On the Transmission of Memory: Inflation Persistence and the Great Moderation*

Christian Conrad†, Simone Elmer‡ and Menelaos Karanasos§

†University of Heidelberg, Heidelberg, Germany

‡KOF Swiss Economic Institute, Zurich, Switzerland

§Brunel University, West London, UK

Very preliminary first draft:
July 2009

Abstract

In this paper we derive the autocorrelation function, the impulse response function and the optimal predictor for the AR(1)-GARCH(1,1)-in-mean-level process. We show that in this model the persistence from the variance is transmitted to the mean and vice versa and, hence, by studying the conditional mean/variance independently one will tend to overestimate the true degree of persistence. Under reasonable assumptions, the AR(1)-GARCH(1,1)-in-mean-level process will be observationally equivalent to an ARMA(2,1) process with the largest autoregressive root being close to one. In particular, unit root tests for the level will not be able to reject the null hypothesis of the process being integrated of order one. We argue that the commonly observed decrease in U.S. inflation persistence and inflation uncertainty can be well explained by our model in combination with a change in monetary policy in the early 1980's.

Keywords: GARCH-in-Mean, persistence, unit root test, conditional heteroscedasticity.

JEL Classification: E31, E58, C12, C22, C52.

*We would like to thank Michael J. Lamla for helpful comments and suggestions.

†Corresponding author: Christian Conrad, University of Heidelberg, Alfred-Weber-Institute, Grabengasse 14, 69117 Heidelberg, Germany; email: christian.conrad@awi.uni-heidelberg.de, tel: +49 (0)6221 54 3173, fax: +49 (0)6221 54 3578.

‡Simone Elmer, KOF Swiss Economic Institute, ETH Zurich, WEH D4, Weinbergstrasse 35, 8092 Zurich, Switzerland; email: elmer@kof.ethz.ch.

§Menelaos Karanasos, Economics and Finance, Brunel University, Uxbridge, West London, UB3 3PH, UK; email: menelaos.karanasos@brunel.ac.uk, tel: +44 (0)1895 265 284, fax: +44 (0)1895 269 770.

1 Introduction

Many economic time series are characterized by an autocorrelation structure which makes it difficult to classify the series as being either stationary I(0) or non-stationary I(1). A primary example for such a series are inflation rates. Conventional wisdom then suggests to employ unit root tests in order to base the econometric analysis either on the level of such a series or on the first difference. Clearly, the decision whether the series is treated as being I(0) or I(1) has important implications for the subsequent modeling, hypothesis testing, forecasting and the like. Besides the observation that many economic time series are strongly dependent over time, there is the stylized fact that for the same series typically GARCH effects with highly persistent volatility are found. Moreover, economic theory often suggests that the level and the second conditional moment of these series should be interrelated. For example, Cukierman and Meltzer (1986) and Holland (1995) argue that inflation uncertainty has either a positive or a negative effect on the level of inflation, while Friedman (1977) and Ball (1992) rationalize an effect of the level of inflation on its second conditional moment. Against this theoretical background the phenomena of persistence in the level and in the conditional variance are usually analyzed and treated independently. For example, standard unit root tests are based either on the assumption that the variance of the series is constant or that some type of heteroscedasticity is at place, but ignore the possibility that the volatility has a direct effect on the level.

In this paper we consider an AR(1)-GARCH(1,1)-in-mean-level process, i.e. a model in which the conditional variance affects the level of the dependent variable and vice versa. This model has been introduced by Engle et al. (1987) and applied in, e.g., Grier and Perry (2000). We provide a new interpretation of the model's properties by arguing that it has an observationally equivalent representation as an ARMA(2,1) process. The highest autoregressive root of the AR part will, under reasonable assumptions, be close to and statistically indistinguishably from one. This means that in empirical applications the process will appear to be an ARIMA(1,1,1). Most importantly, the largest root of the AR part is closely linked to the persistence of the conditional variance of the process. Using a Monte Carlo study, we show that in the presence of volatility spillovers conventional unit root tests fail to reject the null that the underlying process is I(1). In the presence of volatility spillovers, the persistence of the conditional variance is transmitted to the level of the process and procedures which do not distinguish between the different degrees of persistence in the two moments tend to overestimate the persistence either in the mean or variance. We illustrate this important point by deriving the autocorrelation function, the impulse response function and the optimal predictor for the level process.

An empirical application to U.S. inflation data shows that the model accurately explains the changes in inflation persistence in the mid-1980's and the accompanying decrease in volatility which is commonly referred to as the "Great Moderation". Our main result is the observation that the persistence in the level and the conditional variance of inflation are directly linked and a meaningful analysis has to take into account both properties jointly. Our findings concur with Stock and Watson (2007) who suggest that the U.S. rate of price inflation can reasonably well be approximated by an IMA(1,1) process with a change in both the MA(1) parameter and the variance of the error term in the mid-1980's. However, we show that the increase in the MA(1) parameter and the decrease in the error variance are ultimately linked and driven by the same sources: a decrease in inflation persistence and a change in the sign of the effect from inflation uncertainty on the level of inflation. Both can be attributed to a change in U.S. monetary policy in the post 1984 era and is line with the predictions made in Clarida and Waldman (2007).

The outline of the paper is as follows. Section 2 presents the model and its properties, including the autocorrelation structure, measures of persistence and optimal predictors. In Section 3 the model is applied to U.S. inflation data. Finally, Section 4 concludes. All proofs are deferred to the appendix.

2 The Model

The AR(1)-GARCH(1,1)-in-mean-level model is given by

$$(1-\phi L)y_t = \varphi + \vartheta h_t^{\frac{\delta}{2}} + \varepsilon_t, \tag{1}$$

with

$$\varepsilon_t = e_t h_t^{\frac{1}{2}},$$

where $\delta > 0$, $\{e_t\}$ is a sequence of independent, identically distributed random variables with zero mean and unit variance and h_t is the conditional variance of y_t. The power transformed conditional variance, $h_t^{\frac{\delta}{2}}$, is positive with probability one and is a measurable function of $\mathcal{F}_{t-1}$, which in turn is the sigma-algebra generated by $\{y_{t-1}, y_{t-2}, \ldots\}$. We assume that h_t is specified as an APARCH$(1,1)$-level (L) process:

$$(1-\beta L)h_t^{\frac{\delta}{2}} = \omega + \alpha f(\varepsilon_{t-1}) + \gamma y_{t-1}, \tag{2}$$

with

$$f(\varepsilon_{t-1}) = f(\varepsilon_t) = [|\varepsilon_t| - \varsigma\varepsilon_t]^{\delta},$$

where $|\varsigma| < 1$. By including lagged y_t in the conditional variance equation and $h_t^{\frac{\delta}{2}}$ in the mean equation, we allow for simultaneous feedback between the two variables. The following conditions are necessary and sufficient for $h_t > 0$, for all t: $\omega > 0$, $\alpha, \beta, \gamma \geq 0$ and $y_t \geq 0$ for all t. Hereafter, we will denote $\mathbb{E}(h_t^{\frac{\delta r}{2}}) = \mu_r$. Notice that when $\delta \neq 0.5, 1, 2$ then both $\mu_{2/\delta}$ and $\mu_{1+1/\delta}$ are fractional moments and have to be calculated numerically. In addition, μ_1 and μ_2 are given below (see proposition 2 below and equation (21) in the Appendix).

