Regime switches in the volatility and correlation of financial institutions

by K. Boudt, J. Danielsson, S.J. Koopman and A. Lucas

October 2012  No 227
Editorial

On October 11-12, 2012 the National Bank of Belgium hosted a Conference on “Endogenous Financial Risk”.

Papers presented at this conference are made available to a broader audience in the NBB Working Paper Series (www.nbb.be).

Statement of purpose:

The purpose of these working papers is to promote the circulation of research results (Research Series) and analytical studies (Documents Series) made within the National Bank of Belgium or presented by external economists in seminars, conferences and conventions organised by the Bank. The aim is therefore to provide a platform for discussion. The opinions expressed are strictly those of the authors and do not necessarily reflect the views of the National Bank of Belgium.

Orders

For orders and information on subscriptions and reductions: National Bank of Belgium, Documentation - Publications service, boulevard de Berlaimont 14, 1000 Brussels.

Tel +32 2 221 20 33 - Fax +32 2 21 30 42


© National Bank of Belgium, Brussels

All rights reserved.
Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

ISSN: 1375-680X (print)
ISSN: 1784-2476 (online)
Regime Switches in Volatility and Correlation of Financial Institutions

Kris Boudt*  
KU Leuven/Lessius  
V.U. University Amsterdam

Jón Danielsson  
London School of Economics.

Siem Jan Koopman  
V.U. University Amsterdam  
Tinbergen Institute

Andre Lucas  
V.U. University Amsterdam  
Tinbergen Institute

September 3, 2012

Abstract

We propose a parsimonious regime switching model to characterize the dynamics in the volatilities and correlations of US deposit banks’ stock returns over 1994–2011. A first innovative feature of the model is that the within–regime dynamics in the volatilities and correlation depend on the shape of the Student $t$ innovations. Secondly, the across–regime dynamics in the transition probabilities of both volatilities and correlations are driven by macro-financial indicators such as the Saint Louis Financial Stability index, VIX or TED spread. We find strong evidence of time–variation in the regime switching probabilities and the within–regime volatility of most banks. The within–regime dynamics of the equicorrelation seem to be constant over the period.

*This paper was prepared for the 2012 National Bank of Belgium biannual colloquium on Endogenous Financial Risk. Financial support from the National Bank of Belgium and Dutch Science Foundation is gratefully acknowledged. We thank David Ardia, Charles Bos, Hans Dewachter, Sébastien Laurent, Joshua Ulrich, and Raf Wouters for helpful comments. This is still ongoing research. Comments are welcome. Correspondence to: kris.boudt@econ.kuleuven.be.
1 Introduction

The prediction of volatility and correlation of financial assets is key in the study of financial stability as well as for internal risk management of financial institutions. In spite of the ample theoretical and empirical evidence in support of regime switching in both the macro economy and financial markets, standard, non-switching, volatility models, like the GARCH family, remain dominant in practical use. This means that such models are likely to fail when they are perhaps most needed, at the time of a transition between between a low risk and high risk regime. This became especially clear in the transition from the low risk environment before the financial crisis started in the summer of 2007, to the extreme volatility a year and a half later. Addressing this problem is the main motivation of our paper.

The regime switching model we propose has three innovative features. First, using the generalized autoregressive score framework of Haas et al. (2004), we let the shape of the estimated density function directly affect the variance and correlation forecasting functions. In particular, the model accounts for the fact that the heavier the tails of the distribution, the more likely it is that an extreme observation is due to a realization from the (fat) tail rather than a recent volatility increase. As such, the model limits the potential adverse effect of extreme observations on estimation and forecasting of volatility and correlation by downweighting the impact of the most extreme observations. In the special case of normal innovations, the proposed regime switching volatility model coincides with the regime switching GARCH(1,1) model of Haas et al. (2004).

Our second contribution is to assume an equicorrelation structure when modeling the regime switching dependence. This restriction reduces the generality of the universe to which the model can be applied, and in particular, often implies that stocks belong to the same sector (Engle and Kelly, 2012). The loss of flexibility
is compensated by several important advantages. First, the estimated correlations are robust to changes in the universe constituents, which is a crucial requirement for our application to the US deposit bank sector. Second, by averaging across similar pairs of stocks, the impact of single extreme observations on the dynamics is attenuated. Finally, it is computationally more tractable in the presence of many assets. Evaluation of the likelihood function of the proposed model, for example, does not require any matrix inversion.

Our final contribution is to make use of state variables in the modeling and forecasting of switching probabilities. There are several potentially useful candidates for such variables, and we examine the Saint Louis Financial Stability index, the VIX or TED spread. While state variables have generally not been found useful in volatility forecasting, our results indicate that they do make a significant contribution to predicting regime switching probabilities.

Our application is to weekly returns on the largest US headquartered bank holding companies, over January 1, 1994 - June 30, 2011. We find strong evidence of time-variation in the regime switching probabilities and the within-regime volatility of most banks, while the within-regime dynamics of the equicorrelation seem to be constant over the period.

The remainder of the paper is organized as follows. First we give in Section 2 a short review of the literature on changes in the risk regime of financial institutions. Section 3 presents the proposed regime switching volatility-correlation model for analyzing the risk of financial institutions. Section 4 introduces the data. The results of the empirical analysis are presented in Section 5. Section 6 concludes the paper and outlines some directions for further research.
2 Review on changes in the risk regime of financial institutions

Danielsson and Shin (2003) emphasize the distinction between episodes of exogenous and endogenous risk. At times of exogenous risk, price changes can be considered to be solely due to reasons outside the control of market participants. Endogenous risk arises when the behavior of market players creates additional risk with respect to the uncertainty of fundamental news. There are several mechanisms through which endogenous risk can arise. A first one is that during periods of crises, information and beliefs become much more uniform and people behave in a similar way, amplifying the uncertainty in the market. Due to an exogenous adverse shock to capital, for example, the perceived volatility increases and market participants decide to sell their assets, which in turn leads to a drop in prices and a further increase in perceived and realized volatility. When the financial sector is undercapitalized, risk budget constraints enforce financial institutions to shed risky assets and sell at this lower price, which in turn prompts a further fall in prices causing a continuation of the increased volatility and potential further sales, creating a feedback loop of volatility increases and fire sales of risky assets. Another example of such feedback effects is the link between an asset’s market liquidity and traders’ funding liquidity. This liquidity spiral arises under the model of Brunnermeier and Pedersen (2009) when financiers increase margins in response to deteriorations in market liquidity, causing traders to de-lever and decreasing further market liquidity.

These coordinated responses of market participants could be an explanation for the presence of extreme dependence in financial markets. This means that during times of turmoil, the dependence in the movement of asset prices tends to become stronger. Increased dependence makes investing in financial markets riskier, lowers
the benefits of portfolio diversification and increases the proportion of systematic market risk. Campbell et al. (2008) call this phenomenon “Diversification Melt-down”.

Also changes in the fundamentals of the macroeconomy can cause regime switches in financial risk. One of the first to document the relation between U.S. recessions and an increase in aggregate market volatility is Officer (1973) who attributed the very high stock market volatility of the 1930s to the economic recession of that period. Several studies have confirmed this close relation between financial volatility and the economic cycle. Hamilton and Lin (1996), for example, estimated a bivariate regime switching model with high and low real growth states and equity return volatility states and found a strong positive correlation between the estimated probabilities to be in the low real growth state and high equity return volatility state. The implicit view in Hamilton and Lin (1996) is that stock volatility must ultimately be driven by fluctuations in macroeconomic fundamentals, such as GDP, interest rates, oil prices and money supply.

