SUPPLEMENTARY MATERIAL FOR

Generalized Autoregressive Score Trees and Forests

Andrew J. Patton and Yasin Simsek

This Version: February 29, 2024

This appendix contains two sections. Section S.1 presents additional details for the t-GAS copula model considered in Section 3.4. Section S.2 presents some additional tables and figures.

S.1 Derivation of the score function for the t-GAS copula

In this section we present the score function for the t-copula analysis discussed in Section 3. We refer to Creal et al. (2013) for the details of the univariate applications (the GARCH and t-GAS models).

S.1.1 Notation

We adopt the notation of Creal et al. (2011) for ease of comparability with that article. The Kronecker product is denoted by $A \otimes B$ for any matrices A and B. $A_{\otimes}$ stands for $A \otimes A$. The function $vec(A)$ vectorizes matrix A into a column vector, and $vech(A)$ vectorizes just the lower triangle of A, which eliminates duplicates in the case that A is symmetric. The duplication matrix is implicitly defined as the solution to $\mathcal{D}\,\mathrm{vech}(A) = \mathrm{vec}(A)$. Finally, $\mathbb{E}_{t-1}$ denotes the expectation conditional on the information available up to period $t-1$.

S.1.2 The probability density function of t copula

We adopt Student's t copula specification in our empirical analysis and its probability density function is given by

$$c(\mathbf{u}_t;\Sigma_t,\nu)=\frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{|\Sigma_t|}\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^2}\left(1+\frac{\mathbf{x}_t'\Sigma_t^{-1}\mathbf{x}_t}{\nu}\right)^{-\frac{\nu+2}{2}}\prod_{i=1}^{2}\left(1+\frac{x_{i,t}^2}{\nu}\right)^{\frac{\nu+1}{2}} \tag{12}$$

where $\mathbf{x}_t=[x_{1,t},x_{2,t}]=[T_\nu^{-1}(u_{1,t}),T_\nu^{-1}(u_{2,t})]'$ obtained by applying the inverse of the univariate t distribution with ν degrees of freedom, $\Gamma(\cdot)$ is gamma function and Σ_t is 2-by-2 correlation matrix. We denote the off-diagonal element of Σ_t with ρ_t which is the variable of interest:

$$\Sigma_t=\begin{bmatrix}1 & \rho_t\\ \rho_t & 1\end{bmatrix} \tag{13}$$

S.1.3 The score and information matrix

We use inverse information matrix of the score function as a scaling factor in all applications. Given the complex structure of the Student's t copula, derivation of the information matrix requires tedious calculations, but Creal et al. (2011) provide a closed-form formula of both score and information matrix. Based on their results, we can write

$$\begin{aligned}\nabla_t &= \tfrac{\partial \log c_t(y_t|\Sigma_t;\nu)}{\partial f_t}\\ &= \tfrac{1}{2}(\mathcal{D}\Psi_t)'\Sigma_{t\otimes}^{-1}\left[w_t\mathbf{x}_{t\otimes}-\operatorname{vec}(\Sigma_t)\right]\\ \mathcal{I}_{t|t-1} &= \mathbb{E}_{t-1}\left[\nabla_t\nabla_t'\right]\\ &= \tfrac{1}{4}(\mathcal{D}\Psi_t)'J_{t\otimes}'\left[gG-\operatorname{vec}(\mathrm{I})\operatorname{vec}(\mathrm{I})'\right]J_{t\otimes}\mathcal{D}\Psi_t\end{aligned} \tag{14}$$

where $\Psi_t\equiv\frac{\partial\,\mathrm{vech}(\Sigma_t)}{\partial\rho_t}$, J_t is such that $\Sigma_t^{-1}=J_t'J_t$, $w_t\equiv\frac{\nu+2}{\nu-2+\mathbf{x}_t'\Sigma_t^{-1}\mathbf{x}_t}$, $g\equiv\frac{v+2}{v+4}$, and the explicit form of matrix G is

$$G=\begin{bmatrix}3&0&0&1\\0&1&1&0\\0&1&1&0\\1&0&0&3\end{bmatrix}. \tag{15}$$

We define the scaled score functions as $s_t = \mathcal{I}_{t|t-1}^{-1} \nabla_t$. As in Janus et al. (2014) we use a transformation to ensure $\rho_t \in (-1, 1)$, by setting $\rho_t = \frac{1-\exp(-\tilde{\rho}_t)}{1+\exp(-\tilde{\rho}_t)}$, where $\tilde{\rho}_t \in \mathbb{R}$. In order to obtain the scaled score function for $\tilde{\rho}_t$, we multiply the original scaled score with the derivative of the transformation function: $\tilde{s}_t = \frac{\partial \tilde{\rho}_t}{\partial \rho_t} s_t$. When we use the explicit form of each component in equation (14), we obtain the following expression for the scaled score of the t copula:

