether” for systemic risk. To illustrate the information from individual firm EPD estimates, Table 6 below presents the top five and bottom five firms according to their EPD on three dates in our sample period, the first day (January 2, 2006), a middle day (January 26, 2009) and the last day (April 17, 2012). We note that SLM Corporation (“Sallie Mae”, in the student loan business) appears in the “least systemic” group on all three dates, indicating that periods in which it is in distress are, according to our model, generally unrelated to periods of wider distress. Marsh and McLennan (which owns a collection of risk, insurance and consulting firms) and Baxter International (a bioscience and medical firm) each appear in the “most systemic” group for two out of three dates.

Table 6 also provides information on the spread of EPD estimates across firms. At the start of our sample the least systemic firms had EPDs of 2 to 3, indicating that only one to two other firms are expected to be in distress when they are in distress. At the end of our sample the least systemic firms had EPDs of 8 to 12, indicating a wider correlation of distress even among the least correlated. A similar finding is true for the most systemic firms: the EPDs for the most systemic firms rise from 48–53 at the start of the sample to 84–94 at the end. Thus there is a general increase in the correlation between firm distress over this sample period.

[INSERT TABLE 6 ABOUT HERE]

A particular focus in the growing literature on systemic risk is on spillovers of distress from the financial sector to the nonfinancial, or “real,” sector, see Adrian and Brunnermeier (2008) and Acharya, *et al.* (2010) for example. Our collection of 100 firms contains 16 financial firms and 84 nonfinancial firms, and we can examine the “expected proportion in distress” across these two classifications. The lower panel of Figure 5 presents estimates of the expected proportion of financial/nonfinancial firms in distress, conditional on one financial/nonfinancial firm being in distress. This generates four time series of EPD, and allows us to look at spillovers from one sector to the other, and spillovers within a sector. This panel reveals substantial heterogeneity in the sensitivity of firms to other firms being in distress. In particular, we see that a substantially higher proportion of financial firms are expected to be in distress conditional on a nonfinancial firm being in distress, than the other way around. At the peak of the financial crisis, these proportions were 80% and 40% respectively, and more recently have averaged around 60% and 30%. This might be taken as somewhat reassuring, as it indicates that while the probability of distress in the “real” sector does increase conditional on distress in the financial sector, the sensitivity is strongest in the other direction. We also note the relatively high sensitivity of financial firms to other financial firms being in distress, indicating some intra-industry contagion.

6 Conclusion

Motivated by the growing interest in measures of the risk of systemic events, this paper proposes new flexible yet parsimonious models for time-varying high-dimensional distributions. We use copula theory to combine well-known models for univariate distributions with new models of the conditional dependence structure (copula) to obtain dynamic joint distributions. Our proposed new dynamic copula models can be applied in dimensions of 100 or more, which is much greater than commonly considered in the literature. These models draw on successful ideas from the literature on dynamic modeling of high-dimensional correlation matrices, (e.g., Engle (2002)) and on recent work

on models for general time-varying distributions (Creal, *et al.* (2013)). We propose a “variance targeting” type estimator for this class of dynamic copulas to dramatically reduce the number of parameters to be estimated by numerical optimization.

We apply our model to a detailed analysis of a collection of 100 credit default swap (CDS) spreads on U.S. firms. The CDS market has expanded rapidly in recent years, and yields a novel view of the health of the underlying firms. We use our model of CDS spreads to provide insights into systemic risk, and we find, unsurprisingly, that systemic risk was highest during the financial crisis of 2008–09. More interestingly, we also find that systemic risk has remained relatively high, and is substantially higher now than in the pre-crisis period.

Appendix A: Proof of Proposition 1

Proof of Proposition 1. (i) The evolution equation for λ_{it} in equation (9) and stationarity of $\{\lambda_t\}$, which holds by assumption 1(b), implies

$$E[\log \lambda_{it}] = \omega_i + \beta E[\log \lambda_{i,t-1}] = \frac{\omega_i}{1 - \beta} \quad (22)$$

So $\omega_i = E[\log \lambda_{it}](1 - \beta)$, and we can re-write our GAS equation (9) in “variance targeting” form:

$$\log \lambda_{it} = E[\log \lambda_{it}](1 - \beta) + \beta \log \lambda_{i,t-1} + \alpha s_{i,t-1} \quad (23)$$

The objective of this proposition is to find an estimate of $E[\log \lambda_{it}]$ based on observable data, thus reducing the number of parameters to be estimated numerically. Next, note that linear correlations are given by:

$$\begin{aligned} \rho_{ij,X}^L &\equiv \text{Corr}[X_i, X_j] = \frac{\lambda_i \lambda_j}{\sqrt{(1 + \lambda_i^2)(1 + \lambda_j^2)}} \equiv g(\lambda_i, \lambda_j) \\ \text{and } \mathbf{R}_X^L &\equiv \text{Corr}[\mathbf{X}] = G(\boldsymbol{\lambda}) \end{aligned} \quad (24)$$

By assumption 1(a), this is an exactly- ($N = 3$) or over- ($N > 3$) identified system, as we have N parameters $\boldsymbol{\lambda} \equiv [\lambda_1, \dots, \lambda_N]'$ and $N(N - 1)/2$ correlations. By Assumption 1(d) we have a corresponding exactly- or over-identified system for the Spearman rank correlation matrix:

$$\mathbf{R}_X = \varphi(\mathbf{R}_X^L) = \varphi(G(\boldsymbol{\lambda})) \quad (25)$$

(In a slight abuse of notation, we let $\varphi(\mathbf{R}_X^L)$ map the entire linear correlation matrix to the rank correlation matrix.) Define the exponential of the inverse of the function $\varphi \circ G$ as H , so that $\log \boldsymbol{\lambda} = H(\boldsymbol{\rho}_X)$, where $\boldsymbol{\rho}_X \equiv \text{vech}(\mathbf{R}_X)$. The function H is not known in closed form but it can be obtained by a simple and fast optimization problem:

$$H(\boldsymbol{\rho}_X) = \arg \min_{\mathbf{a}} (\text{vech}\{\varphi(G(\exp \mathbf{a}))\} - \boldsymbol{\rho}_X)' (\text{vech}\{\varphi(G(\exp \mathbf{a}))\} - \boldsymbol{\rho}_X) \quad (26)$$

Note that under Assumption 1(a), there is no error in this optimization problem; this is just a means of recovering H from $\varphi \circ G$. This is the GMM analog to the usual method-of-moments estimator used in variance targeting. Under Assumption 1(c) the function $H(\boldsymbol{\rho}_X)$ is linear, so $E[\log \boldsymbol{\lambda}_t] = E[H(\boldsymbol{\rho}_{t,X})] = H(\bar{\boldsymbol{\rho}}_X)$, where $\bar{\boldsymbol{\rho}}_X \equiv E[\boldsymbol{\rho}_{t,X}]$. Finally, we exploit the fact that $\text{RankCorr}[\mathbf{X}]$ is identical to $\text{RankCorr}[\boldsymbol{\eta}]$ by Assumption 1(a) and Theorem 5.1.6 of Nelsen (2006). So we obtain $E[\log \boldsymbol{\lambda}_t] = H(\bar{\boldsymbol{\rho}}_X) = H(\bar{\boldsymbol{\rho}}_\eta)$.