Changes in the risk regimes could explain the formation of asset price cycles. In fact, the current practice in risk models is to use a rolling window volatility estimate or a GARCH model, extrapolating the recent history to the future. This might lead then to asset price bubbles in stable periods, since underestimation of risk leads to an overvaluation of financial assets (Danielsson, 2011). Often the investment and risk management decision are delegated to a financial intermediary such that the ultimate lender is unable to observe how risky the investment of the borrower is of potential agency problems at the level. Asset price bubble might arise when the intermediation is plagued by agency problems due to e.g. the use of debt contracts and the limited liability these involve (Allen and Gale, 2000), or when the investment managers bear limited downside risk because the worst that can happen to them is
that they are fired (Allen and Gorton, 1993).

In a financial system where balance sheets are marked to market, these asset price cycles affect directly the balance sheets of the financial institutions and as noted by Adrian and Shin (2010) banks react procyclically to this. If banks were passive, an increase in the value of their assets, leads to a drop in leverage. Adrian and Shin (2010) observed however that historically banks seem to have a constant leverage target: during the boom, they take on more short term debt and expand their balance sheet. And vice versa during the downturn. In such a manner, asset price cycles create cycles in the balance sheet decisions of financial institutions.

3 A score based regime switching volatility-correlation model

We are interested in the evolution of the volatility and correlation of the weekly returns of the \( N_t \) largest US deposit banks over the period January 1 1994 - June 30 2011. For simplicity in notation, we take \( N \) as constant, but in the empirical application, \( N \) varies between 13 and 15 in function of data availability. Let \( y_t = (y_{1t}, \ldots, y_{Nt})' \) denote the \( N \times 1 \) vector of time \( t \) stock returns with conditional density function \( f_{t|t-1}(y_t; \theta) \). The model will be such that the parameter vector \( \theta \) decomposes into \( N \) parameters \( \theta_i \) associated with margin \( i \) and a parameter \( \theta_* \) for the copula density function. Because of the copula assumption, the conditional density function can be rewritten as the product of the marginal densities \( f_{i|t-1} \) and the copula density denoted \( c_{i|t-1} \):

\[
f_{i|t-1}(y_t; \theta) = \prod_{i=1}^{N} f_{i|t-1}(y_{it}; \theta_i) \cdot c_{i|t-1}(F_{1|t-1}(y_{1t}), \ldots, F_{N|t-1}(y_{Nt}); \theta_*) .
\]

(3.1)
The marginal density functions will be assumed to be standardized student $t$ distributed, and the copula is the standardized multivariate student $t$ distribution with uniform $(0, 1)$ marginals.

In the following two subsections we first present the model for the within–regime dynamics in the volatility and correlation. A large number of volatility and correlation specifications have been proposed in the past, both in the single and multiple regime case. The approach taken here follows the Generalized Autoregressive Score (GAS) framework of Creal et al. (2012) in which the score of the density function is taken as the driver for the time-varying volatility and correlation parameters. It follows that differences in the density function can lead to different volatility and correlation models. The proposed model approach deviates slightly from the general framework of Creal et al. (2012) by using, instead of the full conditional density function, the regime-specific conditional marginal (resp. copula) density function to drive the time-variation in the volatility (resp. correlation parameter) of that regime. This approach has the particular advantage that the regime-specific volatility and correlation process does not depend on the regime paths. It further allows to have different volatility model specifications for each of the bank return series and, combined with the equicorrelation assumption, it makes the estimation feasible in high dimensions. In the Appendix a short introduction to the general GAS framework is given and details on how the application of this framework leads to the models presented below. While the within-regime volatility and correlation dynamics are driven by the past return series, state variables are used to drive the time-variation in the transition probabilities across regimes. Finally, we present a two-step maximum likelihood method for estimating the proposed model.
3.1 Model for the marginal distribution

There is ample empirical evidence for the presence of low and high volatility regimes for many financial assets.\(^1\) Previous research has either assumed that within the regime, volatility is either constant or time-varying according to the regime switching extension of the Bollerslev (1986) GARCH(1,1) model (Gray, 1996; Klaassen, 2002; Haas et al., 2004). More precisely, the regime-switching GARCH(1,1) of Haas et al. (2004) specifies that the conditional variance of asset \(i\) in regime \(k\) is a weighted sum of the realized variance on day \(t - 1\) and the predicted regime \(k\) variance for day \(t - 1\):

\[
h_{it}^k = \omega_{it}^k + \alpha_{it}^k(y_{it-1} - \mu_{it}^k)^2 + \beta_{it}^k h_{it-1}^k, \quad (3.2)
\]

with \(\mu_{it}^k\) the conditional mean of asset \(i\) in regime \(k\) (for simplicity, we assume no within-regime dynamics in the mean) and \(k \in \{I, II\}\).

Several authors (Dueker, 1997; Klaassen, 2002) have then combined these regime switching GARCH(1,1) type of models with Student \(t\) innovations, since they significantly improve the model likelihood and lead to more stable regimes. The standardized student \(t\) density function with mean \(\mu_{it}^k\), conditional variance \(h_{it}^k\) and number of degrees of freedom \(\nu_{it}^k\) is

\[
t_{\nu_{it}^k}(y_{it}, \mu_{it}^k, h_{it}^k) = \frac{1}{\sqrt{h_{it}^k}} \frac{\Gamma \left( \frac{\nu_{it}^k + 1}{2} \right)}{\Gamma \left( \frac{\nu_{it}^k}{2} \right) [\pi (\nu_{it}^k - 2)]^{1/2}} \left[ 1 + \frac{(y_{it} - \mu_{it}^k)^2}{h_{it}^k (\nu_{it}^k - 2)} \right]^{-\frac{1 + \nu_{it}^k}{2}}, \quad (3.3)
\]

where \(\Gamma(\cdot)\) is the gamma function.

In a single regime setup, however, Creal et al. (2012) and Harvey and Chakravarty

\(^1\)Throughout the paper, we will assume that for each series there are either one or two volatility regimes. In the application, the number of regimes is determined using the BIC criterion. The assumption of at most two regimes keeps notation tractable and is in line with previous empirical research showing evidence of a high and low volatility regime. Regimes are indicated with Roman numbers (ie regime I and II).
(2008) have argued that one should design conditional volatility models differently for fat-tailed distribution than for normal distribution. The intuition is clear: the more heavy tailed the distribution, the less likely it becomes that an extreme observation is due to an increase in volatility.

This is appropriately accounted for by the Student t score based volatility model, where the conditional volatility is no longer a linear combination of squared observations, but it downweights observations in function of their extremeness and the tail thickness:

$$h_{it}^k = \omega^k + \alpha^k (1 + 3/\nu^k) \frac{\mu^k + 1}{[\nu^k - 2] + \left[ \frac{(y_{it-1} - \mu^k)^2}{h_{it-1}^k} \right]^{\nu^k - 2} + \left[ \frac{y_{it} - \mu^k}{h_{it-1}^k} \right]^2} + \beta^k h_{it-1}^k. \quad (3.4)$$

Note that for \( \nu^k = \infty \), we retrieve the usual GARCH(1,1) model. In the estimation we impose the usual inequalities \( \omega^k, \alpha^k, \beta^k > 0 \) such that the conditional variance process is guaranteed to be positive. We further assume that \( \nu^k > 2 \) such that in each regime there is a finite variance.

The difference in volatility impact in function of the tail fatness is illustrated in Figure 1. The full line indicates the case of normal innovations, where the model coincides with a GARCH model and the impact of lagged returns on future volatility is quadratic. The dashed and dotted lines correspond to Student t innovations with 10 and 4 degrees of freedom, respectively. We see that the volatility impact is still increasing as a function of the magnitude of the return, but less than quadratic and, the thicker the tails are, the more extreme observations are downweighted compared to the quadratic impact under the normal model.