$$s_t = \left(\frac{2}{1-\rho_t^2}\right)\left(\frac{1+\rho_t^2}{g+(2g-1)\rho_t^2}\right)\left(w_t(x_{1,t}x_{2,t}-\rho_t) - \frac{\rho_t}{1+\rho_t^2}(w_t x_{1,t}^2 + w_t x_{2,t}^2 - 2)\right) \qquad (16)$$

Noting that $(g, w_t) \to (1, 1)$ as $\nu \to \infty$, we thus also obtain the scaled score function for Gaussian copula:

$$s_t = \left(\frac{2}{1-\rho_t^2}\right)\left(x_{1,t}x_{2,t} - \rho_t - \frac{\rho_t}{1+\rho_t^2}(x_{1,t}^2 + x_{2,t}^2 - 2)\right). \qquad (17)$$

S.2 Additional tables and figures

Table S.1: Out-of-sample performance of t-GAS models using QLIKE loss

	Benchmark	DRF	Tiny Tree	Tree	Forest
Benchmark					
DRF	-1.332				
Tiny Tree	-2.006	-0.789			
Tree	-5.277	-3.856	-5.594		
Forest	-3.652	-2.599	-0.871	3.154	
QLIKE	0.403	0.382	0.369	0.324	0.358

Table S.2: Out-of-sample performance of GARCH models using -logL loss

	Benchmark	DRF	Tiny Tree	Tree	Forest
Benchmark					
DRF	-3.025				
Tiny Tree	-2.300	0.350			
Tree	-5.777	-2.356	-5.818		
Forest	-6.703	-2.392	-2.183	1.594	
-logL	1.205	1.177	1.182	1.148	1.162

Figure S.1: Parameter estimates as a function of T10Y state variable for t Copula forest-based GAS

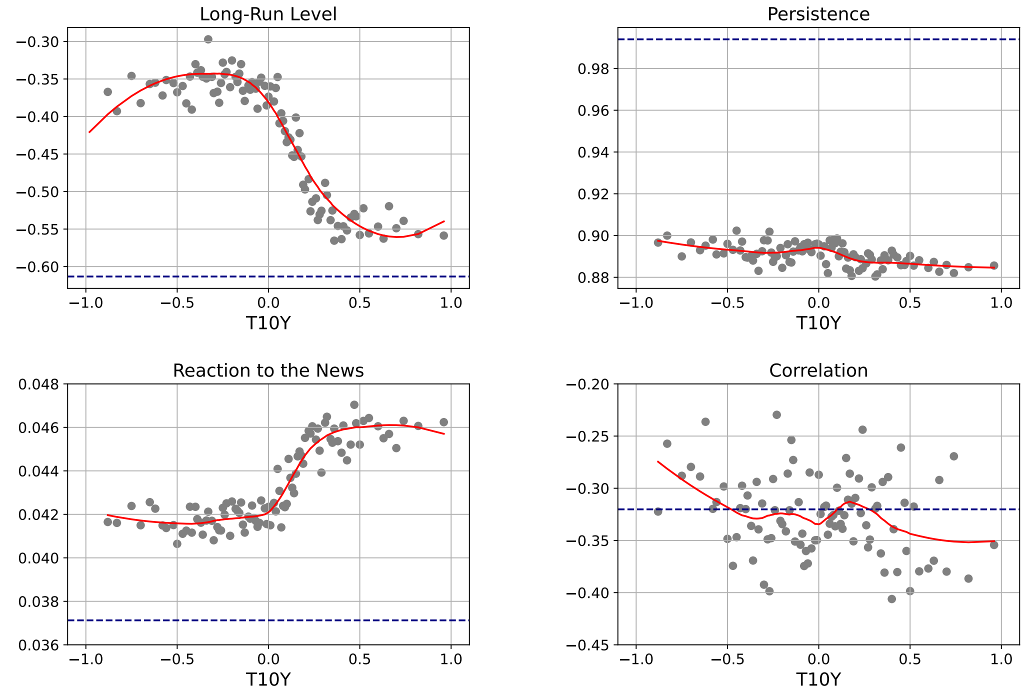

Figure S.2: Parameter estimates as a function of T10Y3M state variable for t Copula forest-based GAS