(ii) We use as our “VT estimator” the sample analog of the above expression:

$$\widehat{\log \boldsymbol{\lambda}} = H(\hat{\boldsymbol{\rho}}_\eta) \quad (27)$$

First note that, since the marginal distributions of $\boldsymbol{\eta}_t$ are known, sample rank correlations are a linear functions of a sample moment, see Nelsen (2006, Chapter 5) for example:

$$\hat{\rho}_{ij,\eta} = -3 + \frac{12}{T} \sum_{t=1}^T F_i(\eta_{i,t}) F_j(\eta_{j,t}) \quad (28)$$

Our estimate of $E[\log \lambda_{it}]$ is obtained in equation (26) as:

$$\widehat{\log \boldsymbol{\lambda}} = \arg \min_{\mathbf{a}} \bar{\mathbf{m}}_T(\mathbf{a})' \bar{\mathbf{m}}_T(\mathbf{a}) \quad (29)$$

$$\text{where } \bar{\mathbf{m}}_T(\mathbf{a}) \equiv \text{vech}\{\varphi(G(\exp \mathbf{a}))\} - \hat{\boldsymbol{\rho}}_\eta^S$$

The element of $\bar{\mathbf{m}}_T$ corresponding to the (i, j) element of the correlation matrix is:

$$\bar{m}_T^{(i,j)}(\mathbf{a}) = [\varphi(G(\exp \mathbf{a}))]_{(i,j)} + 3 - \frac{12}{T} \sum_{t=1}^T F_i(\eta_{i,t}) F_j(\eta_{j,t}) \quad (30)$$

Thus $\widehat{\log \boldsymbol{\lambda}}$ is a standard GMM estimator for $N \geq 3$. ■

Appendix B: Implementation details

B.1: Obtaining a factor copula likelihood

The factor copula introduced in Oh and Patton (2012) does not have a likelihood in closed form, but it is relatively simple to obtain the likelihood using numerical integration. Consider the factor structure in equation (5) and (6). Our objective is to obtain the copula density of \mathbf{X}_t :

$$\mathbf{c}_t(u_1, \dots, u_N) = \frac{\mathbf{g}_t(G_{1t}^{-1}(u_1), \dots, G_{Nt}^{-1}(u_N))}{g_{1t}(G_{1t}^{-1}(u_1)) \cdots g_{Nt}(G_{Nt}^{-1}(u_N))} \quad (31)$$

where $\mathbf{g}_t(x_1, \dots, x_N)$ is the joint density of \mathbf{X}_t , $g_{it}(x_i)$ is the marginal density of X_{it} , and $\mathbf{c}_t(u_1, \dots, u_N)$ is the copula density. To construct the copula density, we need each of the functions that appear on the right-hand side above: $g_{it}(x_i)$, $G_{it}(x_i)$, $\mathbf{g}_t(x_1, \dots, x_N)$ and $G_{it}^{-1}(u_i)$.

The independence of Z and ε_i implies that:

$$\begin{aligned}
f_{X_i|Z,t}(x_i|z) &= f_{\varepsilon_i}(x_i - \lambda_{it}z) \\
F_{X_i|Z,t}(x_i|z) &= F_{\varepsilon_i}(x_i - \lambda_{it}z) \\
\mathbf{g}_{\mathbf{X}|Z,t}(x_1, \dots, x_N|z) &= \prod_{i=1}^N f_{\varepsilon_i}(x_i - \lambda_{it}z)
\end{aligned} \tag{32}$$

With these conditional distributions, one dimensional integration gives the marginals:

$$g_{it}(x_i) = \int_{-\infty}^{\infty} f_{X_i,Z,t}(x_i, z) dz = \int_{-\infty}^{\infty} f_{X_i|Z,t}(x_i|z) dF_{Zt}(z) = \int_{-\infty}^{\infty} f_{\varepsilon_i}(x_i - \lambda_{it}z) dF_{Zt}(z) \tag{33}$$

and similarly

$$\begin{aligned}
G_{it}(x_i) &= \int_{-\infty}^{\infty} F_{\varepsilon_i}(x_i - \lambda_{it}z) dF_{Zt}(z) \\
\mathbf{g}_t(x_1, \dots, x_N) &= \int_{-\infty}^{\infty} \prod_{i=1}^N f_{\varepsilon_i}(x_i - \lambda_{it}z) dF_{Zt}(z)
\end{aligned} \tag{34}$$

We use a change of variables, $u \equiv F_{Zt}(z)$, to convert these to bounded integrals:

$$\begin{aligned}
g_{it}(x_i) &= \int_0^1 f_{\varepsilon_i}(x_i - \lambda_{it}F_{Zt}^{-1}(u)) du \\
G_{it}(x_i) &= \int_0^1 F_{\varepsilon_i}(x_i - \lambda_{it}F_{Zt}^{-1}(u)) du \\
\mathbf{g}_t(x_1, \dots, x_N) &= \int_0^1 \prod_{i=1}^N f_{\varepsilon_i}(x_i - \lambda_{it}F_{Zt}^{-1}(u)) du
\end{aligned} \tag{35}$$

Thus the factor copula density requires the computation of just one-dimensional integrals. (For a factor copula with J common factors the integral would be J -dimensional.) We use Gauss-Legendre quadrature for the integration, using Q ‘‘nodes,’’ (see Judd (1998) for details) and we choose Q on the basis of a small simulation study described below.

Finally, we need a method to invert $G_{it}(x_i)$, and note from above that this is a function of both

x and the factor loading λ_{it} , with $G_{it} = G_{js}$ if $\lambda_{it} = \lambda_{js}$. We estimate the inverse of G_{it} by creating a grid of 100 points for x in the interval $[x_{\min}, x_{\max}]$ and 50 points for λ in the interval $[\lambda_{\min}, \lambda_{\max}]$, and then evaluating G at each of those points. We then use two-dimensional linear interpolation to obtain $G^{-1}(u; \lambda)$ given u and λ . This two-dimensional approximation substantially reduces the computational burden, especially when λ is time-varying, as we can evaluate the function G prior to estimation, rather than re-estimating it for each likelihood evaluation.

We conducted a small Monte Carlo simulation to evaluate the accuracy of the numerical approximations for G and G^{-1} . We use quadrature nodes $Q \in \{10, 50, 150\}$ and $[x_{\min}, x_{\max}] = [-30, 30]$, $[\lambda_{\min}, \lambda_{\max}] = [0, 6]$ for the numerical inversion. For this simulation, we considered the factor copula implied by the following structure:

$$X_i = \lambda_0 Z_t + \varepsilon_i, \quad i = 1, 2 \tag{36}$$

$$\text{where } Z_t \sim \text{Skew } t(\nu_0, \psi_0), \quad \varepsilon_{it} \sim \text{iid } t(\nu_0), \quad Z_t \perp \varepsilon_i \quad \forall i$$

where $\lambda_0 = 1$, $\nu_0^{-1} = 0.25$ and $\psi_0 = -0.5$. At each replication, we simulate $\mathbf{X} = [X_1, X_2]$ 1000 times, and apply the empirical distribution function to transform \mathbf{X} to $\mathbf{U} = [U_1, U_2]$. With this $[U_1, U_2]$ we estimate $[\lambda, \nu^{-1}, \psi]$ by our numerically approximated maximum likelihood method.