The APARCH(1,1)-L formulation in equation (2) can readily be interpreted as an ARMA(1,1)-L process for the conditional variance:

$$(1-cL)h_t^{\frac{\delta}{2}} = \omega + \alpha v_{t-1} + \gamma y_{t-1}, \tag{3}$$

where $c = \alpha\kappa^{(1)} + \beta$, with $\kappa^{(r)} = \mathbb{E}\{[f(e_t)]^r\}$, and $v_t = f(\varepsilon_t) - \kappa^{(1)} h_t^{\frac{\delta}{2}}$ is an uncorrelated term with expected value 0 and $\mathbb{E}(v_t^2) = \sigma_v^2 = \mu_2\widetilde{\kappa}$ with $\widetilde{\kappa} = [\kappa^{(2)} - (\kappa^{(1)})^2]$. Notice also that $\mathbb{E}(\varepsilon_t^2) = \sigma_\varepsilon^2 = \mu_{2/\delta}\mathbb{E}(e_t^2)$ and $\mathbb{E}(\varepsilon_t v_t) = \sigma_{\varepsilon v} = \mu_{1+1/\delta}\overline{\kappa}$ with $\overline{\kappa} = \mathbb{E}[e_t f(e_t)]$. In other words, the covariance matrix of the two 'shocks' ε_t,v_t is given by

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon v} \\ \sigma_{\varepsilon v} & \sigma_v^2 \end{bmatrix} = \begin{bmatrix} \mu_{2/\delta}\mathbb{E}(e_t^2) & \mu_{1+1/\delta}\overline{\kappa} \\ \mu_{1+1/\delta}\overline{\kappa} & \mu_2\widetilde{\kappa} \end{bmatrix}. \tag{4}$$

Remark 1 *Consider the case where e_t is standard normal. Then $\mathbb{E}(e_t^2) = 1$, and $k^{(r)}$, $\overline{\kappa}$ are given by*

$$\begin{aligned}\kappa^{(r)} &= \frac{1}{\sqrt{\pi}}[(1-\varsigma)^{r\delta} + (1+\varsigma)^{r\delta}]2^{(\frac{r\delta}{2}-1)}\Gamma(\frac{r\delta+1}{2}),\\ \overline{\kappa} &= ((1-\varsigma)^{\delta} - (1+\varsigma)^{\delta})2^{(\delta/2)}\Gamma(\frac{\delta}{2}+1).\end{aligned}$$

If in addition $\delta = 2$, $\gamma = 0$ and $\varsigma = 0$ then Σ reduces to

$$\mathbf{\Sigma} = \begin{bmatrix} \mu_1 & 0 \\ 0 & 2\mu_2 \end{bmatrix},$$

where μ_2 is given by equation (23) in the Appendix.

Note, that the parameter c measures the *memory* or *persistence* in the conditional variance. The extreme case in which $c = 1$ is well known as the integrated GARCH model (see Engle and Bollerslev, 1986).

The APARCH(1,1) specification can be expressed as an ARMA(2,1) process. In the model given by expressions (1)-(2), although the conditional variance follows a APARCH(1,1) formulation, due to the simultaneous feedback, it has a univariate ARMA(2,1) representation.

Proposition 1 *The univariate ARMA(2,1) representation of the process y_t and the power transformed conditional variance $h_t^{\frac{\delta}{2}}$ are given by*

$$\begin{aligned}(1-a_1L-a_2L^2)y_t &= \varphi^* + (1-cL)\varepsilon_t + \vartheta\alpha v_{t-1}, &(5)\\ (1-a_1L-a_2L^2)h_t^{\frac{\delta}{2}} &= \omega^* + \gamma\varepsilon_{t-1} + (1-\phi L)\alpha v_{t-1}, &(6)\end{aligned}$$

where $a_1 = \phi + c + \vartheta\gamma$, $a_2 = -\phi c$, $\varphi^* = \varphi(1-c) + \vartheta\omega$ and $\omega^* = \omega(1-\phi) + \varphi\gamma$.

Equation (5) shows that the process y_t is driven by two shocks, namely shocks to the mean ε_t and shocks to the conditional variance v_t. The two shocks will be uncorrelated if $\varsigma = 0$. Equation (5) can be rewritten as

$$(1-a_1L-a_2L^2)y_t = \varphi^* + (1-\theta L)\eta_t$$

where η_t is an uncorrelated error process with mean zero and variance σ_η^2. The parameters $\theta$and σ_η^2 can expressed as

$$\begin{aligned}\sigma_\eta^2 &= \frac{\sigma_1}{-\theta},\\ \theta &= \frac{-\sigma_0 \pm \sqrt{\sigma_0^2 - 4\sigma_1^2}}{2\sigma_1},\end{aligned}$$

where $\sigma_0 = (1+c^2)\sigma_\varepsilon^2 + (\vartheta\alpha)^2\sigma_v^2 - 2\vartheta\alpha\sigma_{\varepsilon v}$, and $\sigma_1 = -c\sigma_\varepsilon^2 + \vartheta\alpha\sigma_{\varepsilon v}$. Notica that (i) θ is real if and only if $\sigma_0^2 > 4\sigma_1^2$, and (ii) when $\sigma_1 \lesseqgtr 0$, that is $\vartheta\alpha\sigma_{\varepsilon v} \lesseqgtr c\sigma_\varepsilon^2$, then $0 \lesseqgtr \theta$.

Assumption A1 (Stationarity) *The inverse roots λ_1 and λ_2 of $(1 - a_1 L - a_2 L^2)$ lie inside the unit circle. Moreover, assume that $\lambda_1 \neq \lambda_2$.*

Assumption (A1) implies that the ARMA(2,1) process given by equation (5) is covariance stationary. Note that it also implies that $a_1 + a_2 < 1$.

