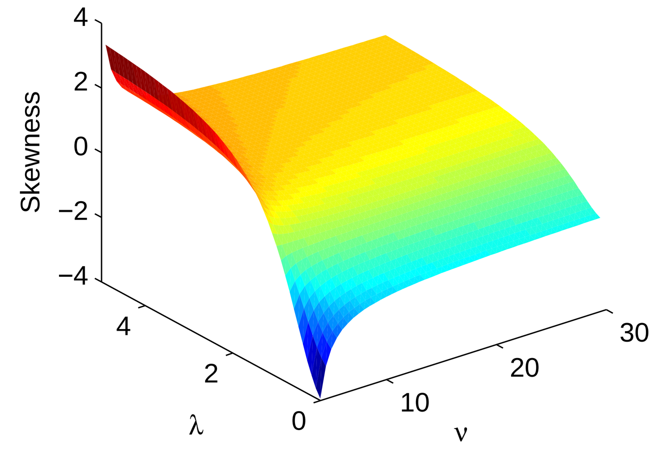
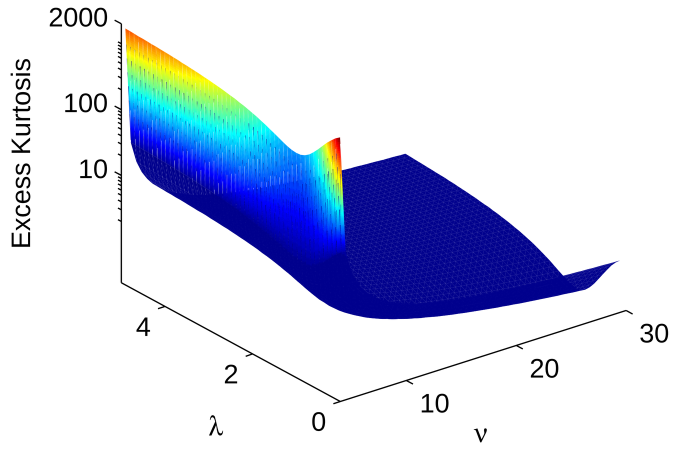




Smooth Mixtures of Asymmetric Student t Densities

Feng Li, Department of Statistics, Stockholm University, Sweden

joint with Mattias Villani and Robert Kohn

Stockholm University

Contributions

A general model is proposed for flexibly estimating the density of a continuous response variable conditional on a possibly high-dimensional set of covariates. The model is a finite mixture of asymmetric student-t densities with covariate-dependent mixture weights. The four parameters of the components, the mean, degrees of freedom, scale and skewness, are all modeled as functions of the covariates.

Applications

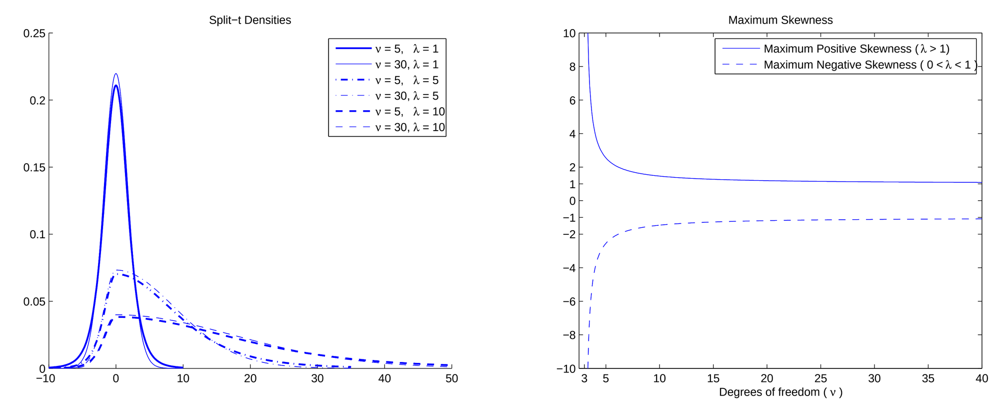
Financial data, such as stock market returns, are typically heavy tailed and subject to volatility clustering. The model is applied to analyse the distribution of daily stock market returns conditional on nine covariates and outperforms widely used GARCH models and other recently proposed mixture models in an out-of-sample evaluation of returns during the recent financial crisis.

Asymmetric Student t

The split-t density is

$$f(y;\mu,\phi,\lambda,\nu)=\begin{cases} c\cdot\kappa(y;\mu,\phi,\nu); & \text{if } y\le\mu,\\ c\cdot\kappa(y;\mu,\lambda\phi,\nu); & \text{if } y>\mu,\end{cases}$$

where $\kappa(y;\mu,\phi,\nu)$ is the kernel of student t density and c is the normalization constant.