Table S3 in the internet appendix contains estimation results for 100 replications. We find that estimation with only 10 nodes introduces a relatively large bias, in particular for ν^{-1} , consistent with this low number of nodes providing a poor approximation of the tails of this density. Estimation with 50 nodes gives accurate results, and is comparable to those with 150 nodes in that bias and standard deviation are small.

Table S4 in the internet appendix studies the Normal factor copula (which can be obtained from the previous equation when $\nu_0 \rightarrow \infty$ and $\psi_0 = 0$) for which we have a closed-form likelihood. We compare estimating λ based on numerical integration with 10, 50 and 150 nodes against using the closed-form likelihood (corresponding to ∞ nodes). We again see that using only 10 nodes leads to

a decrease in accuracy, while 50 and 150 nodes give approximately equally-accurate results, that are just slightly worse than using the analytical likelihood. We thus use 50 nodes throughout the paper.

B.2: “Variance targeting” assumptions

We investigate the plausibility of Assumptions 1(c) and 1(d) via simulations. We use 50,000 observations to estimate the true (unknown) mappings. Note that while the copula is time-varying, we only need to study the mapping for a given day (set of shape parameters), and so we do not consider dynamics in this simulation study. Moreover, the mappings are pair-wise, and so we need only consider the (i, j) bivariate case. We consider three different factor copulas: (i) $\nu_z = \nu_\varepsilon = 5, \psi = 0.1$ (ii) $\nu_z = \nu_\varepsilon = 4, \psi = 0.25$ (iii) $\nu_z = \nu_\varepsilon = \infty, \psi = 0$ (corresponding to the Normal copula). We fix $\lambda_i = \lambda_j = \lambda$, and let λ vary so that the model-implied rank correlation ranges from 0.1 to 0.7, which covers the range of pair-wise rank correlations we observe in our data.

The results are summarized in Figure 1 in the internet appendix. The left panel of this figure reveals that the mapping from rank correlation to linear correlation changes only slightly with the shape parameters $(\nu_z, \nu_\varepsilon, \psi)$, and in all cases the function is strictly increasing, supporting Assumption 1(d). In fact, we observe that for all cases the function is close to being the identity function, and we invoke this approximation in our estimation to increase computational speed. The right panel of the figure plots the mapping from rank correlation to log factor loadings, and shows that the true mapping is reasonably approximated by a straight line, particularly for values of rank correlation near the sample average rank correlation in our application, which is around 0.4, supporting Assumption 1(c). Violations of either of these assumptions, if large relative to sampling variability and other sources of estimation error, would manifest in poor performance of the estimation of the “heterogeneous dependence” model. As discussed in Section 3, the simulation results in Table 1 provide evidence that this estimation has good finite-sample properties.

References

- [1] Acharya, V.V., L.H. Pedersen, T. Philippon and M.P. Richardson, 2010, Measuring Systemic Risk, working paper, NYU Stern.
- [2] Adrian, T. and M.K. Brunnermeier, 2009, CoVaR, Staff Report 348, Federal Reserve Bank of New York.
- [3] Almeida, C., C. Czado, and H. Manner, 2012, Modeling high dimensional time-varying dependence using D-vine SCAR models, working paper, <http://arxiv.org/pdf/1202.2008v1.pdf>.
- [4] Ang, A. and F. A. Longstaff, 2013, Systemic sovereign credit risk: Lessons from the U.S. and Europe, *Journal of Monetary Economics*, 60, 493-510.
- [5] Barclays Capital, 2010, *Standard Corporate CDS Handbook*.
- [6] Bisias, D., M. Flood, A. W. Lo and S. Valavanis, 2012, A Survey of Systemic Risk Analytics, Office of Financial Research Working Paper 0001.
- [7] Blasques, F., S.J. Koopman and A. Lucas, 2014a, Maximum Likelihood Estimation for Generalized Autoregressive Score Processes, Tinbergen Institute Discussion Paper TI 2014-029/111.
- [8] Blasques, F., S.J. Koopman and A. Lucas, 2014b, Stationarity and Ergodicity of Univariate Generalized Autoregressive Score Processes, *Electronic Journal of Statistics*, forthcoming.
- [9] Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.
- [10] Brenner, R.J., R.H. Harjes and K.F. Kroner, 1996, Another look at models of the short-term interest rate, *Journal of Financial and Quantitative Analysis*, 31, 85-107.
- [11] Brownlees, C.T. and R.F. Engle, 2015, SRISK: A Conditional Capital Shortfall Index for Systemic Risk Measurement, working paper, NYU Stern.
- [12] Carr, P. and L. Wu, 2011, A Simple Robust Link Between American Puts and Credit Protection, *Review of Financial Studies*, 24(2), 473-505.
- [13] Chen, X. and Y. Fan, 2006, Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification, *Journal of Econometrics*, 135, 125-154.
- [14] Christoffersen, P., V. Errunza, K. Jacobs and H. Langlois, 2012, Is the potential for international diversification disappearing?, *Review of Financial Studies*, 25, 3711-3751.
- [15] Christoffersen, P., K. Jacobs, X. Jin and H. Langlois, 2013, Dynamic Dependence in Corporate Credit, working paper, Bauer College of Business, University of Houston.
- [16] Conrad, J., R.F. Dittmar and A. Hameed, 2011, Cross-Market and Cross-Firm Effects in Implied Default Probabilities and Recovery Values, working paper, Kenan-Flagler Business School, University of North Carolina.

- [17] Creal, D.D., S.J. Koopman, and A. Lucas, 2011, A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations, *Journal of Business and Economic Statistics*, 29 (4), 552–563.
- [18] Creal, D.D., S.J. Koopman, and A. Lucas, 2013, Generalized Autoregressive Score Models with Applications, *Journal of Applied Econometrics*, 28(5), 777-795.
- [19] Creal, D.D. and R.S. Tsay, 2014, High Dimensional Dynamic Stochastic Copula Models, *Journal of Econometrics*, forthcoming.
- [20] Creal, D.D., R.B. Gramercy and R.S. Tsay, 2014, Market-based Credit Ratings, *Journal of Business and Economic Statistics*, 32(3), 430-444.
- [21] Davis, R. A., W. T. M. Dunsmuir, and S. Streett, 2003, Observation driven models for Poisson counts, *Biometrika*, 90(4), 777–790.
- [22] De Lira Salvatierra, I.A. and A.J. Patton, 2014, Dynamic Copula Models and High Frequency Data, *Journal of Empirical Finance*, forthcoming.
- [23] Duffie, D. and K.J. Singleton, 2003, *Credit Risk: Pricing, Measurement, and Management*, Princeton University Press, New Jersey.
- [24] Engle, R.F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation, *Econometrica*, 50, 987-1007.
- [25] Engle, R.F., 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business & Economic Statistics*, 20(3), 339–350.
- [26] Engle, R.F. and B. Kelly, 2012, Dynamic Equicorrelation, *Journal of Business and Economic Statistics*, 30(2), 212-228.
- [27] Engle, R.F. and J. Mezrich, 1996, GARCH for Groups. *Risk*, 9, 36–40.
- [28] Engle, R.F. and J.R. Russell, 1998, Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica*, 66, 1127-1162.
- [29] Engle, R.F. and K. Sheppard, 2001, Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, working paper, University of California, San Diego.
- [30] Giesecke, K. and B. Kim, 2011, Systemic risk: What defaults are telling us, *Management Science*, 57(8), 1387–1405.
- [31] Giglio, S., B. Kelly and S. Pruitt, 2015, Systemic Risk and the Macroeconomy: An Empirical Evaluation, *Journal of Financial Economics*, forthcoming.
- [32] Glosten, L.R., R. Jagannathan, D.E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48(5), 1779-1801.
- [33] Gonçalves, S., U. Hounyo, A.J. Patton and K. Sheppard, 2013, Bootstrapping two-stage extremum estimators, working paper, Oxford-Man Institute of Quantitative Finance.