Remark 2 *For illustrative purposes consider the case that $\vartheta\gamma = 0$. In this situation, the inverse roots are given by $\lambda_1 = \phi$ and $\lambda_2 = c$. In many empirical applications c is found to be close to and statistically indistinguishably from one (see, e.g., Engle and Bollerslev, 1986). Hence, this example suggests that the model given by equations (1)-(2) leads to observations y_t which empirically may be easily confused with a process that is integrated of order one, in particular the ARIMA(1,1,1) given by:*

$$(1 - \phi L)(1 - L) y_t = \varphi^* + (1 - \theta L)\eta_t$$

Proposition 2 *When Assumption (A1) holds the unconditional expectation of $h_t^{\frac{\delta}{2}}$ exists if $\omega^* > 0$, and it is given by*

$$\mu_1 = \frac{\omega^*}{1 - a_1 - a_2}. \tag{7}$$

Note that the existence of μ_1 guarantees that of $\mu_{2/\delta}$ only if $\delta \geq 1$. Similarly the existence of the second moment μ_2 (see below) guarantees that of $\mu_{1+1/\delta}$ only if $\delta \geq 1$.

Proposition 3 *Let Assumption (A1) hold. Then, equations (5) and (6) admit the Wold representation*

$$\begin{aligned} y_t &= y^* + \psi_{y\varepsilon}(L)\varepsilon_t + \psi_{yv}(L)v_t, & (8) \\ h_t^{\frac{\delta}{2}} &= \mu_1 + \psi_{h\varepsilon}(L)\varepsilon_t + \psi_{hv}(L)v_t, & (9) \end{aligned}$$

where $y^* = \varphi^*/(1 - a_1 - a_2)$, and $\psi_{ij}(L) = \sum_{k=0}^{\infty} \psi_{ij}^{(k)} L^k$, $i = y, h$; $j = \varepsilon, v$ and

$$\begin{aligned}
\psi_{y\varepsilon}^{(0)} &= 1, \; \psi_{y\varepsilon}^{(k)} = \left[\frac{\lambda_1^k(\lambda_1 - c)}{\lambda_1 - \lambda_2} + \frac{\lambda_2^k(\lambda_2 - c)}{\lambda_2 - \lambda_1} \right], \; k \geq 1, \\
\psi_{yv}^{(0)} &= 0, \; \psi_{yv}^{(k)} = \vartheta\alpha \left(\frac{\lambda_1^k}{\lambda_1 - \lambda_2} + \frac{\lambda_2^k}{\lambda_2 - \lambda_1} \right), \; k \geq 1, \\
\psi_{h\varepsilon}^{(0)} &= 0, \; \psi_{h\varepsilon}^{(k)} = \gamma \left(\frac{\lambda_1^k}{\lambda_1 - \lambda_2} + \frac{\lambda_2^k}{\lambda_2 - \lambda_1} \right), \; k \geq 1, \\
\psi_{hv}^{(0)} &= 0, \; \psi_{hv}^{(1)} = \alpha, \; \psi_{hv}^{(k)} = \alpha \left[\frac{\lambda_1^{k-1}(\lambda_1 - \phi)}{\lambda_1 - \lambda_2} + \frac{\lambda_2^{k-1}(\lambda_2 - \phi)}{\lambda_2 - \lambda_1} \right], \; k \geq 2.
\end{aligned}$$

Remark 3 *If $\sigma_{\varepsilon v} = 0$ then $\psi_{y\varepsilon}^{(k)}$ and $\psi_{yv}^{(k)}$ are the 'impulse response functions' of a one unit mean shock ε_t or variance shock v_t to the process y_t. The expression for $\psi_{yv}^{(k)}$ shows that a positive shock to the conditional variance can either imply a positive or a negative cumulative response, depending on the sign of ϑ. To the contrary, a positive shock to the mean always implies a positive cumulative response. We now consider the situation where*

$\vartheta\gamma = 0$ and, hence, $\lambda_1 = \phi$ and $\lambda_2 = c$. For this specific example, the above formulas reduce to:

$$\begin{aligned}\psi_{y\varepsilon}^{(k)} &= \phi^k \\ \psi_{yv}^{(k)} &= \vartheta\alpha\left(\frac{\phi^k - c^k}{\phi - c}\right)\end{aligned}$$

for $k \geq 1$. This nicely illustrates that the speed with which a mean or conditional variance shock decays to zero can be different provided that $\phi < c$. The case where $\sigma_{\varepsilon v} \neq 0$ will be treated in Section 2.2.

In the following subsections we will discuss the covariance structure and the optimal predictors of y_t and h_t.

2.1 Covariance structure

In order to simplify the description of our analysis we will introduce the following notation:

$$\begin{aligned}\lambda &= \frac{1}{(1-\lambda_1\lambda_2)(\lambda_1-\lambda_2)}, \\ \lambda^{(k)} &= \lambda\left[\frac{\lambda_1^{1+k}}{1-\lambda_1^2} - \frac{\lambda_2^{1+k}}{1-\lambda_2^2}\right], \\ \lambda_g^{(k)} &= \left\{\begin{array}{lll}\lambda\left[\frac{\lambda_1(1+g^2-2g\lambda_1)}{1-\lambda_1^2} - \frac{\lambda_2(1+g^2-2g\lambda_2)}{1-\lambda_2^2}\right] & \text{if} & k=0, \\ \lambda\left[\frac{\lambda_1^k(\lambda_1-g)(1-\lambda_1 g)}{1-\lambda_1^2} - \frac{\lambda_2^k(\lambda_2-g)(1-\lambda_2 g)}{1-\lambda_2^2}\right] & \text{if} & k \geq 1\end{array}\right., \quad g = \phi, c.\end{aligned}$$

Note that $\lambda^{(0)} = \frac{(1+\lambda_1\lambda_2)}{(1-\lambda_1\lambda_2)(1-\lambda_1^2)(1-\lambda_2^2)}$ and $\lambda^{(1)} = \frac{(\lambda_1+\lambda_2)}{(1-\lambda_1^2)(1-\lambda_2^2)}$.
Further, for $k \geq 0$ and $g = \phi, c$, we define

$$\begin{aligned}\widetilde{\lambda}_g^{(k)} &= \lambda\left[\frac{\lambda_1^{1+k}(\lambda_1-g)}{1-\lambda_1^2} - \frac{\lambda_2^{1+k}(\lambda_2-g)}{1-\lambda_2^2}\right], \\ \overline{\lambda}_g^{(k)} &= \lambda\left[\frac{\lambda_1^{1+k}(1-g\lambda_1)}{1-\lambda_1^2} - \frac{\lambda_2^{1+k}(1-g\lambda_2)}{1-\lambda_2^2}\right].\end{aligned}$$