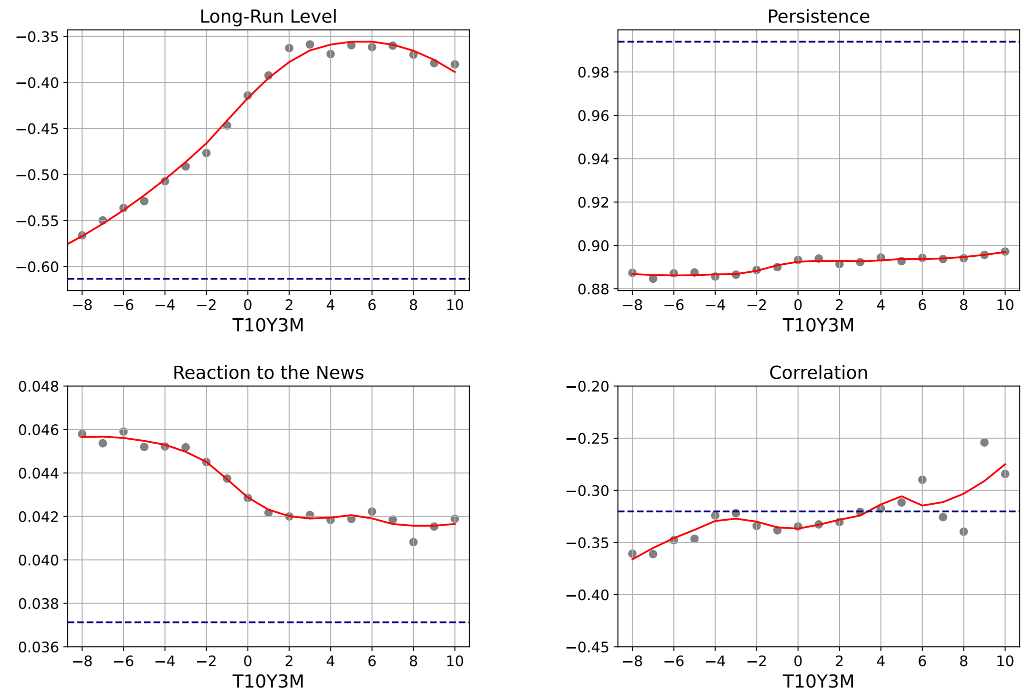

Figure S.3: Parameter estimates as a function of VIX state variable for t Copula forest-based GAS

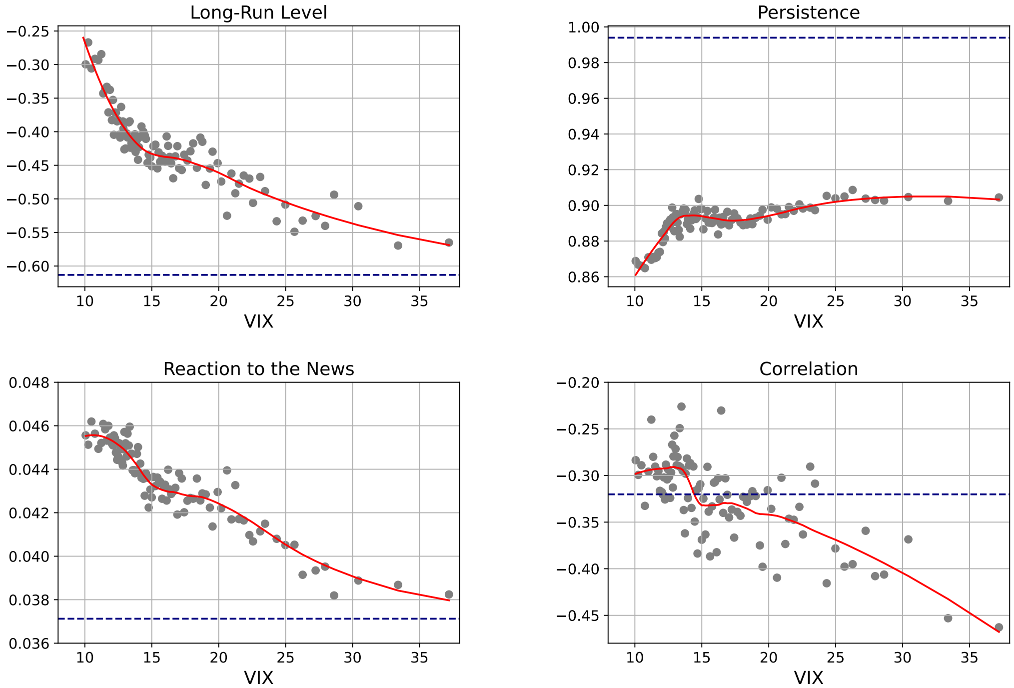

Figure S.4: **The estimated ACD tree model.** This figure depicts the tree structure for the ACD model. The tree's splits are based on SPX and T10Y, which refer to the S&P 500 return and 10 year bond return respectively.