Basic Properties:

1. $\lambda<1$, skewed to the left; $\lambda>1$, skewed to the right;
2. $\lambda=1$, reduces to symmetric student-t distribution;
3. $\nu\to\infty$, reduces to the two-piece normal distribution;
4. $\lambda=1$ and $\nu\to\infty$, reduces to normal distribution.

Mixture of Split-t

Given x, a finite mixture distribution $p(y|x)$ is

$$\sum_{k=1}^{K}\omega_k f_k(y_i|\theta_k),\ i=1,\dots,n.$$

A smooth mixture model is a finite mixture density with weights that are smooth function of the covariates, e.g

$$\omega_k(x)=\frac{\exp(x'\gamma_k)}{\sum_{r=1}^{K}\exp(x'\gamma_r)}$$

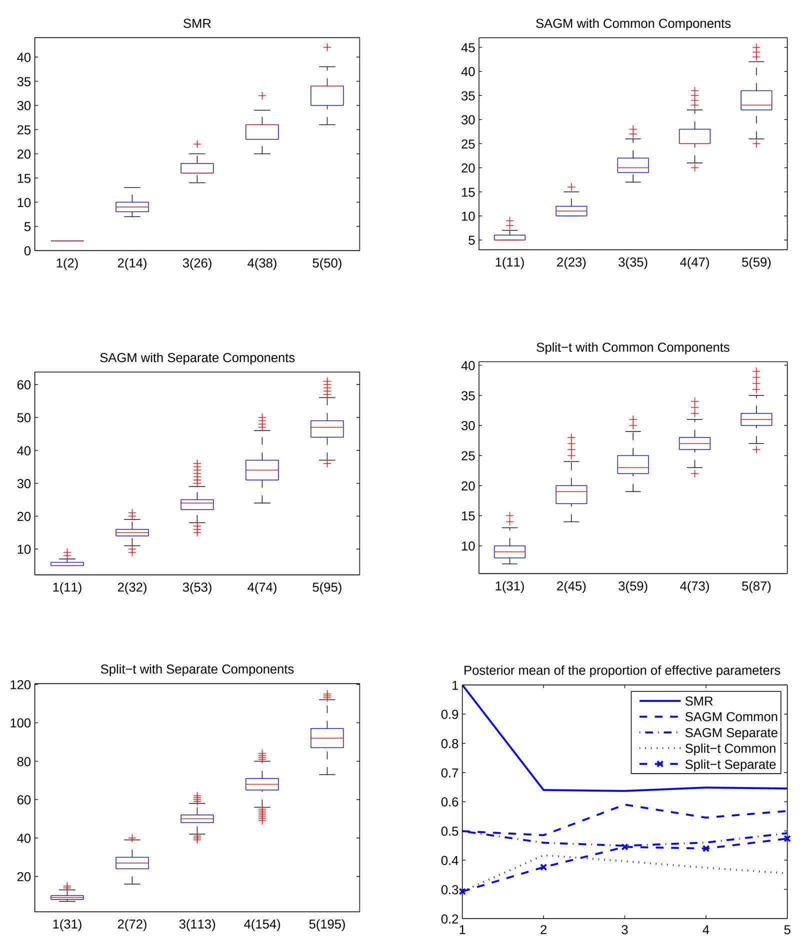
In a smooth mixture of split-t densities, the four components μ,ϕ,λ and ν in each mixture components are connected to covariates as, e.g.,

$$\mu=\beta_{\mu_0}+x_t'\beta_\mu,\quad \ln\phi=\beta_{\phi_0}+x_t'\beta_\phi,$$
$$\ln\nu=\beta_{\nu_0}+x_t'\beta_\nu,\quad \ln\lambda=\beta_{\lambda_0}+x_t'\beta_\lambda.$$

Features are common if only the intercepts are allowed to be different across components. This often an empirically relevant, simplification of the model.