Interestingly, if $\lambda_1 = \phi$ and $\lambda_2 = c$, then the above expressions reduce to

$$\lambda_g^{(k)} = \frac{(\phi c/g)^k}{1-(\phi c/g)^2}, \quad \widetilde{\lambda}_g^{(k)} = \frac{(\phi c/g)^{1+k}}{[1-(\phi c/g)^2](1-\phi c)}, \; \overline{\lambda}_g^{(k)} = \frac{1}{g - \frac{\phi c}{g}}\left[\frac{g^{1+k}}{1-\phi c} - \frac{(\frac{\phi c}{g})^{1+k}}{1-(\frac{\phi c}{g})^2}\right].$$

Proposition 4 *Assume that $\mathbf{\Sigma}$ exists. That is $0 < \mathbf{\Sigma} < \infty$.*[1] *Then the autocovariances of y_t, $\mathbb{C}ov_k(y_t)$, $k \in \mathbb{N}$, are given by*

$$\mathbb{V}(y_t) = \sigma_\varepsilon^2\lambda_c^{(0)} + \sigma_v^2(\vartheta\alpha)^2\lambda^{(0)} + 2\sigma_{\varepsilon v}\vartheta\alpha\widetilde{\lambda}_c^{(0)}, \tag{10}$$

$$\mathbb{C}ov_k(y_t) = \sigma_\varepsilon^2\lambda_c^{(k)} + \sigma_v^2(\vartheta\alpha)^2\lambda^{(k)} + \sigma_{\varepsilon v}\vartheta\alpha(\widetilde{\lambda}_c^{(k)} + \overline{\lambda}_c^{(k-1)}), \; k \geq 1. \tag{11}$$

where all the λ's have been defined above.

[1] Matrix inequality sign $\mathbf{\Sigma} < \infty$ represents element-by-element inequality.

Lemma 1 *Let $\lambda_1 = \phi$ and $\lambda_2 = c$. In this case we have*

$$
\begin{aligned}
\mathbb{V}(y_t) &= \frac{1}{1-\phi^2}\left\{\sigma_\varepsilon^2 + \frac{\vartheta\alpha}{1-\phi c}\left[\frac{\sigma_v^2\vartheta\alpha(1+\phi c)}{1-c^2} + 2\sigma_{\varepsilon v}\phi\right]\right\}, \\
\mathbb{C}ov_k(y_t) &= \frac{\sigma_\varepsilon^2\phi^k}{1-\phi^2} + \sigma_v^2(\vartheta\alpha)^2\lambda^{(k)} + \frac{\sigma_{\varepsilon v}\vartheta\alpha}{(\phi-c)(1-\phi c)}\left[\frac{\phi^k(1+\phi^2-2\phi c)}{1-\phi^2} - c^k\right],\ k \geq 1.
\end{aligned}
$$

Obviously when $\vartheta = 0$, $\mathbb{C}ov_k(y_t) = \frac{\sigma_\varepsilon^2\phi^k}{1-\phi^2}$, $k \geq 0$, which are the autocovariances of an AR(1)-GARCH(1,1) process.

For the specific case that $\delta = 2$, $\gamma = 0$ and $\varsigma = 0$ we will now graphically illustrate the behavior of the autocorrelation function given by $\rho_k(y_t) = \mathbb{C}ov_k(y_t)/\mathbb{V}(y_t)$ for $k = 1, 2, \ldots$. Figures 1, 3 and 5 show the autocorrelations of y_t for $\alpha = 0.1$, $\beta \in \{0.8, 0.85, 0.9\}$, $\vartheta \in \{0, 0.5, 1, 2\}$ and $\phi \in \{0.1, 0.5, 0.9\}$. Figure 1 clearly shows that even for a very low value of ϕ the autocorrelation function of y_t can appear quite persistent if the conditional variance process is sufficiently persistent, i.e. $c = \alpha + \beta$ is close to one, and there is an in-mean effect, i.e. $\vartheta \neq 0$. The cases $\phi = 0.5$ and $\phi = 0.9$ in Figures 3/5 illustrate that for most of the combinations of β and $\vartheta(\neq 0)$ the almost linearly decaying autocorrelations would suggest that y_t is integrated of order one.

The second order partial autocorrelation of y_t, $\rho_2^*(y_t)$, is given by

$$
\rho_2^*(y_t) = \frac{\rho_2(y_t) - \rho_1^2(y_t)}{1-\rho_1^2(y_t)},
$$

where $\rho_k(y_t)$is the kth order autocorrelation of y_t.

In order to obtain some further insights into the persistence properties of y_t, we generated series y_t, $t = 1, \ldots, T$, with $T = 500$ or $T = 1000$ according to our AR(1)-GARCH(1,1)-in-mean model. For simplicity, we set $\delta = 2$, $\gamma = 0$ and $\varsigma = 0$, i.e. we consider a model with no level effect and no asymmetries. The innovation e_t was chosen to be standard normal. We then applied the standard Dickey-Fuller, the Phillips-Perron and the KPSS test to the y_t series. Figures 7 and 8 show the fraction of cases in which the null hypothesis of each test is not rejected. The results are based on 1000 Monte-Carlo replications. The acceptance rate of all three tests clearly depends on the absolute value of the in-mean term ϑ. Both unit root tests tend to accept the null hypothesis of y_t being I(1) with $|\vartheta|$ increasing. The results clearly suggest that unit root tests will have low power against alternatives which allow for in-mean effects in combination with persistent conditional variances. Similarly, the KPSS does no longer accept the null hypothesis of y_t being I(0) with $|\vartheta|$ increasing. That is, the size of the KPSS test is considerably distorted in the presence of in-mean effects in combination with persistent conditional variances.

In summary, our results on the covariance structure of y_t suggest that a process with persistent conditional variance and in-mean effect may easily be confused with a process that is integrated of order one in the level.

2.2 Measures of Persistence

The above considerations suggest that conventional measures of persistence might result in misleading conclusions regarding the persistence of the process y_t. The most often applied measures are (a) the largest autoregressive root (LAR), which we denote by $\lambda^* = \max(\lambda_1, \lambda_2)$ and (b) the sum of the coefficients (SUM) in the autoregressive progress, that is $a_1 + a_2$ (see, e.g., Pivetta and Reis, 2007).[2] Obviously, both measures would ignore the presence of the in-mean effect and, hence, 'overestimate' the persistence in the mean which is partly induced by the persistence in the conditional variance.

We follow Fiorentini and Sentana (1999) who argue that any reasonable measure of persistence of shocks must be based on the impulse response function (IRF). In the following we will derive the IRF for the general case where ε_t and v_t are correlated.

Let $\mathbf{y}_t = \mathbf{\Psi}(L)\varepsilon_t$ denote the Wold representation of the vector process $\mathbf{y}_t = (y_t, h_t)'$, where $\mathbf{\Psi}(L) = [\psi_{ij}(L)]_{i=y,h;j=\varepsilon,v}$, and $\varepsilon_t = (\varepsilon_t, v_t)'$ with covariance matrix $\mathbf{\Sigma}$ (see equation (4)).