\(^2\)An alternative approach for accounting for the proportionally smaller impact of extreme observations is the class of bounded innovation propagation volatility models, proposed by Muler and Yohai (2008), Boudt and Croux (2010) and Boudt et al. (2012). This approach only specifies the central model generating the majority of the return observations and estimates this model using a robust estimation techniques to ensure that the parameter estimates are not much affected by a few extreme price movements that are unlikely to occur under the central model.
3.2 Model for the conditional copula density function

All assets considered here are major US deposit banks. For such a homogeneous group of assets, it is reasonable to assume that the correlation parameter in the copula density function is the same for all pairs of stocks. This so-called “dynamic equicorrelation parameter” has been extensively studied by Engle and Kelly (2012), under the assumption that there are no shifts in the unconditional equicorrelation. In our application on financial institutions, this assumption is likely to be violated. Because of the interconnectedness between financial institutions, episodes of high systemic risk are characterized by dramatic changes in the volatility and correlations of financial institutions. See e.g. the recent empirical evidence in Leonidas and Italo de Paula (2011) and the review in Section 2. Ang and Chen (2002) find that
regime-switching models perform best in explaining the difference in correlation in normal versus extreme downside regimes.

We allow for two correlation regimes and the dependence in each regime is modeled through a Student $t$ copula. More precisely, we assume that in the correlation regime $s_{st} = k$, the copula is a Student $t$ copula with $\nu^k_s$ degrees of freedom and correlation matrix $R^k_{st}$.\footnote{This correlation matrix is not the conditional correlation of the stock returns, as the marginal distributions and the copula have different degrees of freedom parameters.} Denote $u_{it} \in (0, 1)$ the probability integral transform of $y_{it}$ and $\tilde{y}^k_{it}$ the $u_{it}$ quantile of the univariate standardized Student $t$ distribution with $\nu^k_s$ degrees of freedom. The corresponding Student $t$ copula density in regime $s_{st} = k$ is:

$$c_{t|t-1}(u_t, \nu^k_s, R^k_t) = (\text{det} R^k_t)^{-1/2} t_N\left[\left(\tilde{y}^k_t\right)'(R^k_t)^{-1}\left(\tilde{y}^k_t\right), \nu^k_s\right] \prod_{i=1}^N t_1\left[(\tilde{y}^k_{it})^2, \nu^k_s\right],$$

with $t_N(z, \nu)$ the standardized multivariate student $t$ density function:

$$t_N(z, \nu) = \frac{\Gamma \left(\frac{\nu+N}{2}\right)}{\Gamma \left(\frac{\nu}{2}\right) \pi^{N/2}} \left[1 + \frac{z}{\nu - 2}\right]^{-\frac{N+\nu}{2}},$$

where $\Gamma(\cdot)$ is the gamma function.

Under the equicorrelation assumption, the correlation matrix in the correlation regime $k$ can be rewritten as:

$$R^k_t = [1 - \rho^k_t]I_N + \rho^k_t J_N,$$

where $I_N$ is the $N$-dimensional identity matrix and $J_N$ the $N \times N$ matrix of ones. Besides computational convenience\footnote{We have the following computationally convenient formulas for the determinant and Maha-}, an important advantage of this restriction is that the time-variation in $R^k_t$ can still be estimated, even if the universe changes on...
a regular basis.

To make sure that the estimated correlations are bounded, we specify $\rho^k_t$ as the hyperbolic tangent of an underlying process $q^k_t$:

$$
\rho^k_t = \frac{\exp(2q^k_t) - 1}{\exp(2q^k_t) + 1}.
$$

(3.8)

The $q^k_t$ process is further truncated to make sure that $R^k_t$ is always positive definite.\(^5\)

The specification for the dynamics of $q^k_t$ is obtained under the $t$ GAS framework in Creal et al. (2011). Under this framework, $q^k_t$ is modeled as a linear function of the score of the Student $t$ copula density function:

$$
q^k_t = \omega^k_t + \alpha^k_t (q^k_{t-1} + S^k_{t-1} \nabla^k_{t-1}) + \beta^k_t q^k_{t-1},
$$

(3.9)

where $\nabla^k_t = \partial \log c_{t|t-1}(u_t, \nu^k_t, R^k_t)/\partial q^k_t$ is the score of the Student $t$ copula density function in correlation regime $k$ and $S^k_t$ is the inverse of the conditional standard deviation of the score, and $\alpha^k_*, \beta^k_* > 0$.

In Appendix, we show that the standardized score is given by:

$$
S^k_t \nabla^k_t = m^k_t \left[ b^k_t \left( \frac{w^k_t}{(N - 1)N} \sum_{i=1}^{N} \sum_{j \neq i} \tilde{y}_{it} \tilde{y}_{jt} - \rho^k_t \right) + a^k_t \left( \frac{w^k_t}{N} \sum_{i=1}^{N} \tilde{y}_{it}^2 - 1 \right) \right],
$$

(3.10)

lanobis distance:

$$
\begin{align*}
\det R^k_t &= \left[ 1 - \rho^k_t \right]^{N-1} \left[ 1 + (N - 1)\rho^k_t \right] \\
(\tilde{y}_t^k)'(R^k_t)^{-1}(\tilde{y}_t^k) &= \frac{1}{1 - \rho^k_t} \left[ \sum_i (\tilde{y}_{it}^k)^2 - \frac{\rho^k_t}{1 + (N - 1)\rho^k_t} \left( \sum_i (\tilde{y}_{it}^k)^2 \right) \right].
\end{align*}
$$

The use of these formulae lends itself to an efficient evaluation of the log-density of the sample, whereby no matrix inversion is required.

\(^5\)A necessary and sufficient condition for positive definiteness of $R^k_t$ is that $-1/(N - 1) < \rho^k_t < 1$. In practice, we imposed this condition in the optimization by a lower and upper truncation of $q^k_t$ at $0.5(\log(N + 0.01) - \log(N + 2.01))$ and 2.75.
with $a_t^k = -\rho_t^k (2 + \rho_t^k (N - 2))$, $b_t^k = 1 + (\rho_t^k)^2 (N - 1)$, and importantly, the weighting factor

$$w_t^k = \frac{[N + \nu_*^k]}{[\nu_*^k - 2 + (\tilde{y}_t^k)' (R_t^k)^{-1} (\tilde{y}_t^k)]}.$$ 

The scaling factor $n_t^k > 0$ is defined in the Appendix.

The score has three main components: (i) the excess value of the cross-product of weighted devolatilized returns and the conditional correlation, (ii) the excess value of the Euclidean norm of those returns and unity and (iii) the weights applied to the devolatilized returns. The first component is probably the most intuitive, as it enforces an increase in the conditional correlation process when the average cross-product of the devolatilized returns exceeds the conditional correlation $\rho_t^k$. This is as expected, and very much in line with the dynamic conditional correlation specifications in the DCC model of Engle (2002) and DECO model of Engle and Kelly (2012).

A drawback of using the average cross-product of devolatilized returns as a proxy for correlations is that it does not correct for the dispersion. In fact, the higher the dispersion, the less informative high values of the cross-products are about increases in correlation. It is intuitively clear that in a bivariate setting, the correlation signal of $(1, 1)$ is much stronger than $(1/4, 4)$, even though their cross-product is the same. The second term in the score corrects for this. For $\rho_t^k > 0$, it reduces the score when the Euclidean norm of the weighted devolatilized returns exceeds its expected value of unity.