- [34] Hafner, C.M. and H. Manner, 2012, Dynamic Stochastic Copula Models: Estimation, Inference and Applications, *Journal of Applied Econometrics*, 27(2), 269-295.
- [35] Hansen, B.E., 1994, Autoregressive conditional density estimation, *International Economic Review*, 35(3), 705-730.
- [36] Hartmann, P., S. Straetmans, and C. de Vries, 2006, Banking system stability: A cross-atlantic perspective, in M. Carey and R.M. Stulz (eds), *The Risks of Financial Institutions*, University of Chicago Press, Chicago, 133-192.
- [37] Harvey, A.C., 2013, *Dynamic Models for Volatility and Heavy Tails*, Econometric Society Monograph 52, Cambridge University Press, Cambridge.
- [38] Harvey, A.C. and G. Sucarrat, 2014, EGARCH models with fat tails, skewness and leverage, *Computational Statistics and Data Analysis*, 76, 320-339.
- [39] Huang, X., H. Zhou and H. Zhu, 2009, A framework for assessing the systemic risk of major financial institutions, *Journal of Banking and Finance*, 33(11), 2036-2049.
- [40] Hull, J.C., 2012, *Risk Management and Financial Institutions*, Third Edition, John Wiley & Sons, New Jersey.
- [41] Jondeau, E. and M. Rockinger, 2006, The copula-GARCH model of conditional dependencies: An international stock market application, *Journal of International Money and Finance*, 25, 827-853.
- [42] Judd, K.L, 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, USA.
- [43] Liu, L.Y., 2014, Estimating Loss Given Default from CDS under Weak Identification, working paper, Duke University.
- [44] Lucas, A., B. Schwaab and X. Zhang, 2014, Conditional Euro Area Sovereign Default Risk, *Journal of Business and Economic Statistics*, 32(2), 271-284.
- [45] Manner, H. and J. Segers, 2011, Tails of correlation mixtures of elliptical copulas, *Insurance: Mathematics and Economics*, 48(1), 153-160.
- [46] Markit Group, 2009, The CDS Big Bang: Understanding the Changes to the Global CDS Contract and North American Conventions, research report.
- [47] Nelsen, R.B., 2006, *An Introduction to Copulas*, Second Edition, Springer, U.S.A.
- [48] Oh, D.H. and A.J. Patton, 2012, Modelling Dependence in High Dimensions with Factor Copulas, working paper, Duke University.
- [49] Oh, D.H. and A.J. Patton, 2013, Simulated method of moments estimation for copula-based multivariate models, *Journal of the American Statistical Association*, 108(502), 689-700.
- [50] Patton, A.J., 2006, Modelling asymmetric exchange rate dependence, *International Economic Review*, 47(2), 527-556.

- [51] Patton, A.J., 2013, Copula methods for forecasting multivariate time series, in G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 2, Elsevier, Oxford.
- [52] Politis, D.N., and J.P. Romano, 1994, The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.
- [53] Rémillard, B., 2010, Goodness-of-fit tests for copulas of multivariate time series, working paper.
- [54] Schwaab, B., 2010, New Quantitative Measures of Systemic Risk, *Financial Stability Review Special Feature E*, European Central Bank.
- [55] Segioviano, M.A. and C. Goodhart, 2009, Banking stability measures, IMF Working Paper WP/09/4.
- [56] Shephard, N., 2005, *Stochastic Volatility: Selected Readings*, Oxford University Press, Oxford.
- [57] Stöber, J. and C. Czado, 2012, Detecting regime switches in the dependence structure of high dimensional financial data, working paper, Technische Universität München.
- [58] White, H., 1994, *Estimation, Inference and Specification Analysis*, Econometric Society Monographs No. 22, Cambridge University Press, Cambridge.

Table 1: Simulation results

	True	Bias	Std	Median	90%	10%	Diff (90%-10%)
<i>Panel A: Equidependence</i>							
ω	0.000	0.005	0.015	0.001	0.027	-0.003	0.030
α	0.050	0.000	0.003	0.050	0.051	0.048	0.003
β	0.980	0.002	0.004	0.980	0.989	0.979	0.010
ν^{-1}	0.200	0.001	0.006	0.200	0.206	0.195	0.010
ψ_z	0.100	0.005	0.017	0.100	0.118	0.097	0.021
<i>Panel B: Block equidependence</i>							
ω_1	-0.030	0.000	0.005	-0.030	-0.025	-0.035	0.010
ω_2	-0.023	-0.001	0.004	-0.024	-0.020	-0.030	0.010
ω_3	-0.017	0.000	0.005	-0.017	-0.011	-0.023	0.012
ω_4	-0.010	0.000	0.004	-0.011	-0.005	-0.016	0.011
ω_5	-0.003	0.001	0.004	-0.002	0.004	-0.007	0.011
ω_6	0.003	0.001	0.004	0.004	0.009	0.000	0.009
ω_7	0.010	0.002	0.005	0.012	0.018	0.007	0.012
ω_8	0.017	0.001	0.005	0.017	0.025	0.012	0.013
ω_9	0.023	0.001	0.005	0.024	0.030	0.018	0.012
ω_{10}	0.030	0.003	0.006	0.033	0.040	0.024	0.015
α	0.050	0.001	0.005	0.051	0.057	0.045	0.012
β	0.980	-0.001	0.002	0.978	0.981	0.976	0.004
ν^{-1}	0.200	-0.005	0.008	0.196	0.202	0.184	0.018
ψ_z	0.100	0.004	0.025	0.103	0.138	0.071	0.068

Notes: This table presents results from the simulation study described in Section 3. Panel A contains results for the equidependence model and Panel B for the “block equidependence” model.