Next denote $P_\infty(y_t|\varepsilon_t) = \sum_{k=0}^{\infty}\left(\psi_{y\varepsilon}^{(k)}\right)^2$, $P_\infty(y_t|v_t) = \sum_{k=0}^{\infty}\left(\psi_{yv}^{(k)}\right)^2$ and $P_\infty(y_t|\sqrt{\varepsilon_t v_t}) = \sum_{k=0}^{\infty}\psi_{y\varepsilon}^{(k)}\psi_{yv}^{(k)}$. Define the lower triangular matrix $\widetilde{\mathbf{\Sigma}}$ with diagonal elements $\widetilde{\sigma}_\varepsilon = \sigma_\varepsilon$, $\widetilde{\sigma}_v = \sigma_v\sqrt{1-\rho_{\varepsilon v}}$, and off diagonal elemen $\sqrt{\widetilde{\sigma}_{\varepsilon v}} = \sigma_v \rho_{\varepsilon v}$ such that $\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{\Sigma}}' = \mathbf{\Sigma}$. That is, $\sigma_\varepsilon^2 = \widetilde{\sigma}_\varepsilon^2$, $\sigma_v^2 = \widetilde{\sigma}_v^2 + \widetilde{\sigma}_{\varepsilon v}$, $j = \varepsilon, v$, and $\sigma_{\varepsilon v} = \widetilde{\sigma}_\varepsilon\sqrt{\widetilde{\sigma}_{\varepsilon v}}$. Then, the infinite moving average representation of $\mathbf{y}_t$ in terms of the standardized orthogonal innovations $\widetilde{\varepsilon}_t = \widetilde{\mathbf{\Sigma}}^{-1}\varepsilon_t$ is $\mathbf{y}_t = \widetilde{\mathbf{\Psi}}(L)\widetilde{\varepsilon}_t$ where $\widetilde{\mathbf{\Psi}}(L) = \mathbf{\Psi}(L)\widetilde{\mathbf{\Sigma}}$ and the covariance matrix of $\widetilde{\varepsilon}_t$ is the identity matrix (see Fiorentini and Sentana, 1999).

We can then define the persistence of a shock to $\widetilde{\varepsilon}_t = \frac{\widetilde{\sigma}_v\varepsilon_t - \sqrt{\widetilde{\sigma}_{\varepsilon v}}v_t}{\widetilde{\sigma}_\varepsilon\widetilde{\sigma}_v - \widetilde{\sigma}_{\varepsilon v}}$ and $\widetilde{v}_t = \frac{\widetilde{\sigma}_\varepsilon v_t - \sqrt{\widetilde{\sigma}_{\varepsilon v}}\varepsilon_t}{\widetilde{\sigma}_\varepsilon\widetilde{\sigma}_v - \widetilde{\sigma}_{\varepsilon v}}$ on the y_t variable as

$$P_\infty(y_t|\widetilde{\varepsilon}_t) = \sum_{k=0}^{\infty}\left(\widetilde{\psi}_{y\varepsilon}^{(k)}\right)^2 = \widetilde{\sigma}_\varepsilon^2 P_\infty(y_t|\varepsilon_t) + \widetilde{\sigma}_{\varepsilon v}P_\infty(y_t|v_t) + 2\widetilde{\sigma}_\varepsilon\sqrt{\widetilde{\sigma}_{\varepsilon v}}P_\infty(y_t|\sqrt{\varepsilon_t v_t}), \tag{12}$$

$$P_\infty(y_t|\widetilde{v}_t) = \sum_{k=0}^{\infty}\left(\widetilde{\psi}_{yv}^{(k)}\right)^2 = \widetilde{\sigma}_v^2 P_\infty(y_t|v_t) \tag{13}$$

The algebra of this measure is simple and its interpretation straightforward since

$$P_\infty(y_t|\widetilde{\varepsilon}_t) = \mathbb{V}(y_t) - P_\infty(y_t|\widetilde{v}_t),$$

that is, the part of the variance of y_t due to $\widetilde{\varepsilon}_t$.

Proposition 5 *If $0 < \mathbf{\Sigma} < \infty$ then $P_\infty(y_t|\widetilde{\varepsilon}_t)$ and $P_\infty(y_t|\widetilde{v}_t)$ are given by equations (12) and (13) respectively where*

$$P_\infty(y_t|\varepsilon_t) = 1 + \frac{1}{(\lambda_1-\lambda_2)^2}\left[\frac{\lambda_1^2(\lambda_1-c)^2}{1-\lambda_1^2} + \frac{\lambda_2^2(\lambda_2-c)^2}{1-\lambda_2^2} - \frac{2\lambda_1\lambda_2(\lambda_1-c)(\lambda_2-c)}{1-\lambda_1\lambda_2}\right], \tag{14}$$

$$P_\infty(y_t|v_t) = \frac{(\vartheta\alpha)^2(1+\lambda_1\lambda_2)}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}, \tag{15}$$

$$P_\infty(y_t|\sqrt{\varepsilon_t v_t}) = \frac{\vartheta\alpha[\lambda_2 - c + \lambda_1(1-c\lambda_2)]}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}. \tag{16}$$

[2]Note that if $\vartheta\gamma > 0$ then $a_1 + a_2 = \phi + c(1-\phi) + \vartheta\gamma > \phi + c(1-\phi) > \phi$,$c$.

Lemma 2 *Interestingly, if $\lambda_1 = \phi$ and $\lambda_2 = c$, then*

$$\begin{aligned} P_\infty(y_t|\varepsilon_t) &= \frac{1}{1-\phi^2},\ P_\infty(y_t|v_t) = \frac{(\vartheta\alpha)^2(1+\phi c)}{(1-\phi^2)(1-c^2)(1-\phi c)}, \\ P_\infty(y_t|\sqrt{\varepsilon_t v_t}) &= \frac{\vartheta\alpha\phi}{(1-\phi c)(1-\phi^2)}, \end{aligned}$$

Obviously, in this case the expression for $P_\infty(y_t|\varepsilon_t)$ is the one for the AR(1) process (see Fiorentini and Sentana, 1999). Further, notice that if $\vartheta = 0$ then $P_\infty(y_t|v_t) = P_\infty(y_t|\sqrt{\varepsilon_t v_t}) = 0$. Finally, if $\phi = 0$ then

$$P_\infty(y_t|\varepsilon_t) = 1,\ P_\infty(y_t|v_t) = \frac{(\vartheta\alpha)^2}{1-c^2},\ P_\infty(y_t|\sqrt{\varepsilon_t v_t}) = 0.$$

2.3 Optimal predictors

In this section we provide a formula for the n-step ahead predictor $\mathbb{E}_t(y_{t+n})$ of y_t given the information available at time t.