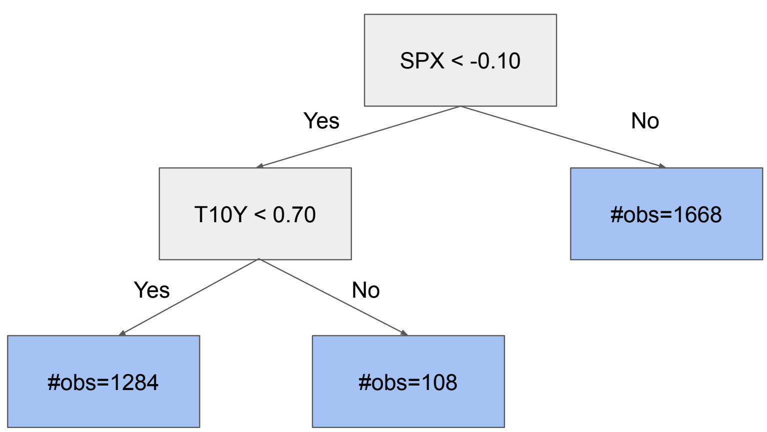

Figure S.5: Parameter estimates as a function of VIX state variable for ACD forest-based

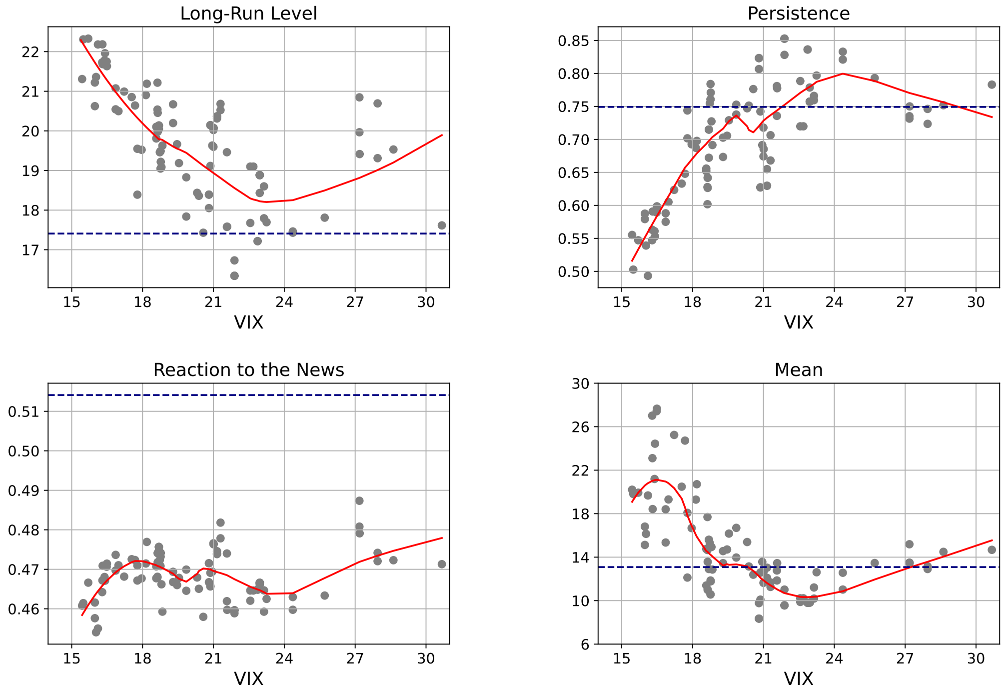

Figure S.6: Parameter estimates as a function of T10Y3M state variable for ACD forest-based