Table 1: Simulation results (continued)

	True	Bias	Std	Median	90%	10%	Diff (90%-10%)
<i>Panel C: Heterogeneous dependence</i>							
ω_1	-0.030	0.004	0.017	-0.022	-0.005	-0.052	0.047
ω_5	-0.028	0.004	0.016	-0.020	-0.005	-0.046	0.041
ω_{10}	-0.025	0.002	0.016	-0.019	-0.005	-0.041	0.036
ω_{15}	-0.022	0.002	0.013	-0.019	-0.004	-0.043	0.039
ω_{20}	-0.019	0.002	0.011	-0.015	-0.003	-0.033	0.030
ω_{25}	-0.016	0.000	0.010	-0.014	-0.003	-0.030	0.027
ω_{30}	-0.012	0.001	0.008	-0.010	-0.002	-0.022	0.020
ω_{35}	-0.009	0.000	0.008	-0.008	-0.002	-0.020	0.018
ω_{40}	-0.006	-0.001	0.005	-0.006	-0.002	-0.015	0.014
ω_{45}	-0.003	-0.001	0.005	-0.003	0.000	-0.010	0.010
ω_{50}	0.000	-0.002	0.004	-0.002	0.001	-0.007	0.008
ω_{55}	0.003	-0.001	0.004	0.001	0.007	-0.003	0.010
ω_{60}	0.006	-0.002	0.005	0.003	0.010	0.000	0.010
ω_{65}	0.009	-0.002	0.006	0.005	0.013	0.000	0.013
ω_{70}	0.012	-0.004	0.007	0.007	0.017	0.001	0.016
ω_{75}	0.015	-0.004	0.008	0.009	0.019	0.002	0.017
ω_{80}	0.018	-0.004	0.009	0.012	0.026	0.002	0.024
ω_{85}	0.021	-0.006	0.011	0.014	0.032	0.002	0.030
ω_{90}	0.024	-0.006	0.012	0.016	0.036	0.003	0.033
ω_{95}	0.027	-0.006	0.014	0.018	0.040	0.004	0.036
ω_{100}	0.030	-0.007	0.016	0.021	0.040	0.004	0.036
α	0.050	-0.006	0.015	0.045	0.062	0.023	0.039
β	0.980	0.002	0.012	0.983	0.997	0.966	0.031
ν^{-1}	0.200	-0.002	0.009	0.199	0.209	0.186	0.023
ψ_z	0.100	0.008	0.032	0.111	0.152	0.064	0.088

Notes: This table presents results from the simulation study described in Section 3. Panel C contains results for the “heterogeneous dependence” model. In the interests of space, Panel C only reports every fifth intercept parameter (ω_i). A table with all 100 intercept parameters is available in the internet appendix.

**Table 2: Summary statistics for daily CDS spreads
and log-differences of daily CDS spreads**

	Mean	5%	25%	Median	75%	95%
<i>Panel A: Cross-sectional distribution of CDS spreads</i>						
Mean	96.953	37.212	53.561	74.957	123.785	200.346
Std dev	69.950	17.344	27.245	47.508	84.336	180.618
1 st -order autocorrel	0.996	0.992	0.995	0.997	0.998	0.998
Skewness	1.203	0.095	0.695	1.280	1.587	2.488
Kurtosis	5.113	2.198	2.943	4.937	6.477	9.486
5%	23.883	9.021	11.741	18.926	29.851	60.538
25%	42.274	20.373	25.212	35.314	47.473	104.704
Median	85.310	35.098	50.105	69.399	113.762	166.208
75%	122.061	46.250	65.862	93.622	154.729	251.112
95%	245.497	72.514	102.554	168.500	313.585	631.924
99%	338.676	80.414	122.885	231.295	435.224	827.098
<i>Panel B: Cross-sectional distribution of log-differences of CDS spreads</i>						
Mean	5.589	-1.634	2.559	5.529	8.521	13.817
Std dev	378.892	308.636	347.627	373.460	400.385	476.533
1 st -order autocorrel	0.161	0.030	0.121	0.164	0.217	0.267
Skewness	1.087	-0.285	0.354	0.758	1.488	3.629
Kurtosis	25.531	7.717	10.286	14.557	25.911	74.843
5%	-514.574	-622.282	-551.334	-509.554	-474.027	-415.651
25%	-144.195	-172.319	-155.635	-145.415	-134.820	-111.993
Median	-2.324	-9.045	-3.644	-0.726	0.000	0.000
75%	132.127	95.168	120.514	131.019	144.363	174.645
95%	570.510	457.775	537.093	568.331	612.769	684.984
<i>Panel C: Autocorrelation in CDS spreads</i>						
	# of rejections					
	Level	Log-diff	Squared log-diff			
ADF test of unit root	12	100	–			
LB test for autocorrel	–	98	89			

Notes: Panels A and B of this table presents summary statistics across the $N = 100$ marginal distributions of daily CDS spreads (Panel A) and the log-differences of CDS spreads (Panel B), measured in basis points in both cases. The columns present the mean and quantiles from the cross-sectional distribution of the measures listed in the rows. Panel C shows the number of rejections (at the 0.05 level) across the 100 firms for augmented Dickey-Fuller tests of the null of a unit root, as well as Ljung-Box tests for autocorrelation up to 10 lags.

Table 3: Marginal distribution parameter estimates

	<i>Cross-sectional distribution</i>					
	Mean	5%	25%	Median	75%	95%
ϕ_0	3.029	-3.760	0.247	3.116	5.861	10.165
ϕ_1	0.005	-0.179	-0.062	0.010	0.082	0.153
ϕ_2	0.025	-0.039	-0.001	0.025	0.050	0.084
ϕ_3	-0.002	-0.058	-0.028	-0.004	0.021	0.064
ϕ_4	0.006	-0.046	-0.014	0.006	0.033	0.054
ϕ_5	0.004	-0.055	-0.022	0.005	0.027	0.060
ϕ_m	0.387	0.163	0.303	0.372	0.480	0.638
$\omega \times 10^4$	5.631	1.401	3.111	5.041	7.260	13.381
β	0.741	0.595	0.699	0.746	0.794	0.845
α	0.114	0.052	0.087	0.106	0.141	0.181
δ	0.022	0.000	0.000	0.000	0.042	0.086
α_m	0.223	0.037	0.137	0.206	0.297	0.494
δ_m	0.072	0.000	0.000	0.059	0.114	0.233
ν	3.620	2.877	3.293	3.571	3.921	4.496
ψ	0.043	-0.003	0.024	0.042	0.062	0.089
					<u># of rejections</u>	
LB test for standardized residuals					9	
LB test for squared standardized residuals					2	
KS test of skew t dist of standardized residuals					11	

Notes: The table presents summaries of the estimated AR(5)-GJR-GARCH(1,1)-Skew $t(\nu, \psi)$ models estimated on log-difference of daily CDS spreads. The columns present the mean and quantiles from the cross-sectional distribution of the parameters listed in the rows. The bottom panel shows the number of rejections (at the 0.05 level) across 100 firms from Ljung-Box tests for serial correlation up to 10 lags. The first row is for standardized residuals of log-difference of daily CDS spreads and the second row for squared standardized residuals. The bottom panel shows the number of rejections across 100 firms from the Kolmogorov–Smirnov test of the Skew $t(\nu, \psi)$ distribution used for the standardized residuals.