Proposition 6 *Under Assumption (A1) the n-step ahead predictor of y_{t+n} is readily seen to be*

$$\begin{aligned} \mathbb{E}_t(y_{t+n}) &= \frac{1}{\lambda_1-\lambda_2}\{\widetilde{\phi} + [(\lambda_1^{n+1}-\lambda_2^{n+1})y_t - (\lambda_1^{n+1}\lambda_2 - \lambda_2^{n+1}\lambda_1)y_{t-1}] + \\ &\quad (\lambda_1^n - \lambda_2^n)(\vartheta\alpha v_t - c\varepsilon_t)\}, \end{aligned}$$

where

$$\widetilde{\phi} = \rho^* \frac{\lambda_1[1-\lambda_1^n(1-\lambda_2)] + \lambda_2[1-\lambda_2^n(1-\lambda_1)]}{(1-\lambda_1)(1-\lambda_2)}.$$

The associated forecast error is given by

$$\begin{aligned} y_{t+n} - \mathbb{E}_t(y_{t+n}) &= \vartheta\alpha \sum_{l=0}^{n-2} \frac{1}{\lambda_1-\lambda_2}\left[\lambda_1^{l+1}L^l - \lambda_2^{l+1}L^l\right] v_{t+n-1} + \left\{1 + \frac{1}{\lambda_1-\lambda_2}\right. \\ &\quad \left.\left[\lambda_1(\lambda_1-c)\sum_{l=1}^{n-1}\lambda_1^{l-1}L^l - \lambda_2(\lambda_2-c)\sum_{l=1}^{n-1}\lambda_2^{l-1}L^l\right]\right\}\varepsilon_{t+n}. \end{aligned}$$

If $0 < \mathbf{\Sigma} < \infty$, then the variance of the forecast error is given by

$$\begin{aligned} \mathbb{V}[y_{t+n} - \mathbb{E}_t(y_{t+n})] &= P_n(y_t|\varepsilon_t)\sigma_\varepsilon^2 + P_n(y_t|v_t)\sigma_v^2 + 2P_n(y_t|\sqrt{\varepsilon_t v_t})\sigma_{\varepsilon v} = \\ &\quad \widetilde{\sigma}_\varepsilon^2 P_n(y_t|\varepsilon_t) + \widetilde{\sigma}_{\varepsilon v}P_n(y_t|v_t) + 2\widetilde{\sigma}_\varepsilon\sqrt{\widetilde{\sigma}_{\varepsilon v}}P_n(y_t|\sqrt{\varepsilon_t v_t}) + \\ &\quad \widetilde{\sigma}_{\varepsilon v}P_n(y_t|\varepsilon_t) + \widetilde{\sigma}_v^2 P_n(y_t|v_t) + 2\widetilde{\sigma}_v\sqrt{\widetilde{\sigma}_{\varepsilon v}}P_n(y_t|\sqrt{\varepsilon_t v_t}) \\ &= P_n(y_t|\widetilde{\varepsilon}_t) + P_n(y_t|\widetilde{v}_t), \end{aligned}$$

where

$$
\begin{aligned}
P_n(y_t|\varepsilon_t) &= P_\infty(y_t|\varepsilon_t) - \frac{1}{(\lambda_1-\lambda_2)^2}\left[\frac{\lambda_1^{2(n+1)}(\lambda_1-c)^2}{1-\lambda_1^2} - \right.\\
&\qquad \left.\frac{\lambda_2^{2(n+1)}(\lambda_2-c)^2}{1-\lambda_2^2} - \frac{2(\lambda_1\lambda_2)^{n+1}(\lambda_1-c)(\lambda_2-c)}{1-\lambda_1\lambda_2}\right],\\
P_n(y_t|v_t) &= P_\infty(y_t|v_t) - \frac{(\vartheta\alpha)^2}{(\lambda_1-\lambda_2)^2}\left[\frac{\lambda_1^{2n}}{1-\lambda_1^2} - \frac{\lambda_2^{2n}}{1-\lambda_2^2} - \frac{2\lambda_1^n\lambda_2^n}{1-\lambda_1\lambda_2}\right],\\
P_n(y_t|\sqrt{\varepsilon_t v_t}) &= P_\infty(y_t|\sqrt{\varepsilon_t v_t}) - \frac{\vartheta\alpha}{(\lambda_1-\lambda_2)(1-\lambda_1\lambda_2)}\left[\frac{\lambda_1^{n+1}(\lambda_1-c)}{1-\lambda_1^2} - \frac{\lambda_2^{n+1}(\lambda_2-c)}{1-\lambda_2^2}\right].
\end{aligned}
$$

Obviously, for a covariance stationary processes $\lim_{n\to\infty} P_n(y_t|\widetilde{\varepsilon}_t) = P_\infty(y_t|\widetilde{\varepsilon}_t)$ and $\lim_{n\to\infty} P_n(y_t|\widetilde{v}_t) = P_\infty(y_t|\widetilde{v}_t)$. Sometimes it is more interesting to look at the effect of a shock on a variable n periods after its occurrence (see Fiorentini and Sentana, 1998, and the references therein). For this purpose, $P_n(y_t|\widetilde{\varepsilon}_t) = \sum_{k=0}^{n}\left(\widetilde{\psi}_k^{(y\varepsilon)}\right)^2$ can be used as a measure of the interim persistence of the shock ε_t. Unlike the $P_\infty(y_t|\widetilde{\varepsilon}_t)$, the n-period measure $P_n(y_t|\widetilde{\varepsilon}_t)$ can be used and interpreted for nonstationary processes as well (Fiorentini and Sentana, 1998).

3 Empirical Application

In this section we apply our model to U.S. inflation data. Seasonally adjusted monthly consumer price index data was obtained from the Federal Reserve Bank of St. Louis for the period 01/1947 - 12/2008. The CPI inflation rate was calculated as $\pi_t = 1200 \times [\ln(P_t) - \ln(P_{t-1})]$, where P_t denotes the price level in month t.

Table 1: Summary Statistics, CPI Inflation.