A final important feature in the score is the weighting of the returns in function of their extremeness as measured by the Mahalanobis distance, with weights that depend on the dimension of the series $N$, the value of the correlation coefficient $\rho_t^k$ and the tail thickness parameter $\nu_*^k$. The thicker the tails, the more likely it is
that abnormally large values of the realized covariance compared to the predicted correlation are due to the heavy-tailed feature of the distribution rather than changes in correlation. For this reason, we see that the effect of the downweighting increases when the tail parameter is small and the observation is more extreme. For $\rho^k_t = 0$ and $\rho^k_t = 0.5$, we illustrate this in Figure 2, where we plot the values of the standardized score $S^k_{t-1} \nabla^k_{t-1}$ for a bivariate vector $u$ with component values between -2 and 2. The score is shown for the bivariate Gaussian copula (left panel) and the Student $t$ copula with 4 degrees of freedom (right panel). From the graph, the importance of the impact of the choice of density on the score, and hence the correlation dynamics, is clear. The more thick-tailed the copula is, the more extreme observations are downweighted. Because the weights depends on the Mahalanobis distance, the value of $\rho^k_t$ impacts the curvature and which values are considered as extreme.
Figure 2: Standardized score under the equicorrelation normal (left panel) and Student $t_4$ (right panel) copula model for $N = 2$. 

Normal copula with $\rho = 0$

Student $t_4$ copula with $\rho = 0$

Normal copula with $\rho = 0.5$

Student $t_4$ copula with $\rho = 0.5$
3.3 Dynamics across volatility and correlation regimes

It can be expected that the likelihood of staying in a regime depends on the changing market conditions, and that the VIX and TED spread are useful state variables determining the intensity of regime changes.

Besides the VIX and TED spread, there are many potential candidates for economic variables driving the time variation in the transition probabilities of the volatility and correlation regimes. An aggregate index for financial stability was developed by the Federal Reserve of Saint Louis. This so-called Saint Louis Financial Stability Index (STLFSI) is defined as the first principle component of eighteen major financial time series capturing some aspect of financial stress.\(^6\)

Henceforth, let \(x\) be the series driving the time-variation in the transition probabilities. Denote \(\psi_{it} = (\psi^I_{it}\psi^{II}_{it})'\) the 0/1 probability value to be in the low and high volatility regime for asset \(i\). We assume the states follow a Markov process with the \(2 \times 2\) dynamic transition matrix \(P_{it}\). The diagonal elements of this matrix are parameterized using the logit transformation of the time-varying quantities \(\pi^I_{it}\) and \(\pi^{II}_{it}\):

\[
P_{(11)it} = \frac{\exp(\pi^I_{it})}{[1 + \exp(\pi^I_{it})]}
\]

\[
P_{(22)it} = \frac{\exp(\pi^{II}_{it})}{[1 + \exp(\pi^{II}_{it})]}
\]

\(^6\)The complete list of variables is composed of 7 interest rates (Effective federal funds rate, 2-year Treasury, 10-year Treasury, 30-year Treasury, Baa-rated corporate, Merrill Lynch High-Yield Corporate Master II Index, Merrill Lynch Asset-Backed Master BBB-rated), 6 yield spreads (10-year Treasury minus 3-month Treasury, Corporate Baa-rated bond minus 10-year Treasury, Merrill Lynch High-Yield Corporate Master II Index minus 10-year Treasury, 3-month London Interbank Offering RateOvernight Index Swap (LIBOR-OIS) spread, 3-month Treasury-Eurodollar (TED) spread, 3-month commercial paper minus 3-month Treasury bill) and 5 other indicators (J.P. Morgan Emerging Markets Bond Index Plus, Chicago Board Options Exchange Market Volatility Index (VIX), Merrill Lynch Bond Market Volatility Index (1-month), 10-year nominal Treasury yield minus 10-year Treasury Inflation Protected Security yield (breakeven inflation rate), Vanguard Financials Exchange-Traded Fund (equities)). For further information, see http://research.stlouisfed.org/publications/net/NETJan2010Appendix.pdf.
We follow the standard approach to generate dynamics in the transition probabilities by using exogenous variables such as the VIX or TED spread as the source of time variation (Connolly et al., 2005; Diebold et al., 1994; Schaller and Van Norden, 1997) and to assume a linear specification:

\[
\pi_{I,t} = c_{I}^{I} + d_{I}^{I} x_{t-1},
\]

\[
\pi_{II,t} = c_{II}^{I} + d_{II}^{I} x_{t-1},
\]

with \( x_{t-1} \) the time \( t - 1 \) value of the exogenous variable. Readers interested in endogenous switching whereby the transition probability depends on \( x_{t} \) rather than \( x_{t-1} \) are referred to the works of Filardo (1994) and Kim et al. (2008), requiring an additional model assumption of the dependence between the shocks in \( x_{t} \) and \( y_{it} \).

Like for the volatility, we consider a low and high correlation regime, with \( \psi_{st} = (1 0) \)' in the low correlation regime and \( \psi_{st} = (0 1) \)' in the high correlation regime. \( P_{st} \) is a 2 \( \times \) 2 transition matrix for the correlation regimes and \( \psi_{st|t-1} = P_{st}\psi_{t-1} \). We parameterize the transition matrix using the logit transformation of the time-varying quantities \( \pi_{st}^{I} \) and \( \pi_{st}^{II} \), with \( \pi_{st}^{I} = c_{s}^{I} + d_{s}^{I} x_{t-1}, \pi_{st}^{II} = c_{s}^{II} + d_{s}^{II} x_{t-1} \).

### 3.4 Estimation

The parameters are estimated by maximizing the log-likelihood function:

\[
LLH(\theta) = \sum_{t=1}^{T} \log c_{t|t-1}(F_{1t|t-1}(y_{1t}), \ldots, F_{Nt|t-1}(y_{Nt}); \theta_{s}) + \sum_{i=1}^{N} \sum_{t=1}^{T} f_{it|t-1}(y_{it}; \theta_{i}).
\]

Since the joint estimation of all parameters is computationally too demanding, we follow Patton (2006), Jondeau and Rockinger (2006) and Pelletier (2006) in estimating the parameter vector \( \theta = (\theta_{1}, \ldots, \theta_{N}, \theta_{s})' \) by a two-step maximum likelihood
procedure, whereby in the first step the parameters pertaining to the marginal distributions are estimated, and in the second step the dependence structure is estimated given the estimated margins.

For the estimation of the margins, the function to be maximized is the log of the expected likelihood of each series:

$$LLH_i(\theta_i) = \sum_{t=1}^{T} \log \psi'_{it|t-1}(\theta_i)\eta_{it}(\theta_i),$$

with $\eta_{it}(\theta_i)$ the 2-dimensional vector holding the conditional densities under each regime. The conditional probabilities to be in each state $\psi_{it|t-1}(\theta_i)$ are obtained using the usual Hamilton filter:

$$\bar{\psi}_{it|t} = \frac{\psi_{it|t-1} \odot \eta_{it}}{\psi'_{it|t-1} \odot \eta_{it}}, \quad (3.11)$$

$$\psi_{it+1|t} = P'_{it+1} \bar{\psi}_{it|t}, \quad (3.12)$$

where $\odot$ denotes element-wise multiplication and $\iota$ is a vector of ones.

Maximization of $LLH_i(\theta_i)$ yields $\hat{\theta}_i$. Given $\hat{\theta}_i$, the probability integral transform of $y_{it}$ can then be computed. Denote these for the $N$ series as $\hat{u}_t = (\hat{u}_{1t}, \ldots, \hat{u}_{Nt})'$. For each possible value of $\theta_*$, the copula log likelihood function can now be evaluated, using the Hamilton filter to do the inference on the regime probabilities. Maximization of the expected copula log likelihood gives the estimate for the dependence parameter $\theta_*$.

To safeguard the analysis against being trapped in local optima, good starting values were first obtained using a global genetic-type of optimizer called Differential Evolution developed by Price et al. (2006). It combines in each generation three random members to improve the current solution. The first generation is calibrated to a combination of optimized solutions to restricted versions of the model, first
step estimates and some random parameter solutions. Gradient techniques are then used to further improve locally the estimates. Bounds on parameters are indirectly imposed through variable transformations. The implementation and analysis was done in the R environment, using the DEoptim (Ardia et al., 2012), Performance-Analytics (Carl et al., 2012), Rcpp (Eddelbuettel and François, 2011) and xts (Ryan and Ulrich, 2012) packages.

