Supporting Information published online only to accompany the paper "A Gaussian mixture autoregressive model for univariate time series" by Leena Kalliovirta, Mika Meitz, and Pentti Saikkonen.

1. Non-ergodicity of a process having all finite dimensional distributions being Gaussian mixtures

In the discussion following Theorem 1 it was pointed out that non-ergodicity would be an undesirable implication for a process having all finite dimensional distributions being Gaussian mixtures. To see that this holds in a particular special case, suppose all finite dimensional distributions of a process x_t, say, are Gaussian mixtures of the form (9) so that, for any $T \geq 1$, the distribution of a realization $(x_1, \ldots, x_T)$ is

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \sum_{m=1}^{M} \alpha_m \mathsf{n}_T \left(\boldsymbol{x}; \boldsymbol{\vartheta}_m\right),$$

where the density function $\mathsf{n}_T \left(\boldsymbol{x}; \boldsymbol{\vartheta}_m\right)$ is a T–dimensional analog of that in (7). The process x_t is clearly stationary. For simplicity, consider the special case where $M = 2$, $\varphi_{m,i} = 0$ $(i = 1, \ldots, p,\ m = 1, 2)$, $\sigma_1 = \sigma_2 = \sigma$, and $\varphi_{1,0} \neq \varphi_{2,0}$. Then $\mathsf{n}_T \left(\boldsymbol{x}; \boldsymbol{\vartheta}_i\right)$ is the joint density of T independent Gaussian random variables with mean $\varphi_{i,0}$ and variance σ^2 $(i = 1, 2)$. This means that, for every T,

$$(x_1, \ldots, x_T) \sim \begin{cases} \mathsf{n}_T \left(\boldsymbol{x}; \boldsymbol{\vartheta}_1\right), & \text{with probability } \alpha_1 \\ \mathsf{n}_T \left(\boldsymbol{x}; \boldsymbol{\vartheta}_2\right), & \text{with probability } 1 - \alpha_1. \end{cases}$$

This implies that, for every T, the sample mean $\bar{X}_T = T^{-1} \sum_{t=1}^{T} x_t$ is distributed as $N(\varphi_{1,0}, \sigma^2/T)$ with probability α_1 and as $N(\varphi_{2,0}, \sigma^2/T)$ with probability $1 - \alpha_1$. As $\varphi_{1,0} \neq \varphi_{2,0}$ and $0 < \alpha_1 < 1$ is assumed, it is therefore immediate that no law of large numbers holds, and consequently the process x_t cannot be ergodic. Indeed, it is not difficult to check that $\bar{X}_T$ converges in distribution to a random variable taking the values $\varphi_{1,0}$ and $\varphi_{2,0}$ with probability α_1 and $1 - \alpha_1$, respectively.

2. Graphical illustration of the different mixing weights

Here we provide a graphical illustration comparing the mixing weights in the mixture AR models discussed in Section 2.4. We also make some remarks on the mixing weights used in the model of Lanne and Saikkonen (2003) and on the transition functions used in some STAR models.

In the top panels of Figure A.1, we plot the mixing weight $\alpha_{1,t}$ of the GMAR model as a function of $y_{t-1} = y$ in the case $M = 2$, $p = 1$, with certain parameter combinations. The bottom left panel shows $\alpha_{1,t}$ in some cases for the LMAR model of Wong and Li (2001b); in the model of Lanne and Saikkonen (2003) $\alpha_{1,t}$ behaves in a comparable way (no picture presented). The two pictures on the left illustrate that the three models can produce mixing weights of similar monotonically increasing patterns. The figure in the top left panel also illustrates the fact that, other things being equal, a decrease in the value of α_m decreases the value of the mixing weight $\alpha_{m,t}$ of the GMAR model. In the conditional expectation of a basic logistic two-regime STAR model, referred to as the LSTAR1 model in Teräsvirta et al. (2010, Sec. 3.4.1), the transition function, which is the counterpart of the mixing weight $\alpha_{1,t}$, also behaves in a similar monotonically increasing way. Given these observations it is interesting that with suitable parameter values the GMAR model can produce nonmonotonic mixing weights even in the case $M = 2$. The top right panel illustrates this. The models of Wong and Li (2001b) and Lanne and Saikkonen (2003) can produce mixing weights of this form only when $M > 2$. Similarly, in LSTAR models a transition function of this form cannot be obtained with a LSTAR1 model. For that one needs an LSTAR2 model (see Teräsvirta et al. (2010, Sec. 3.4.1)) or some other similar model such as the exponential autoregressive model of Haggan and Ozaki (1981). Thus, once the number of component models is specified the GMAR model appears more flexible in terms of the form of mixing weights than the aforementioned previous mixture models and the same is true when the mixing weights of the GMAR model are compared