Table 4: Model estimation results

	Equidependence			Block equidependence			Heterogeneous Dependence				
	Normal		Factor	Normal		Factor	Normal		Factor		
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	
$\omega_{1 \rightarrow G}$	-0.013 (0.047)	-0.027 (0.033)	0.005 (0.009)	-0.010 (0.017)	-	-	-	-	-	-	
α_{GAS}	0.024 (0.009)	0.022 (0.003)	0.026 (0.003)	0.026 (0.003)	0.036 (0.015)	0.029 (0.001)	0.033 (0.011)	0.026 (0.010)	0.144 (0.062)	0.170 (0.022)	0.171 (0.015)
β_{GAS}	0.988 (0.060)	0.847 (0.070)	0.994 (0.026)	0.907 (0.035)	0.996 (0.044)	0.976 (0.042)	0.998 (0.015)	0.992 (0.031)	0.994 (0.065)	0.994 (0.031)	0.982 (0.040)
ν_z^{-1}	-	-	0.095 (0.025)	0.078 (0.031)	-	-	0.010 (0.029)	0.010 (0.020)	-	0.010 (0.027)	0.020 (0.021)
ν_ϵ^{-1}	-	-	0.170 (0.009)	0.178 (0.010)	-	-	0.206 (0.011)	0.190 (0.009)	-	0.204 (0.014)	0.179 (0.010)
ψ_z	-	-	-0.018 (0.049)	-0.015 (0.049)	-	-	0.088 (0.093)	0.093 (0.056)	-	0.091 (0.074)	0.124 (0.058)
$\log \mathcal{L}$	38,395		40,983		38,518		41,165		39,361		41,913
AIC	-76,778		-81,942		-77,008		-82,290		-78,314		-83,406
BIC	-76,771		-81,927		-76,991		-82,266		-78,066		-83,151

Notes: This table presents parameter estimates for two versions of the factor copula (Normal and Skew t - t), each with one of three degrees of heterogeneity of dependence (equidependence, block equidependence, and heterogeneous dependence). Standard errors, that account for the multi-stage estimation, based on the stationary bootstrap of Politis and Romano (1994) are presented below the estimated parameters. All models are allowed to have a structural break on April 8, 2009 (see Section 4.4), and we denote parameters from the first and second sub-samples as “Pre” and “Post.” The log-likelihood at the estimated parameters and the Akaike and Bayesian Information criteria are presented in the bottom three rows. The intercept parameters (ω_i) for the block equidependence and heterogeneous dependence models are not reported to conserve space.

Table 5: Model comparison results

		Normal		Factor (ν, ψ_z, ν)		Factor $(\nu_z, \psi_z, \nu_\varepsilon)$	
		Static	GAS	Static	GAS	Static	GAS
<i>Equi-dependence</i>	# param	2	6	6	10	8	12
	$\log \mathcal{L}$	36,185	38,395	39,223	40,688	39,508	40,983
	AIC	-72,366	-76,778	-78,434	-81,356	-79,000	-81,942
	BIC	-72,364	-76,771	-78,427	-81,344	-78,990	-81,927
<i>Block equi-dependence</i>	# param	10	14	14	18	16	20
	$\log \mathcal{L}$	36,477	38,518	39,441	40,799	39,757	41,165
	AIC	-72,934	-77,008	-78,854	-81,562	-79,482	-82,290
	BIC	-72,922	-76,991	-78,837	-81,540	-79,463	-82,266
<i>Heterogeneous dependence</i>	# param	200	204	204	208	206	210
	$\log \mathcal{L}$	37,652	39,361	40,357	41,522	40,628	41,913
	AIC	-74,904	-78,314	-80,306	-82,628	-80,844	-83,406
	BIC	-74,661	-78,066	-80,058	-82,375	-80,594	-83,151

Notes: This table presents the log-likelihood at the estimated parameters, as well as the Akaike and Bayesian Information criteria, for a variety of copula models. The preferred model according to each of these criteria is highlighted in bold. Also presented is the number of estimated parameters; note that this accounts for the fact that we allow for a structural break in these parameters, and so the number reported is twice as large as it would be in the absence of a break. We consider models with three degrees of heterogeneity of dependence (equidependence, block equidependence, and heterogeneous dependence); with and without dynamics (static and GAS); and three versions of the factor copula (Normal, Skew t - t with a common degrees of freedom parameter, and Skew t - t with separately estimated degrees of freedom parameters).

Table 6: Estimates of systemic risk

	<i>2 January 2006</i>		<i>26 January 2009</i>		<i>17 April 2012</i>	
	EPD	Firm	EPD	Firm	EPD	Firm
Most systemic	53	Marsh & McLennan	78	Lockheed Martin	94	Wal-Mart
2	50	Hewlett-Packard	77	Campbell Soup	88	Baxter Int'l
3	50	IBM	75	Marsh & McLennan	88	Walt Disney
4	49	Valero Energy	75	Baxter Int'l	87	Home Depot
5	48	Bristol-Myers Squibb	74	Goodrich	84	McDonald's
96	3	Aetna	35	Vornado Realty	12	MetLife
97	3	Xerox	34	Gen Elec Capital	11	The GAP
98	3	Comp Sci Corp	34	Johnson Controls	11	Sallie Mae
99	3	Sallie Mae	34	Alcoa	11	Comp Sci Corp
Least systemic	2	Freeport-McMoRan	33	Sallie Mae	8	Pitney Bowes

Note: This table presents the five firms with the largest estimated expected proportion in distress (EPD), defined in equation (21), and the five firms with the smallest EPD, for three dates in our sample period.