For the estimation of the copula likelihood, we further imposed that the degrees of freedom parameter is the same in the two correlation regimes. Additionally, the optimization method discussed above was implemented inside a loop on the degrees of freedom parameter.
Data

Our application is to weekly returns on the largest US headquartered bank holding companies, over January 1 1994 - June 30 2011. For each year, the price data was collected for the fifteen largest bank holding companies, headquartered in the US, in terms of US domestic deposits. The starting year coincides with the first year for which the Federal Deposit Insurance Corporation publishes the yearly report on US domestic deposits as of June 30 of each year on its website.\footnote{http://www2.fdic.gov/sod/index.asp} Table 1 lists the selected bank holding companies and the years for which they belonged to the top 15 of largest deposit holding companies in the US, for at least three consecutive years. Because of missing data, the eventual number of banks per weekly time unit is between 11 and 14. Returns are computed as the weekly log difference of the adjusted price series, downloaded from CRSP, and expressed in percentage points.

We first conduct an exploratory analysis to visualize the time-variation in universe, the volatility and the correlation.

The absolute value of the weekly return series of different length is plotted in the left panel of Figure 7 in the Appendix, together with the sample standard deviation and the single regime volatility estimates under the t-GARCH model of Bollerslev (1987) and the t-GAS model of Creal et al. (2012). Comparing the t-GARCH and t-GAS volatility forecasts, we see that for almost all series, accounting for the thickness of the tails, leads to a lower forecast of the conditional volatility and a reduced sensitivity to extreme returns. There is a large commonality in the volatility dynamics of the different series. Figure 3 reports the time series plot of the weekly cross-sectional mean absolute returns. In that Figure, shaded areas were added to indicate the periods identified by the NBER as economic contractions in the US, namely March 2001-November 2001 and December 2007-June 2009. Additionally,
Table 1: List of US bank holding companies, together with the first and last year for which they belong to the top 15 of largest deposit banks in the US.

<table>
<thead>
<tr>
<th>PERMNO</th>
<th>Name</th>
<th>First</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>34746</td>
<td>Fifth Third Bancorp</td>
<td>2001</td>
<td>2011</td>
</tr>
<tr>
<td>35044</td>
<td>Regions Financial Corp</td>
<td>2005</td>
<td>2011</td>
</tr>
<tr>
<td>36469</td>
<td>First Union Corp, Wachovia Corp</td>
<td>1994</td>
<td>2008</td>
</tr>
<tr>
<td>38703</td>
<td>Norwest Corp, Wells Fargo &amp; Co</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>47079</td>
<td>Citicorp</td>
<td>1994</td>
<td>1998</td>
</tr>
<tr>
<td>47896</td>
<td>Chemical Banking Corp, Chase Manhattan Corp, JP Morgan Chase &amp; Co</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>49656</td>
<td>Bank of New York Mellon Corp</td>
<td>2008</td>
<td>2011</td>
</tr>
<tr>
<td>50024</td>
<td>Wells Fargo &amp; Co (before merger with Norwest)</td>
<td>1994</td>
<td>1997</td>
</tr>
<tr>
<td>56232</td>
<td>National City Corp</td>
<td>1996</td>
<td>2008</td>
</tr>
<tr>
<td>58827</td>
<td>Bankamerica Corp</td>
<td>1994</td>
<td>1998</td>
</tr>
<tr>
<td>59408</td>
<td>Nationsbank Corp, Bankamerica Corp, Bank of America Corp</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>60442</td>
<td>PNC Bank Corp, PNC Financial Services GRP Inc</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>61284</td>
<td>Barnett Banks Inc</td>
<td>1994</td>
<td>1997</td>
</tr>
<tr>
<td>64995</td>
<td>Keycorp</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>65138</td>
<td>Bank One Corp</td>
<td>1994</td>
<td>2004</td>
</tr>
<tr>
<td>66157</td>
<td>US Bancorp</td>
<td>1998</td>
<td>2011</td>
</tr>
<tr>
<td>68144</td>
<td>Suntrust Banks Inc</td>
<td>1994</td>
<td>2011</td>
</tr>
<tr>
<td>69032</td>
<td>Morgan Stanley</td>
<td>2009</td>
<td>2011</td>
</tr>
<tr>
<td>70519</td>
<td>Citigroup</td>
<td>1999</td>
<td>2011</td>
</tr>
<tr>
<td>71563</td>
<td>Southern National Corp NC, BB&amp;T Corp</td>
<td>2000</td>
<td>2011</td>
</tr>
<tr>
<td>81055</td>
<td>Capital One Financial Corp</td>
<td>2006</td>
<td>2011</td>
</tr>
</tbody>
</table>

PERMNO is the CRSP identifier.

Vertical lines were added to indicate important economic events over the period.\(^8\)

The ratio between the cross-sectional covariance between the returns in the universe and the variance of these returns can be seen as a proxy for the equicorrelation:

\[
 rt = \left[ \frac{1}{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} I_{it} I_{jt}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} y_{it} y_{jt} I_{it} I_{jt} \right] \left[ \frac{1}{\sum_{i=1}^{N} I_{it}} \sum_{i=1}^{N} y_{it}^2 I_{it} \right],
\]

with \( I_{it} \) the dummy variable indicating that bank \( i \) belongs to the top 15 of US deposit bank holding firms in year \( t \). The time series of \( r_t \) is plotted in the top panel of Figure 4. The realized equicorrelation reveals to be very noisy with little

Figure 3: Time series of weekly values of the mean absolute returns across US deposit bank holding companies

Persistence. Taking the rolling four-week weekly average shows the dynamics in the equicorrelation in a more clear way. There seems to be an increase in correlation over the 2008 financial crisis period. The average equicorrelation over the period is 44%. There are also relatively large swings in the equicorrelation around the announcement on October 27, 2003 of the acquisition of Fleetboston by Bank of
America.\textsuperscript{9}

Figure 4: Time series of realized equicorrelations across US deposit bank holding companies, its rolling 4-week average value and the sample average.

Finally, let us inspect the time series plot of the VIX, TED spread and Saint Louis Financial Stability Index (STLFSI) in Figure 5. Comparing the time series of the VIX and TED spread, we see that the TED spread has relatively reacted more to

\textsuperscript{9}See, for example, the CNN journal article at \url{http://money.cnn.com/2003/10/27/news/companies/ba_fleet/index.htm}
the 2007-2008 credit crunch following the subprime crisis than the VIX. Correlations between STLFSI and VIX and TED spread are 83% and 56%, respectively, while correlation between VIX and TED spread is 52%.
5 Results

5.1 Analysis of volatility dynamics

For each financial institution we estimate the GARCH and the t-GAS volatility models under the assumption of 1 regime, 2 regimes with constant volatility in each regime, two regime with one regime with constant volatility and another one with dynamic volatility and 2 regimes with dynamic volatility. Regime switching probabilities are either constant or driven by the STLFSI, TED spread or VIX.

For each bank, we selected the “best” model as the one with the lowest value for the Bayesian information criterion (BIC). In the right panel of Figure 7 in the Appendix, we show the volatility forecast under the model with the lowest BIC for each deposit bank.

A first interesting results is that, of all models considered, the lowest BIC is achieved by a double regime volatility model, with time-varying transition probabilities. The STLFSI, TED spread and VIX are selected 6, 6, and 10 times respectively. For shorter return series, at least one of the regimes tends to be characterized by constant volatility. The t-garch model is selected for 8 banks, the t-gas model for 10 banks.