to the transition functions of LSTAR models.

As far as the mixing weights of the C-STAR model of Dueker et al. (2007) are concerned they can be nonmonotonic, as illustrated in the bottom right panel of Figure A.1. After trying a number of different parameter combinations it seems, however, that (at least in the case $p = 1$) nonmonotonic mixing weights are rather special for this model. The first four (monotonic) graphs in the bottom right panel correspond to parameter configurations in Table 2 of Dueker et al. (2007). The fourth one is interesting in that it produces a nearly constant graph (the graph would be constant if the values of the standard deviations σ_1 and σ_2 were changed to be equal). Finally, note that a common convenience of the GMAR model as well as of the models of Wong and Li (2001b) and Dueker et al. (2007) is that there is no need to choose a threshold variable (typically y_{t-d}) as in the model of Lanne and Saikkonen (2003) or in TAR and STAR models.

3. Graphical analysis of quantile residuals of the GMAR$^{\mathrm{C}}$ model in Table 1

Here we provide additional graphical analysis (mentioned in Section 4.1 of the paper) of the quantile residuals of the GMAR$^{\mathrm{C}}$ model in Table 1. According to the diagnostic tests based on quantile residuals (see Table 1), the restricted GMAR$^{\mathrm{C}}$ model provides a good fit to the data. To further investigate the properties of the quantile residuals, Figure A.2 depicts time series and QQ–plots of the quantile residuals as well as the first ten standardized sample autocovariances of quantile residuals and their squares (the employed standardization is such that, under correct specification, the distribution of the sample autocovariances is approximately standard normal). The time series of quantile residuals computed from a correctly specified model should resemble a realization from an independent standard normal sequence. The graph of quantile residuals and the related QQ–plot give no obvious reason to suspect this, although some large positive quantile residuals occur. According to the approximate 99% critical bounds only two somewhat larger autocovariances are seen but even they are found at larger lags (we use 99% critical