	01/1947 - 12/1983		01/1984 - 12/2008	
Standard deviation of π_t	4.81		3.19	
	ACF	PACF	ACF	PACF
of $\Delta\pi_t$ at lag				
1	-0.382	-0.382	-0.187	-0.187
2	-0.087	-0.273	-0.268	-0.314
3	0.020	-0.160	-0.006	-0.150
4	-0.052	-0.172	0.009	-0.135
5	0.018	-0.121	-0.115	-0.229

Table 2 presents the estimation results for the full sample and the two subsamples. For the full sample, the AR(1)-AGARCH(1,1)-M implies a positive and significant in-mean term, i.e. more inflation uncertainty leads to higher rates of inflation. The AR(1) coefficient is 0.47 and, as expected, highly significant. The asymmetry term ς is found to be negative and significant, suggesting that positive inflation shocks create more uncertainty than negative ones of the same size. The parameter values for α and β are rather typical for inflation data, and the persistence parameter $\lambda_2 = c = 0.96$ implies strong persistence in the conditional variance but still ensures the existence of the unconditional second moment of y_t. For the estimated ARMA(2,1) model the AR(1) coefficient is larger than one and the AR(2) coefficient is negative. Both observations are in line with our AR(1)-GARCH(1,1)-M specification. Also the two estimated roots λ_1 and λ_2 from the ARMA(2,1) model are close to their counterparts from the AR(1)-GARCH(1,1)-M.

Table 2: Model Estimates.

	01/1947 - 12/2008		01/1947 - 12/1984		01/1984 - 12/2008	
	AR(1)-GARCH(1,1)-M	ARMA(2,1)	AR(1)-GARCH(1,1)-M	ARIMA(1,1,1)	AR(1)-GARCH(1,1)-M	ARIMA(1,1,1)
	Mean Equation					
$const.$	2.97 (0.20)	3.26 (0.72)	2.92 (0.42)	−0.01 (0.03)	3.25 (0.24)	−0.01 (0.003)
AR(1)	0.47 (0.04)	1.23 (0.10)	0.43 (0.07)	0.18 (0.10)	0.34 (0.07)	0.32 (0.08)
AR(2)		−0.26 (0.09)				
MA(1)		−0.82 (0.07)		−0.80 (0.06)		−0.99 (0.01)
ϑ	0.05 (0.02)		0.11 (0.03)		−0.06 (0.03)	
	Variance Equation					
ω	0.27 (0.14)		0.31 (0.14)		0.60 (0.30)	
α	0.12 (0.05)		0.06 (0.04)		0.27 (0.10)	
ς	−0.17 (0.09)		−0.55 (0.32)		−0.15 (0.13)	
β	0.84 (0.05)		0.88 (0.04)		0.67 (0.10)	
AIC	5.014	5.251	5.245	5.415	4.630	4.328
SIC	5.058	5.276	5.311	5.442	4.717	4.370
λ_1	0.47	0.96	0.44	1.00	0.34	1.00
λ_2	0.96	0.27	0.96	0.18	0.95	0.32

Notes: The numbers in parenthesis are Bollerslev-Wooldrige robust standard errors.

The results from the first subsample are quite similar to the ones from the full sample. However, we make two interesting observations. First, the in-mean term is considerable larger and highly significant, implying that the effect of inflation uncertainty on the level of inflation was stronger in the first subsample. The evidence is in line with . . . . Moreover, the asymmetry term is more than three times bigger in the first subsample than in the full sample.

To the contrary, our estimates for the second subsample suggest that the in-mean was negative from 1984 onwards, reflecting a change in monetary policy. The asymmetry term is no longer significant. Most importantly, the value of β is by far smaller than in the full sample or the first subsample. While in the first subsample the volatility at time t is almost entirely determined by h_{t-1} and only to a small fraction by ε_t, h_{t-1} looses some of its importance in the second subsample while ε_t gains importance.

Considering the corresponding ARIMA(1,1,1) models, the parameter estimates are line with our AR(1)-GARCH(1,1)-M specification. Interestingly, the MA(1) coefficient is in absolute value bigger in the second subsample, an observation also made in Stock and Watson (2007).

4 Conclusions

We discuss the persistence properties of the AR(1)-GARCH(1,1)-in-mean-level model. This model allows for an in-mean effect as well as a level effect. Both effects are in line with economic theory which, e.g., suggests that inflation uncertainty should have an effect on the level of inflation and vice versa. Our main result is that the commonly observed persistence in the mean/conditional variance of many economic times series may be a result of a transmission mechanism. If this mechanism is ignored, conventional procedures for estimating the persistence in the mean/variance may lead to upward biased estimates. In particular, unit root tests will falsely indicate a unit root and, hence, suggest the modeling of the differenced series rather than the level series. Our primary empirical example are inflation rates and we argue that the decrease in both U.S. inflation persistence and inflation uncertainty from the mid 1980's onwards can be well explained by our specification.

References

[1] Christensen, B. J., and M. Ø. Nielsen (2007). "The effect of long memory in volatility on stock market fluctuations." *Review of Economics and Statistics* 89, 684–700.

[2] Clarida, R., and D. Waldman (2007). "Is bad news about inflation good news for the exchange rate?" in J. Y. Campbell (ed.) *Asset Prices and Monetary Policy*, University of Chicago Press, forthcoming.

[3] Conrad, C., and M. Karanasos (2005). "Dual long memory in inflation dynamics across countries of the Euro area and the link between inflation uncertainty and macroeconomic performance." *Studies in Nonlinear Dynamics and Econometrics* 9(4), Article 5.

[4] Cook, S. (2006). "GARCH, heteroscedasticity-consistent covariance matrix estimation and (non)linear unit root testing." *Applied Financial Economics Letters* 2(4), 217–222.

[5] Engle, R. F., D. M. Lilien, and R. P. Robins (1987). "Estimating time varying risk premia in the term structure." *Journal of Business and Economic Statistics* 9, 345–359.

[6] Engle, R. F., and T. Bollerslev (1986). "Modelling the persistence of conditional variances." *Econometric Reviews* 5, 1–50.

[7] Fiorentini, G., and E. Sentana (1998). "Conditional means of time series processes and time series processes for conditional means." *International Economic Review* 39, 1101–1118.

[8] Grier, K., and M. Perry (2000). "The effects of real and nominal uncertainty on inflation and output growth: some GARCH-M evidence." *Journal of Applied Econometrics* 15, 45–58.

[9] Karanasos, M. (2007). "The covariance structure of some financial time series models." *Quantitative and Qualitative Analysis in Social Sciences* 1(2), 71–87.

[10] Karanasos, M. (2001). "Prediction in ARMA models with GARCH in mean effects." *Journal of Time Series Analysis* 22(5), 555–576.

[11] Kim, K., and P. Schmidt (1993). "Unit root tests with conditional heteroskedasticity." *Journal of Econometrics* 59, 287–300.

[12] Kourogenis, N., and N. Pittis (2008). "Testing for a unit root under errors with just barely infinite variance." *Journal of Time Series Analysis* 29(6), 1066–1087.

[13] Pivetta, F., and R. Reis (2007). "The persistence of inflation in the United States." *Journal of Economic Dynamics and Control* 31, 1326–1358.