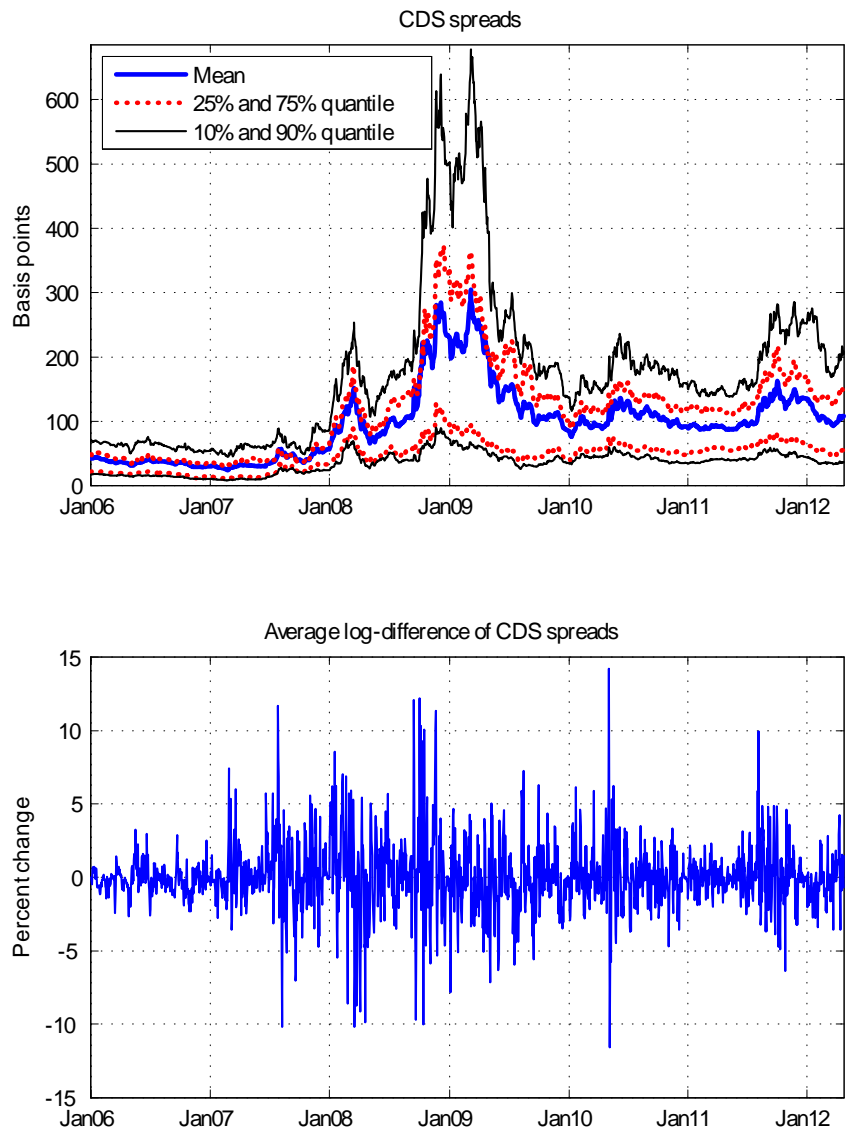


Figure 1: *The upper panel plots the mean and 10%, 25%, 75% and 90% quantiles across the CDS spreads for 100 U.S. firms over the period January 2006 to April 2012. The lower panel reports the average (across firms) percent change in CDS spreads for the same time period.*

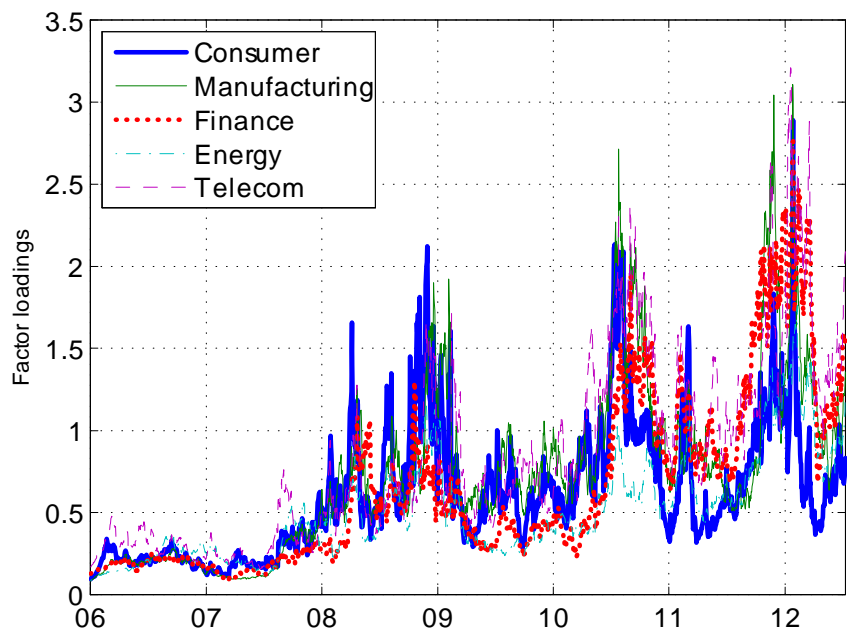


Figure 2: *This figure plots the estimated factor loadings (λ_t) from the heterogeneous dependence factor copula model, averaged across firms in the same industry.*

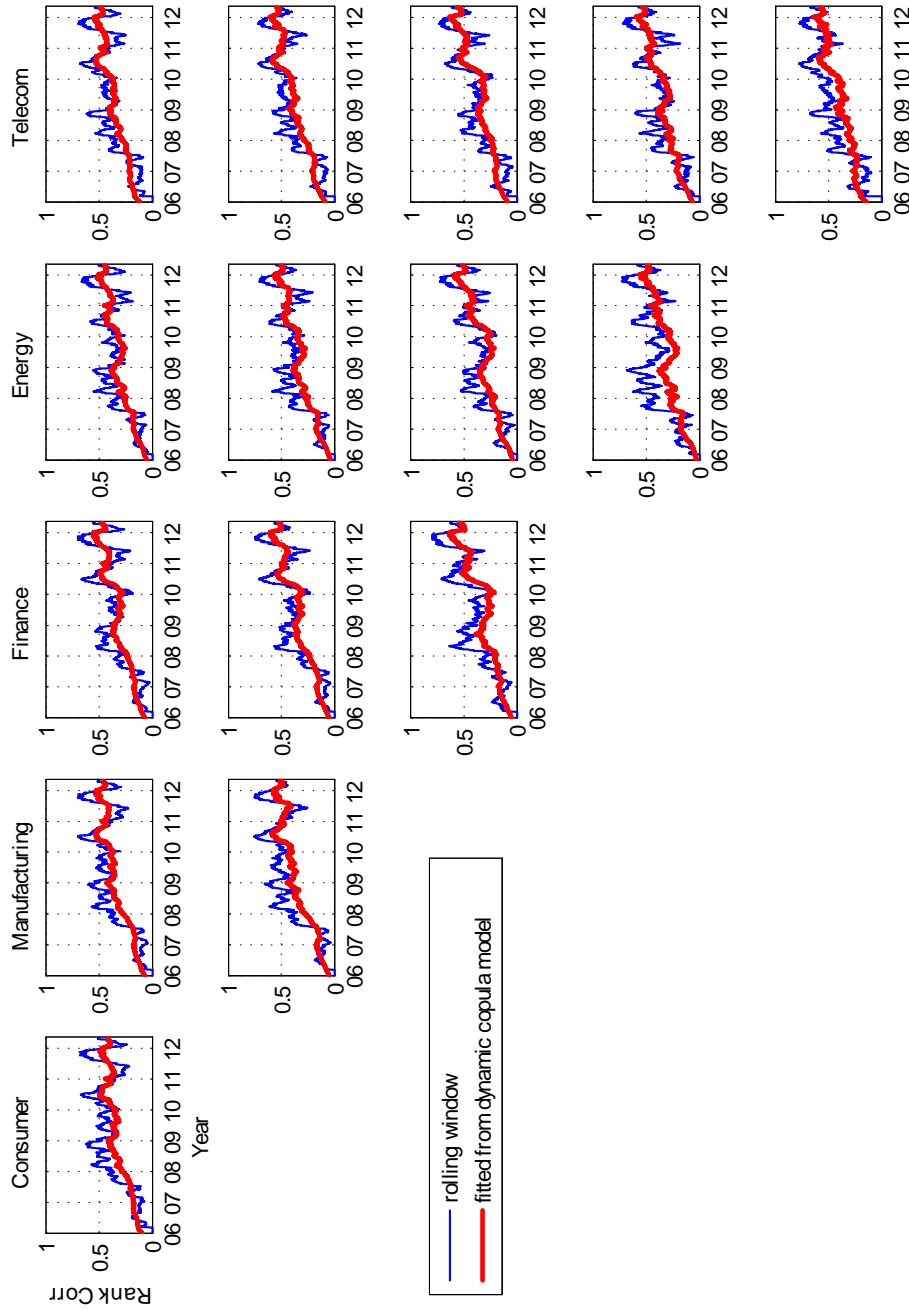


Figure 3: This figure plots the time-varying rank correlations implied by the heterogeneous dependence factor copula, as well as 60-day rolling window rank correlations. Each element of this figure is an average of all pair-wise correlations for firms in the same pair of industries. The diagonal elements present rank correlations between firms in the same industry; the off-diagonal elements for firms in different industries.