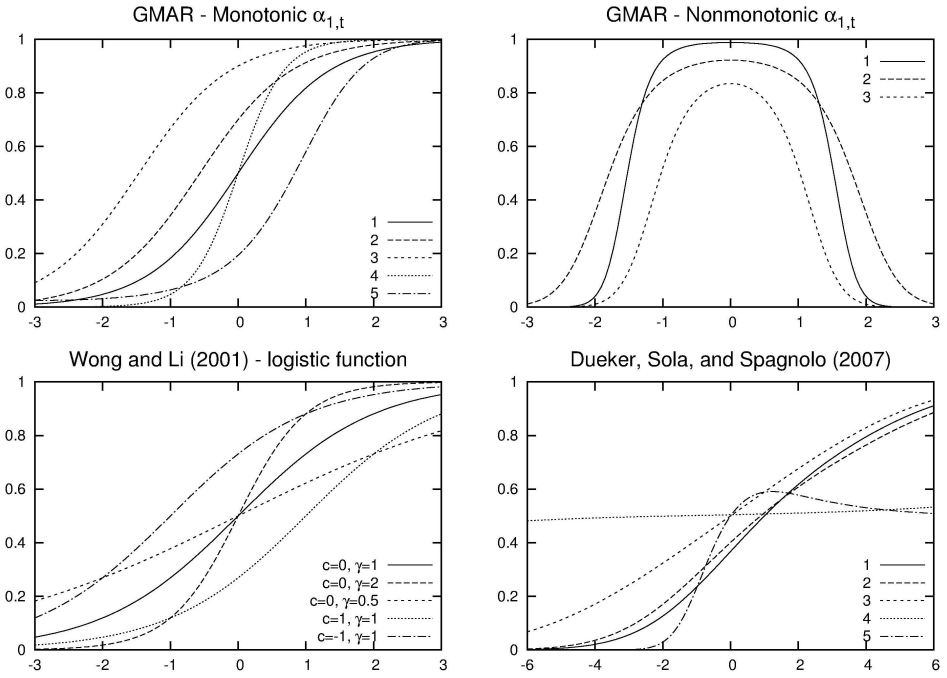

Figure A.1: **Top left panel:** $\alpha_{1,t}$ in the GMAR model ($p = 1$) as a function of y_{t-1}. Parameter values used: model 1: $\varphi_{1,0} = 0.5$, $\varphi_{2,0} = -0.5$, $\varphi_{1,1} = \varphi_{2,1} = 0.5$, $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_1 = 0.5$; model 2: same as model 1 but $\alpha_1 = 0.7$; model 3: same as model 1 but $\alpha_1 = 0.9$; model 4: $\varphi_{1,0} = 1$, $\varphi_{2,0} = -1$, $\varphi_{1,1} = \varphi_{2,1} = 0.5$, $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_1 = 0.5$; model 5: $\varphi_{1,0} = \varphi_{2,0} = 0.5$, $\varphi_{1,1}0.75$, $\varphi_{2,1} = 0.25$, $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_1 = 0.5$. **Top rigth panel:** $\alpha_{1,t}$ in the GMAR model ($p = 1$) as a function of y_{t-1}. Parameter values used: model 1: $\varphi_{1,0} = \varphi_{2,0} = 0$, $\varphi_{1,1} = 0.2$, $\varphi_{2,1} = 0.9$, $\sigma_1^2 = 0.25$, $\sigma_2^2 = 4$, $\alpha_1 = 0.5$; model 2: same as model 1 but $\sigma_1^2 = \sigma_2^2 = 0.5$, $\alpha_1 = 0.7$; model 3: same as model 1 but $\sigma_2^2 = 0.25$. **Bottom left panel:** $\alpha_{1,t}$ in the model of Wong and Li (2001b) as a function of y_{t-1}. Logistic equation assumed to be of the form $\log(\alpha_{1,t}/\alpha_{2,t}) = \gamma(y_{t-1} - c)$, or equivalently, $\alpha_{1,t} = \frac{1}{1+e^{-\gamma(y_{t-1}-c)}}$. Note that this is exactly the standard form of the logistic function. Curves correspond to different values of c and γ. **Bottom right panel:** $\alpha_{2,t} = 1-\alpha_{1,t}$ in the model of Dueker et al. (2007) as a function of y_{t-1}. Parameter values used: model 1: $c = 1$, $\varphi_{1,0} = -0.5$, $\varphi_{2,0} = 0.5$, $\varphi_{1,1} = \varphi_{2,1} = 0.9$, $\sigma_1 = 3$, $\sigma_2 = 2$; model 2: $c = 1$, $\varphi_{1,0} = -1$, $\varphi_{2,0} = 1$, $\varphi_{1,1} = \varphi_{2,1} = 0.9$, $\sigma_1 = 3$, $\sigma_2 = 2$; model 3: $c = 0$, $\varphi_{1,0} = -1$, $\varphi_{2,0} = 1$, $\varphi_{1,1} = \varphi_{2,1} = 0.9$, $\sigma_1 = \sigma_2 = 3$; model 4: $c = 0$, $\varphi_{1,0} = -10$, $\varphi_{2,0} = 10$, $\varphi_{1,1} = \varphi_{2,1} = 0.7$, $\sigma_1 = 5$, $\sigma_2 = 4$; model 5: $c = 0$, $\varphi_{1,0} = \varphi_{2,0} = 0$, $\varphi_{1,1} = -0.3$, $\varphi_{2,1} = 0.3$, $\sigma_1 = 1$, $\sigma_2 = 0.25$.