Key Features

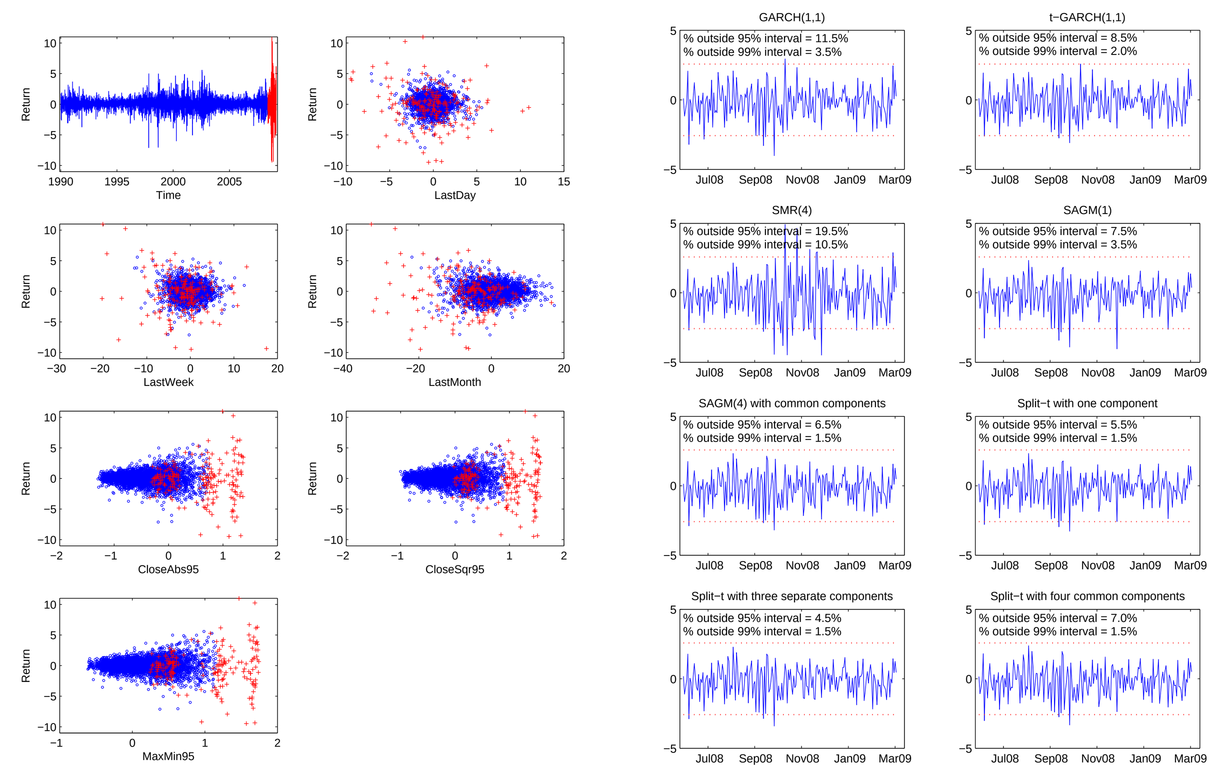
1. *Complex-but-few* approach – Enough flexibility is used within the mixture components so that the number of components can be kept to a minimum.
2. *Bayesian variable selection* are used in all five sets of covariates in mean, scale, skewness and kurtosis parameters and gating function to automatically determine important variables.



The figure displays the posterior distributions of the number of included parameters in the S&P 500 data. On first five subplots, the horizontal axis measures the number of components (potential number of parameters in parentheses) and the vertical axis is the total number of effective parameters after variable selection. The right-bottom subplot is the posterior mean of the proportion of effective parameters in each model.

3. *Finite-step Newton's method* within variable selection is used to sample the parameters which can greatly speed up the algorithm.
4. *Five-fold cross-validation* of the log predictive density score(LPDS) is used in model comparison.

Daily S&P 500 Returns

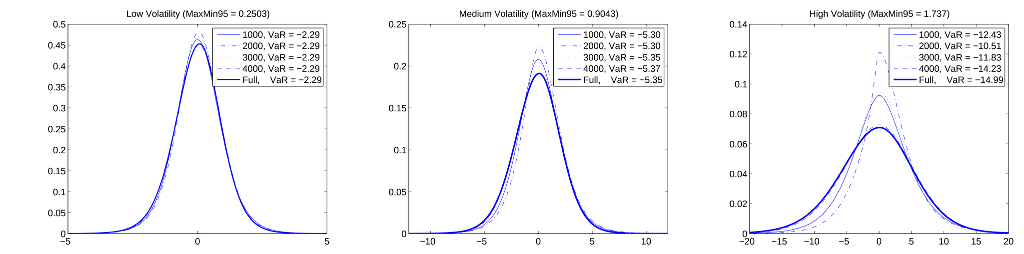


We use normalized residuals to assess the quality of the predictive densities. The SMR (*simple-and-many* approach) with largest LPDS produces much to large residuals during the most volatile period, and so does the GARCH(1,1) and t-GARCH(1, 1). The one-component split-t model is doing remarkably well during this very difficult time period.