The average BICs (relatively to the BIC of the Gaussian constant volatility model) are in Table 2.\textsuperscript{10} The first four lines showing the model fit in the single regime case confirm the stylized fact of fat tails and time-varying volatility in the returns of financial institutions.

The next three lines show the average model fits of using a double regime constant volatility model. In case of constant transition probabilities, the two regime models underperform compared to the the single regime time varying volatility models. But

\textsuperscript{10}For the univariate models, all BICs were positive. Standardizing is needed to account for the different number of observations.
Table 2: Average model fit (BIC) of univariate volatility models for US deposit banks. The BIC is expressed relatively to the Gaussian constant volatility model. The conditional mean return in each regime is constant. The conditional variance is either constant or time-varying according to the GAS specification.

<table>
<thead>
<tr>
<th>R1</th>
<th>R2</th>
<th>(v_1 = v_2)</th>
<th>BIC</th>
<th>BIC-STLFSI</th>
<th>BIC-TED</th>
<th>BIC-VIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-constant</td>
<td>g-GAS</td>
<td></td>
<td>0.965</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g-GAS</td>
<td>t-constant</td>
<td></td>
<td>0.933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-gas</td>
<td>g-gas</td>
<td></td>
<td>0.923</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-GARCH</td>
<td>t-constant</td>
<td></td>
<td>0.931</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g-constant</td>
<td>g-constant</td>
<td></td>
<td>0.946</td>
<td>0.914</td>
<td>0.898</td>
<td>0.884</td>
</tr>
<tr>
<td>t-constant</td>
<td>t-constant</td>
<td></td>
<td>0.947</td>
<td>0.893</td>
<td>0.887</td>
<td>0.832</td>
</tr>
<tr>
<td>t-constant</td>
<td>t-constant</td>
<td>YES</td>
<td>0.945</td>
<td>0.88</td>
<td>0.842</td>
<td>0.84</td>
</tr>
<tr>
<td>g-gas</td>
<td>g-constant</td>
<td></td>
<td>0.938</td>
<td>0.93</td>
<td>0.939</td>
<td>0.945</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-constant</td>
<td></td>
<td>0.941</td>
<td>0.859</td>
<td>0.83</td>
<td>0.8</td>
</tr>
<tr>
<td>t-garch</td>
<td>t-constant</td>
<td></td>
<td>0.94</td>
<td>0.855</td>
<td>0.83</td>
<td>0.818</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-constant</td>
<td>YES</td>
<td>0.939</td>
<td>0.857</td>
<td>0.829</td>
<td>0.798</td>
</tr>
<tr>
<td>t-garch</td>
<td>t-constant</td>
<td>YES</td>
<td>0.939</td>
<td>0.857</td>
<td>0.827</td>
<td>0.814</td>
</tr>
<tr>
<td>g-gas</td>
<td>g-gas</td>
<td></td>
<td>0.935</td>
<td>0.885</td>
<td>0.868</td>
<td>0.854</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-gas</td>
<td></td>
<td>0.935</td>
<td>0.852</td>
<td>0.825</td>
<td>0.789</td>
</tr>
<tr>
<td>t-garch</td>
<td>t-garch</td>
<td></td>
<td>0.945</td>
<td>0.855</td>
<td>0.828</td>
<td>0.788</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-gas</td>
<td>YES</td>
<td>0.942</td>
<td>0.862</td>
<td>0.827</td>
<td>0.802</td>
</tr>
<tr>
<td>t-garch</td>
<td>t-garch</td>
<td>YES</td>
<td>0.943</td>
<td>0.86</td>
<td>0.829</td>
<td>0.801</td>
</tr>
</tbody>
</table>

The lowest BIC values per column are bolded. For the multiple regime models, the lowest BIC value per row is underlined.

If we allow for time varying models, the BIC is significantly reduced. The lowest average relative BIC is obtained using the VIX as the variables driving the transition probabilities and Student \(t\) innovations.

The last ten lines correspond to models with time varying volatility in at least one of the regimes. Assuming constant transition probabilities, the single regime Student \(t\) GAS model still has the best model fit. Using the STLFSI or TED spread as a driver for the time-varying transition probabilities, the 2-regime t-GAS model has the lowest BIC. Using the VIX the lowest BIC is obtained using a two regime t-GARCH volatility model, but the difference in BIC with the t-GAS model is negligible.
5.2 Analysis of equicorrelation dynamics

We now turn to the analysis of the dynamics in the equicorrelation parameter of the Student \( t \)-copula. We estimated the \( t \)-GAS equicorrelation model under the assumption of 1 regime, 2 regimes with constant equicorrelation in each regime, two regime with one regime with constant equicorrelation and another one with dynamic volatility and 2 regimes with dynamic equicorrelation. Regime switching probabilities are either constant or driven by the STLFSI, TED spread or VIX. Based on an initial grid search, the degrees of freedom parameter is set to 100 for all models and regimes.

Table 3: BIC associated to the estimated equicorrelation \( t \)-copula for US deposit banks. The equicorrelation parameter is either constant or time-varying according to the GAS specification.

<table>
<thead>
<tr>
<th>R1</th>
<th>R2</th>
<th>STLFSI</th>
<th>TED</th>
<th>VIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-constant</td>
<td>24036.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-gas</td>
<td>24047.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-constant</td>
<td>t-constant</td>
<td>23691.31</td>
<td>23127.65</td>
<td>23101.59</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-constant</td>
<td>23704.71</td>
<td>23144.48</td>
<td>23114.93</td>
</tr>
<tr>
<td>t-gas</td>
<td>t-gas</td>
<td>23718.00</td>
<td>23160.06</td>
<td>23128.42</td>
</tr>
</tbody>
</table>

The lowest BIC values per column are bolded. For the multiple regime models, the lowest BIC value per row is underlined.

Table 3 reports the BIC associated to those equicorrelation \( t \)-copula models. All models without time-varying transition probabilities underperform compared to the constant equicorrelation model. There does not seem to be much to gain from modeling the within-regime dynamics, which is in support of the regime switching constant correlation model of Pelletier (2006), but with time-varying transition probabilities. The best model is a two-regime equicorrelation model with time-variation driven by the VIX and (an almost negligible) time-varying correlation in the low correlation regime. The corresponding predicted weekly equicorrelation parameter
Figure 6: Time series of dynamic equicorrelation parameter under the VIX driven two-regime constant correlation t-copula model (upper plot) and the probability to be in the high correlation regime (lower plot).

is plotted in Figure 6. We see that the estimated equicorrelation fluctuates between 42% and 75% and that the likelihood of being in the high correlation regime tends to be high when the VIX is relatively high.
6 Conclusion

In spite of ample theoretical and empirical evidence in support of regime switching in both the macro economy and financial markets, standard, non-switching, volatility models, like the GARCH family, remain dominant in practical use. This means that such models are likely to fail when they are perhaps most needed, at the time of a transition between a low risk and high risk regime. This motivates us to propose a regime switching volatility-correlation model to predict the volatility and correlation of financial assets. The within-regime dynamics of volatility and correlation are driven by the score of the density function, while state variables such as the VIX and TED spread affect the transition probabilities between volatility and correlation regimes.

The model is applied to a time-varying universe of US deposit bank holding companies, over 1994 - 2011. We find strong evidence of time-variation in the regime switching probabilities and the within-regime volatility of most banks, while the within-regime dynamics of the equicorrelation seem to be constant over the period.