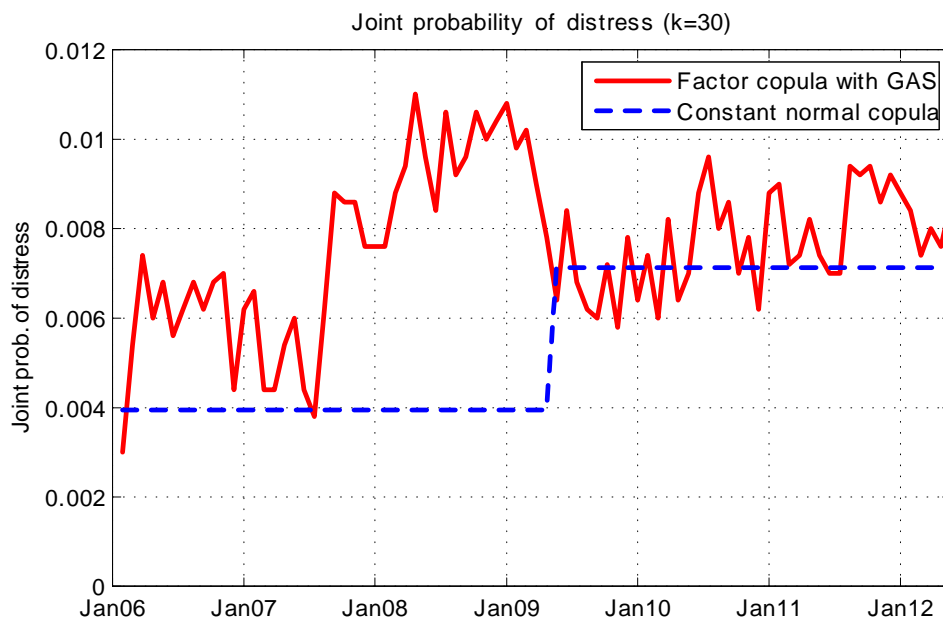


Figure 4: *This figure shows the joint probability of distress (JPD), for two copula models. Both models allow for a structural break on April 8, 2009. Distress is defined using the 99% quantile, and we consider the probability of $k=30$ firms, out of 100, in distress. The sample period is January 2006 to April 2012.*

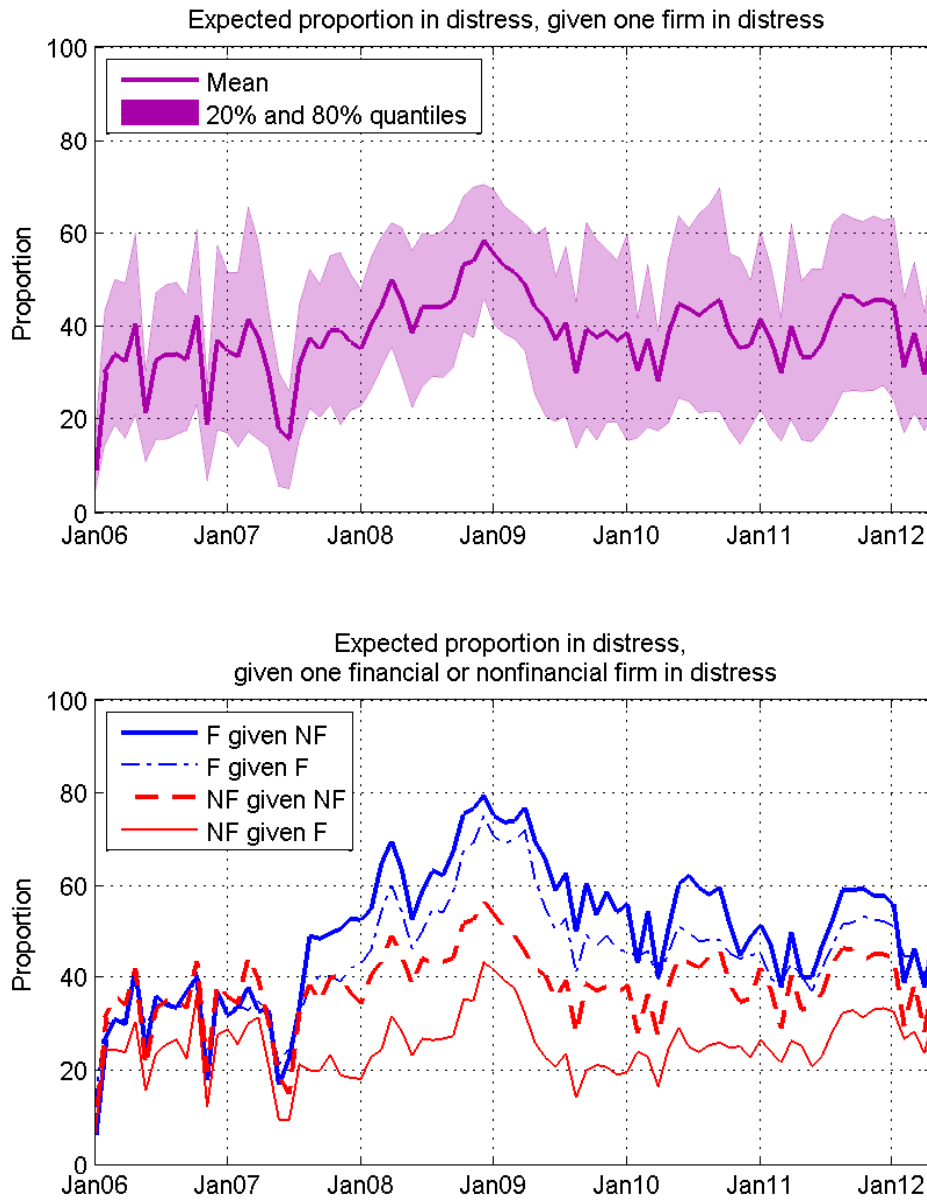


Figure 5: *This figure shows the expected proportion (in percent) of firms in distress, given firm i in distress, averaged across firms. The upper panel covers all 100 firms, and reports the mean and cross-sectional 20% and 80% quantiles. The lower panel reports the proportion of 16 financial firms in distress given one financial or nonfinancial firm in distress, and the proportion of 84 nonfinancial firms in distress given one financial or nonfinancial firm in distress. The sample period is January 2006 to April 2012.*