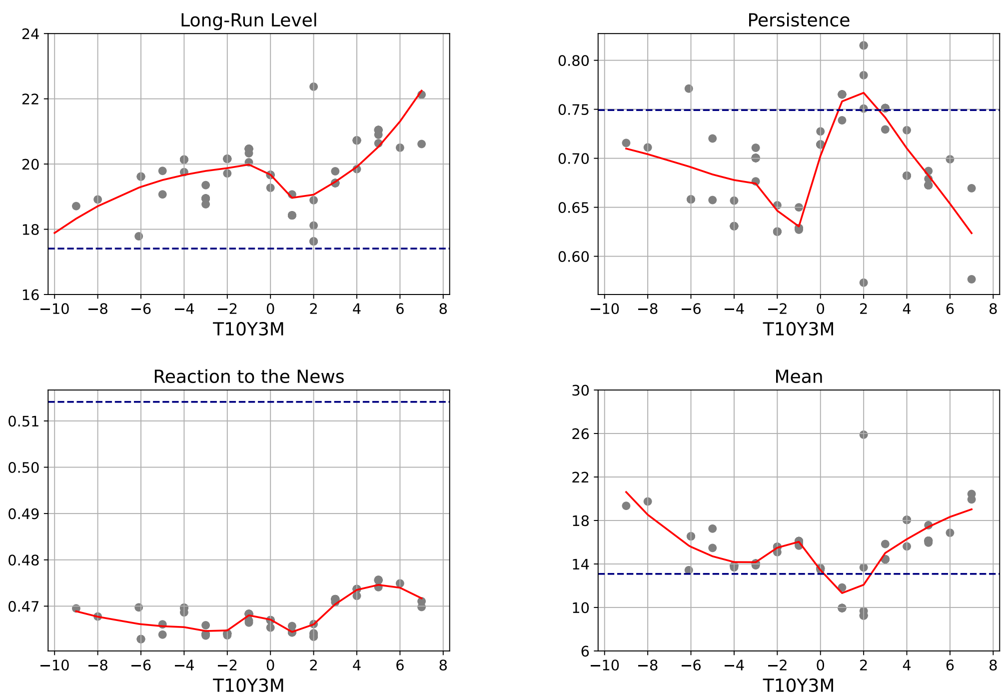

Figure S.7: Parameter estimates as a function of DURATION state variable for ACD forest-based

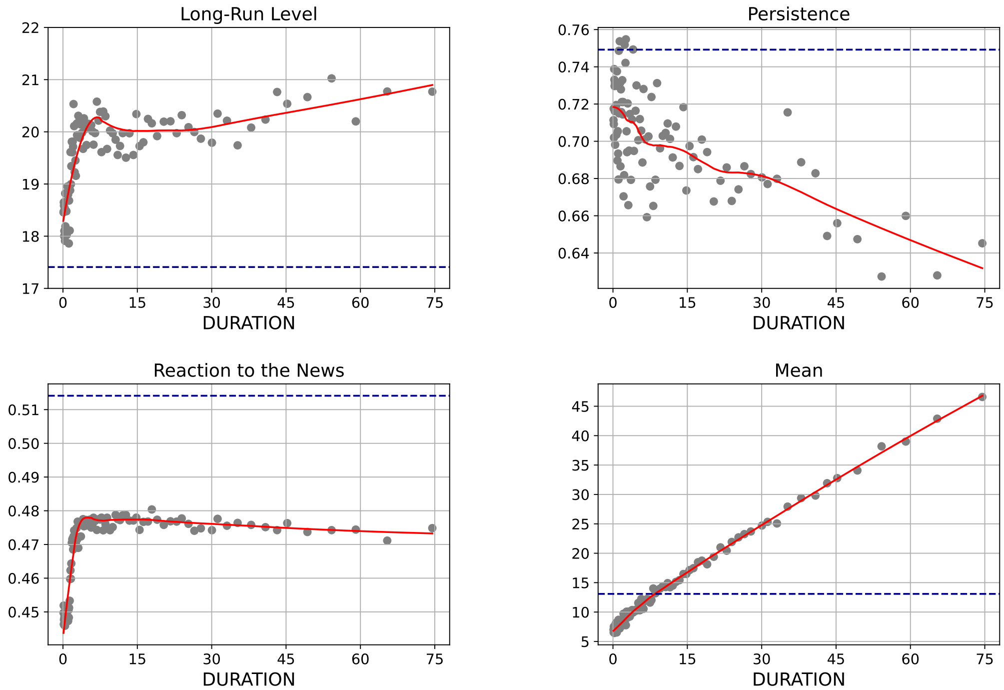

Figure S.8: Parameter estimates as a function of LIQUIDITY state variable for ACD forest-based