Several topics stand out as interesting avenues for further research. A first one relates to the use of US deposit bank balance sheet variables as a driver for the transition probabilities, and in particular the banks’ leverage. A model replicating the joint dependence between balance sheet capacity, volatility and risk premium was recently proposed by Danielsson et al. (2011). Leverage is also a key variable in the systemic risk calculation method proposed by Brownlees and Engle (2011). Secondly, the model could be generalized, by allowing for endogenous regime switching as in Filardo (1994) and Kim et al. (2008), or considering other distributions. Through a skewed distribution, leverage effects could be introduced in the within regime
volatility dynamics, as in Ardia (2009). Finally, the proposed model does not allow for spillover effects between the volatility dynamics of the two regimes. This has the computational advantage that the conditional volatility predictions in each regime do not depend on the path of the volatility of the other regime. But other choices are possible, as in Gray (1996) or Klaassen (2002).
Appendix on GAS models

1. The Generalized Autoregressive Score framework in Creal et al. (2012)

Recall that we denote \( f_{t|t-1}(y_t; \lambda_t, \theta) \) the conditional observation density and suppose there is no regime switching such that the conditional variances and the copula correlation coefficient are the only time-varying parameters, which we generically denote by \( \lambda_t \). Creal et al. (2012) recommend to use the score to drive the time-variation in the parameter vector \( \lambda_t \), since the score defines the steepest ascent direction for improving the model’s local fit in terms of the likelihood at time \( t \) given the current position of the parameter \( \lambda_t \).

In determining the magnitude of the parameter update, Creal et al. (2012) recommend an autoregressive specification, as well as using a scaled version of the score. The GAS model of order 1 for \( \lambda_t \) is then given by:

\[
\lambda_t = \omega + A_s t - 1 (\lambda_{t-1}, \theta) + B \lambda_{t-1} \tag{6.1}
\]

where \( \omega \) is a vector of constants, \( A \) and \( B \) are coefficient matrices and \( S(\lambda_{t-1}; \theta) \) is a scaling matrix. Creal et al. (2012) recommend to scale the score using a power of the variance of the score:

\[
S^k_t = \frac{1}{\{E_{t-1}[((\nabla_t(\lambda_t, \theta))(\nabla_t(\lambda_t, \theta))^\prime)]\}^d}, \tag{6.4}
\]

with \( E_{t-1}[\cdot] \) the conditional expectation operator with the information available at time \( t - 1 \) as information set. Several choices of powers are considered by the authors: \( d = 0 \) correspond to a steepest ascent step, \( d = 1 \) can be associated to Gauss-Newton updating, while \( d = 1/2 \) appears in the work of Nelson and Foster (1994) as asymptotically optimal in some cases.

This model was applied by Creal et al. (2011) to model the conditional covariance matrix in a single regime case with four assets. A direct application of this approach to the universe of 15 largest US deposit banks is computationally infeasible, because of the curse of dimensionality this model faces. These could be partly solved by imposing parameter restrictions, such as diagonality of the matrices \( A \) and \( B \), but the path dependence of the score then still remains a challenging issue in the multiple regime case. In this paper, we opt to circumvent this problem by using the regime-specific score to drive the parameters of the corresponding regime.

2. The regime-switching \( t \)-GAS volatility model

The parameter of interest is \( h_{it}^k \). The score of the standardized student \( t \) density function of volatility regime \( k \) in (3.6) with respect to \( h_{it}^k \) is:

\[
\nabla_{it}^k = \frac{\partial \log t_{\nu^k}(y_{it}|\mu_{it}^k, h_{it}^k)}{\partial h_{it}^k} = \frac{\nu^k + 1}{2} \frac{1}{(\nu^k - 2) + (y_{it} - \mu_{it}^k)^2/h_{it}^k} (y_{it} - \mu_{it}^k)^2 (h_{it}^k)^{-2} - \frac{1}{2h_{it}^k}. \]

31
We have that $E_{t-1}[^{\nabla k_i}_{it} | s_{it} = k; \mu^k_i, h^k_{it}] = 0$. Conditionally on $s_{it} = k$, $(\nu^k_i/(\nu^k_i - 2))(y_{it} - \mu^k_i)^2/h^k_{it}$ is $F(1, \nu^k_i)$ distributed. Hence, as noted by Harvey and Chakravarty (2008), the score can be rewritten as

$$\nabla^k_{it} = \frac{1}{2h^k_{it}}[(\nu^k_i + 1)b^k_{it} - 1], \quad b^k_{it} = \frac{\nu^k_i (y_{it} - \mu^k_i)^2}{\nu^k_i - 2} h^k_{it}$$

with $b^k_{it}$ a random variable that conditionally on $s_{it} = k$ has a beta distribution of the first kind with shape parameters $1/2$ and $\nu^k_i/2$. The conditional expectation and variance of $b^k_{it}$ are $1/(\nu^k_i + 1)$ and $2\nu^k_i/[(\nu^k_i + 1)^2(\nu^k_i + 3)]$, respectively. Hence, $E_{t-1}[^{\nabla k_i}_{it} | s_{it} = k; \mu^k_i, h^k_{it}] = ((h^k_{it})^2/2\nu^k_i)/(\nu^k_i + 3)$.

Using the asymptotic variance to standardize the score gives the following update equation for the within regime variance when $s_t = k$:

$$h^k_{it} = \omega^k_i + \alpha^k_i \left( 1 + 3/\nu^k_i \right) \left( \frac{\nu^k_i + 1}{\nu^k_i - 2} + \frac{\nu^k_i}{2} \frac{[(y_{it} - \mu^k_i)^2 - 3]}{h^k_{it}} \right) + \beta^k_i h^k_{it-1},$$

(6.5)

with $\omega^k_i$, $\alpha^k_i$ and $\beta^k_i$ time invariant parameters. This is equivalent to the proposed variance equation in (3.4) by setting $\beta^k_i = \beta^k_i - 3\alpha^k_i/\nu^k_i$.

3. The regime-switching t-GAS correlation model

The parameter of interest is $\rho^k_i$. The score of the t-copula function when $s_{st} = k$ is:

$$\nabla^k_i = \frac{\partial \log c_{it-1}(u_{it}, \nu^k_{*t}, R^k_i)}{\partial \rho^k_i}$$

(6.6)

$$= -\left[ \frac{N + \nu^k_{*t}}{\nu^k_{*t} - 2} + (y^k_{it} - \mu^k_i)^2/h^k_{it} \right] \frac{\partial (y^k_{it})'(R^k_i)^{-1}(y^k_{it})}{\partial \rho^k_i} - \frac{1}{2} \frac{\partial \log \det R^k_i}{\partial \rho^k_i} \frac{\partial \rho^k_i}{\partial \rho^k_i}$$

To derive a more explicit formula for the score, we use the intermediary result that:

$$\frac{\partial (y^k_{it})'(R^k_i)^{-1}(y^k_{it})}{\partial \rho^k_i} = -(c^k_i)^2 \left[ b^k_i \sum_{i=1}^{N} \sum_{j \neq i} y^k_{it} y^k_{jt} + a^k_i \sum_{i=1}^{N} (y^k_{it})^2 \right]$$

(6.7)

$$\frac{\partial \log \det (R^k_i)}{\partial \rho^k_i} = (c^k_i)^2 [b^k_i + a^k_i](N - 1)N,$$

(6.8)

where $a^k_i = -\rho^k_i(2 + \rho^k_i(N - 2))$, $b^k_i = (1 + (\rho^k_i)^2(N - 1))$ and $c^k_i = 1/[(1 + (N - 1)\rho^k_i)(1 - \rho^k_i)]$. To obtain these results, we use standard matrix differential calculus\textsuperscript{11} to rewrite:

$$\frac{\partial \log \det (R^k_i)}{\partial \rho^k_i} = \text{tr} \left[ (R^k_i)^{-1} \frac{\partial R^k_i}{\partial \rho^k_i} \right]$$

and

$$\frac{\partial (y^k_{it})'(R^k_i)^{-1}(y^k_{it})}{\partial \rho^k_i} = -\text{tr} \left[ (y^k_{it})' \frac{\partial (y^k_{it})'(R^k_i)^{-1}}{\partial \rho^k_i} (y^k_{it})^{-1} \right] \frac{\partial R^k_i}{\partial \rho^k_i} \frac{\partial \rho^k_i}{\partial \rho^k_i}.$$
\( JJ = NJ \), we further have:

\[
(R^k_t)^{-1} \frac{\partial R^k_t}{\partial \rho^k_t} = c_t^k [a^k_t (J_N - I_N) - \rho^k_t (N - 1) J],
\]

\[
(R^k_t)^{-1} \frac{\partial R^k_t}{\partial \rho^k_t} (R^k_t)^{-1} = (c_t^k)^2 [(a^k_t)^2 (J_N - I_N) + ((\rho^k_t)^2 (N - 1) N - a^k_t \rho^k_t 2 (N - 1)) J_N]
\]

\[
= (c_t^k)^2 [- (a^k_t)^2 I_N + (1 + (\rho^k_t)^2 (N - 1)) J_N].
\]

Taking the trace of these matrices leads to the expression in (6.7)-(6.8).