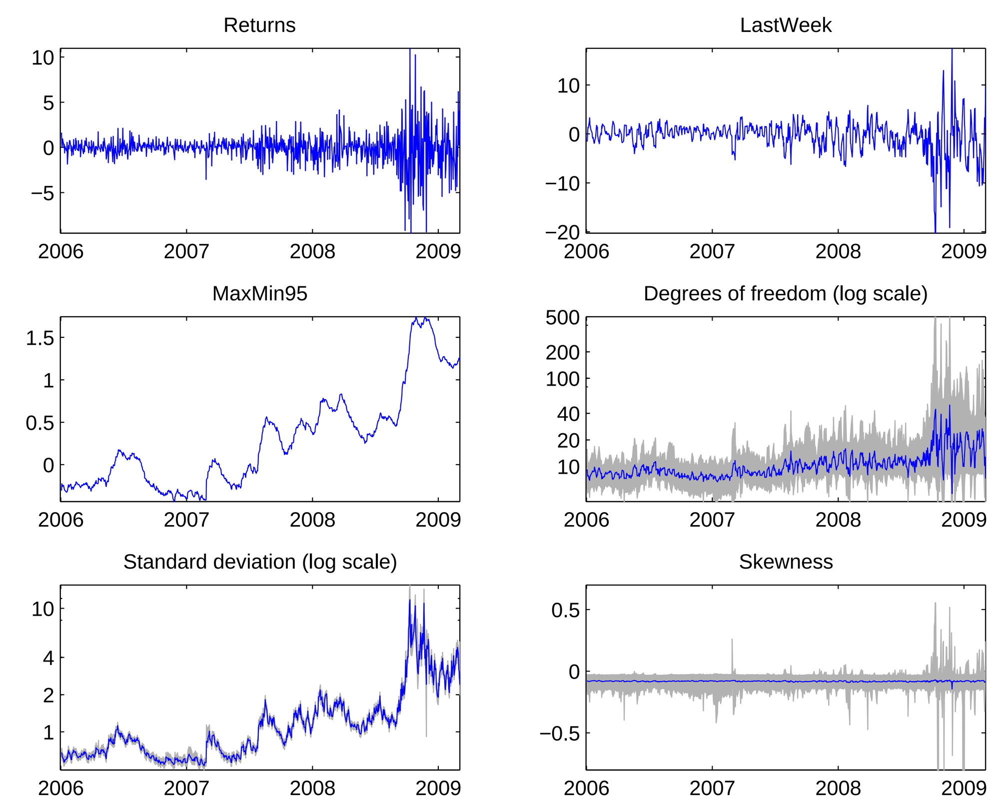
Model	$K=1$	$K=2$	$K=3$	$K=4$	$K=5$	Max n.s.e.
SMR	**−1044.78**	−638.89	−505.74	−487.11	−489.19	0.98 (3)
+ Skew	−540.91	−525.07	−513.85	−506.68	−506.13	0.82 (2)
+ DF	−544.00	−518.71	−498.93	−500.14	−494.29	0.89 (1)
+ Skew + DF	−530.86	−504.63	−498.03	−498.83	**−496.87**	0.88 (5)
SAGM Common	−477.73	−473.10	−473.12	−470.30	−472.86	0.26 (2)
+ Skew	−474.18	−467.29	−468.75	−467.93	−467.22	0.35 (4)
+ DF	−474.74	−472.92	−470.51	−469.40	−468.87	0.34 (4)
+ Skew + DF	**−472.37**	−468.92	−469.30	−466.21	**−465.86**	0.53 (4)
SAGM Separate		−469.21	−469.50	−470.53	−471.02	0.49 (3)
+ Skew		−468.48	−466.93	−467.48	−468.02	0.58 (4)
+ DF		−469.08	−469.24	**−462.03**	−467.78	0.72 (5)
+ Skew + DF		**−466.84**	**−462.56**	−462.47	−474.58	0.74 (5)
GARCH(1,1)	−479.03					
t-GARCH(1,1)	−477.39					

LPDS shows that the SMR model does poorly, even with a large number of components, and is outperformed by the GARCH(1, 1) and t-GARCH(1, 1) models. The split-t with covariate dependent scale, skewness and degrees of freedom is the best one-component model.

Posterior Study



We estimate one-component split-t model using first 1000, 2000, 3000, 4000 trading days and the full sample. Then we obtain conditional predictive densities for three sets of covariates low, medium and high volatility. It is shown that the predictive densities is stable even with very high volatility data.



The median of the degrees of freedom actually increased during the most volatile part of the financial crisis (but at the same time the scale parameter rose dramatically to bring about a very large boost in standard deviation of returns), but, during some spells, the posterior distribution of ν also has a long left tail with substantial probability mass on very small values of ν.

Footnotes

Li (feng.li@stat.su.se) is a Ph.D student at Stockholm University. Villani (mattias.villani@gmail.com) is a researcher at Sveriges Riksbank and senior lecturer at Stockholm University. Kohn (r.kohn@unsw.edu.au) is Scientia Professor at the Australian School of Business, University of New South Wales.

Kohn was partially supported by ARC Grant DP0667069. This paper is forthcoming in *Journal of Statistical Planning and Inference*, doi:10.1016/j.jspi.2010.04.031.