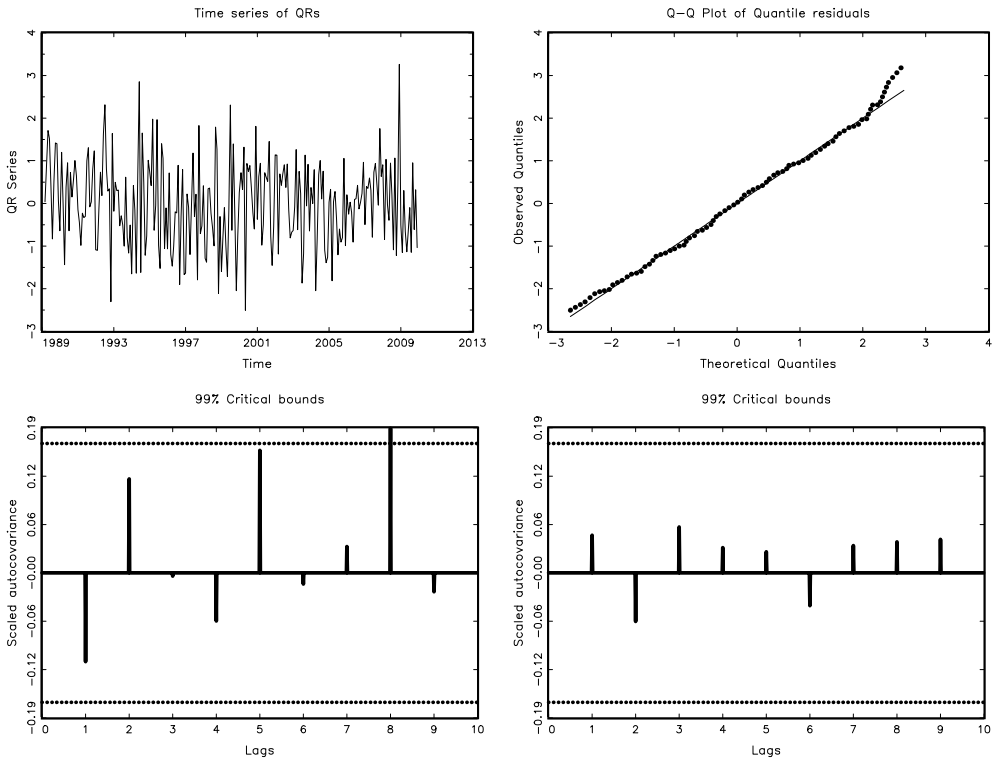

Figure A.2: Diagnostics of the restricted GMAR$^{\mathrm{C}}$ model described in Table 1: Time series of quantile residuals (top left panel), QQ-plot of quantile residuals (top right panel), and ten first scaled autocovariances of quantile residuals and squared quantile residuals (bottom left and right panels, respectively). The lines in the bottom panels show approximate 99% critical bounds.

bounds because, from the viewpoint of statistical testing, several tests are performed). It is particularly encouraging that the GMAR$^{\mathrm{C}}$ model has been able to accommodate for the conditional heteroskedasticity in the data (see the bottom right panel of Figure A.2), unlike the considered linear AR models (see the diagnostic tests for AR(4) model in Table 1). Thus, unlike the linear AR models, the GMAR$^{\mathrm{C}}$ model seems to provide an adequate description for the interest rate differential series.

Additional Reference

HAGGAN, V., AND T. OZAKI (1981): "Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model," *Biometrika*, 68, 189–196.