Combining (6.6) with (6.7)-(6.8) yields:

\[
\nabla^k t = \frac{1}{2} (c_t^k)^2 (N - 1) N \left[ \frac{w^k_i}{(N - 1) N} \sum_{i=1}^{N} \sum_{j \neq i} \tilde{y}^k_{ij} (\tilde{y}^k_{ij} - \rho^k_t) + a^k_t \left( \frac{w^k_i}{N} \sum_{i=1}^{N} \tilde{y}^2_{ij} - 1 \right) \right] \frac{\partial \rho^k_t}{\partial \rho_t},
\]

with \( w^k_i = (N + \nu^k_t)/(\nu^k_t - 2 + (\tilde{y}^k_t)'(R^k_t)^{-1}(\tilde{y}^k_t)) \).

We will use the square root of the conditional variance of the score of the Student \( t \) copula in regime \( k \) to standardize the corresponding score. To derive this conditional variance, we first rewrite the score (6.9) in matrix notation:

\[
\nabla^k t = \frac{1}{2} (c_t^k)^2 (N - 1) N \frac{\partial \rho^k_t}{\partial \rho_t} \lambda^k \det(J^k_t) [w^k_t \text{vec}(\tilde{y}^k_t (\tilde{y}^k_t)') - \text{vec}(R^k_t)],
\]

with \( \lambda \) a \( N^2 \times 1 \) vector of ones, \( \lambda^k \) a \( N^2 \times N^2 \) diagonal matrix with element \((i - 1)N + i\) equal to \( a^k_t \) (for \( i = 1, \ldots, N \)) and \( b^k_t \) otherwise. Conditional on \( s_{st} = k \), \( (R^k_t)^{-1/2} \tilde{y}^k_t \) follows a Student \( t \) distribution with mean 0, covariance matrix \( \lambda^k \) and \( \nu^k_t \) degrees of freedom. It follows from the proof of Theorem I in Creal et al. (2011) that the conditional variance of \( \nabla^k t \) when \( s_{st} = k \) is given by:

\[
S^k_t = \frac{1}{4} (c_t^k)^4 (N - 1)^2 N^2 \left( \frac{\partial \rho^k_t}{\partial \rho_t} \right)^2 \lambda^k \det(J^k_t) [g^k \text{vec}(\tilde{y}^k_t (\tilde{y}^k_t)') - \text{vec}(R^k_t)],
\]

where \( J^k_t = (R^k_t)^{1/2} \otimes (R^k_t)^{1/2}, g^k = (\nu^k_t + N)/(\nu^k_t + 2 + N) \) for the Student \( t \) and the element \( G[i, j] \) of the matrix \( G \) is given by

\[
G[(i - 1)N + l, (j - 1)N + m] = \delta_{ij}\delta_{lm} + \delta_{il}\delta_{jm} + \delta_{im}\delta_{jl},
\]

for \( i, j, l, m = 1, \ldots, N \) and where Kronecker delta \( \delta_{ij} \) is unity when \( i = j \) and 0 otherwise.

Note that for \( \rho^k_t = 0 \) we have \( \lambda^k \det(G(J^k_t)'(A^k_t)'(A^k_t)') = N((2 + N)(a^k_t)^2 + 2(N - 1)(b^k_t)^2) \) and \( \lambda^k \det(G(J^k_t)'(A^k_t)'(A^k_t)') = N^2(a^k_t)^2 \). Unfortunately, there does not seem to be an explicit expression available for \( R^k_t^{1/2} \) for general \( N \) and \( \rho^k_t \). To avoid having to use Cholesky decompositions or other methods to calculate \( (R^k_t)^{1/2} \) in the likelihood functions, we approximate \( \lambda^k \det(G(J^k_t)'(A^k_t)'(A^k_t)') = N^2(a^k_t)^2 \). Unfortunately, there does not seem to be an explicit expression available for \( R^k_t^{1/2} \) for general \( N \) and \( \rho^k_t \). To avoid having to use Cholesky decompositions or other methods to calculate \( (R^k_t)^{1/2} \) in the likelihood functions, we approximate \( \lambda^k \det(G(J^k_t)'(A^k_t)'(A^k_t)') \) and \( \lambda^k \det(G(J^k_t)'(A^k_t)'(A^k_t)') \) with the fit of a piecewise third order polynomial regression of these values on \( \rho^k_t \).
Appendix with additional figures

Figure 7: Left panel shows the time series of absolute weekly stock returns of US deposit banks, together with the sample standard deviation and the single regime volatility estimates under the t-GARCH model of Bollerslev (1986) and the t-GAS model of Creal et al. (2012). Right panel reports the volatility estimates under the model with lowest BIC.
Double regime student t-gas vol model, VIX
volatility regime I
predicted volatility
volatility regime II

Double regime student t-garch vol model, $\nu_1 = \nu_2$, VIX
volatility regime I
predicted volatility
volatility regime II

Double regime student t constant vol model, $\nu_1 = \nu_2$, TED
volatility regime I
predicted volatility
volatility regime II
Mixed student t gas vol model, $\nu_1 = \nu_2$,
volatility regime I
predicted volatility
volatility regime II

Double regime student t gas vol model, VIX
volatility regime I
predicted volatility
volatility regime II

Double regime student t constant vol model, TED
volatility regime I
predicted volatility
volatility regime II
Double regime student t-garch vol model, $\nu_1 = \nu_2$, VIX volatility regime I
predicted volatility
volatility regime II

Double regime student t-garch vol model, $\nu_1 = \nu_2$, TED volatility regime I
predicted volatility
volatility regime II

Double regime student t-garch vol model, $\nu_1 = \nu_2$, TED volatility regime I
predicted volatility
volatility regime II

Double regime student t-garch vol model, $\nu_1 = \nu_2$, TED volatility regime I
predicted volatility
volatility regime II
68144 / Suntrust Banks
Double regime student t-gas vol model, VIX
volatility regime I
predicted volatility
volatility regime II

69032 / Morgan Stanley
Double regime gaussian-gas vol model, TED
volatility regime I
predicted volatility
volatility regime II

70519 / Citigroup
Double regime student t-garch vol model, STLFSI
volatility regime I
predicted volatility
volatility regime II
Double regime student t garch vol model, STLFSI
volatility regime I
predicted volatility
volatility regime II
References


98. "Dynamics on monetary policy in a fair wage model of the business cycle", by D. De la Croix, G. de Walque and R. Wouters, Research series, October 2006.
102. "Fiscal sustainability indicators and policy design in the face of ageing", by G. Langenus, Research series, October 2006.


224. "Asymmetric information in credit markets, bank leverage cycles and macroeconomic dynamics", by A. Rannenberg, Research series, April 2012.