[14] Rodrigues, P. M. M., and A. Rubia (2005). "The performance of unit root tests under level-dependent heteroskedasticity." *Economics Letters* 89, 262-268.

[15] Stock, J. H., and M. W. Watson (2007). "Why has U.S. inflation become harder to forecast?" *Journal of Money, Credit and Banking* 39(1), 3–33.

A Appendix

A.1 Correlation Structure of $h_t^{\frac{\delta}{2}}$ and Cross-correlations

Proposition 7 *Suppose that* $\mathbf{\Sigma}$ *exists. Then the autocovariances of* $h_t^{\frac{\delta}{2}}$, $\mathbb{C}ov_k(h_t^{\frac{\delta}{2}})$, $k \in \mathbb{N}$, *are given by*

$$\begin{aligned} \mathbb{V}(h_t^{\frac{\delta}{2}}) &= \sigma_\varepsilon^2\gamma^2\lambda^{(0)} + \sigma_v^2\alpha^2\lambda_\phi^{(0)} + 2\sigma_{\varepsilon v}\gamma\alpha\widetilde{\lambda}_\phi^{(0)}, & (17)\\ \mathbb{C}ov_k(h_t^{\frac{\delta}{2}}) &= \sigma_\varepsilon^2\gamma^2\lambda^{(k)} + \sigma_v^2\alpha^2\lambda_\phi^{(k)} + \sigma_{\varepsilon v}\gamma\alpha(\overline{\lambda}_\phi^{(k)} + \widetilde{\lambda}_\phi^{(k-1)}),\ k \geq 1. & (18)\end{aligned}$$

Moreover, the autocovariances of $f(\varepsilon_t)$ *are given by*

$$\begin{aligned} \mathbb{V}[f(\varepsilon_t)] &= \mathbb{V}(h_t^{\frac{\delta}{2}}) + \sigma_v^2,\\ \mathbb{C}ov_k[f(\varepsilon_t)] &= \mathbb{C}ov_k(h_t^{\frac{\delta}{2}}) + \sigma_v^2\psi_k^{(hv)},\ k \geq 1.\end{aligned}$$

Lemma 3 *When $\lambda_1 = \phi$ and $\lambda_2 = c$, we have*

$$\begin{aligned}
\mathbb{V}(h_t^{\frac{\delta}{2}}) &= \frac{1}{1-c^2}\left\{\sigma_v^2\alpha^2 + \frac{\gamma}{1-\phi c}\left[\frac{\sigma_\varepsilon^2\gamma(1+\phi c)}{1-\phi^2} + 2\sigma_{\varepsilon v}\alpha c\right]\right\}, \\
\mathbb{C}ov_k(h_t^{\frac{\delta}{2}}) &= \sigma_\varepsilon^2\gamma^2\lambda^{(k)} + \frac{\sigma_v^2\alpha^2 c^k}{1-c^2} + \frac{\sigma_{\varepsilon v}\gamma\alpha}{(1-\phi c)(\phi - c)}\left[\phi^{1+k} + \frac{c^k(\phi - 2c + \phi c^2)}{1-c^2}\right].
\end{aligned}$$

Further, if $\gamma = 0$, then the above expressions reduce to $\mathbb{C}ov_k(h_t^{\frac{\delta}{2}}) = \frac{\sigma_v^2\alpha^2c^k}{1-c^2}$, $k \geq 0$, which are the autocovariances of the APARCH$(1,1)$ model.

Next, define

$$\begin{aligned}
\lambda_{c\phi}^{(0)} &= \lambda\left\{\frac{\lambda_1[-c + (1+c\phi)\lambda_1 - \phi\lambda_1^2]}{(1-\lambda_1^2)} - \frac{\lambda_1[-c + (1+c\phi)\lambda_1 - \phi\lambda_1^2]}{(1-\lambda_2^2)}\right\}, \\
\lambda_{c\phi}^{(1)} &= \lambda\left\{\frac{\lambda_1^2[-c + (1+c\phi)\lambda_1^{-1} - \phi]}{(1-\lambda_1^2)} - \frac{\lambda_2^2[-c + (1+c\phi)\lambda_2^{-1} - \phi]}{(1-\lambda_2^2)}\right\}, \\
\lambda_{c\phi}^{(k)} &= \lambda\left\{\frac{\lambda_1^{1+k}[-c + (1+c\phi)\lambda_1^{-1} - \phi\lambda_1^{-2}]}{(1-\lambda_1^2)} - \frac{\lambda_2^{1+k}[-c + (1+c\phi)\lambda_2^{-1} - \phi\lambda_2^{-2}]}{(1-\lambda_2^2)}\right\}, \qquad k \geq 2.
\end{aligned}$$

Interestingly, if $\lambda_1 = \phi$ and $\lambda_2 = c$, then the above expressions reduce to

$$\lambda_{c\phi}^{(0)} = \frac{\phi}{1-\phi c},\ \lambda_{c\phi}^{(1)} = \frac{1}{1-\phi c},\ \lambda_{c\phi}^{(k)} = \frac{c^k}{1-\phi c}.$$

Proposition 8 *If $0 < \mathbf{\Sigma} < \infty$ then the covariances between y_{t-k} and h_t, $k \in \mathbb{N}$, are given by*

$$\begin{aligned}
\mathbb{C}ov(y_t, h_t^{\frac{\delta}{2}}) &= \sigma_\varepsilon^2\gamma\widetilde{\lambda}_c^{(0)} + \sigma_v^2\vartheta\alpha^2\overline{\lambda}_\phi^{(0)} + \sigma_{\varepsilon v}\alpha(\widetilde{\lambda}_{c\phi}^{(0)} + \gamma\vartheta\lambda^{(0)}), && (19) \\
\mathbb{C}ov(y_{t-k}, h_t^{\frac{\delta}{2}}) &= \sigma_\varepsilon^2\gamma\overline{\lambda}_c^{(k-1)} + \sigma_v^2\vartheta\alpha^2\widetilde{\lambda}_\phi^{(k-1)} + \sigma_{\varepsilon v}\alpha(\widetilde{\lambda}_{c\phi}^{(k)} + \gamma\vartheta\lambda^{(k)}), \qquad k \geq 1. && (20)
\end{aligned}$$

Lemma 4 *Interestingly if $\lambda_1 = \phi$, $\lambda_2 = c$ then*

$$\begin{aligned}
\mathbb{C}ov(y_t, h_t^{\frac{\delta}{2}}) &= \frac{\sigma_\varepsilon^2\phi\gamma}{(1-\phi c)(1-\phi^2)} + \frac{\sigma_v^2\vartheta\alpha^2}{(1-\phi c)(1-c^2)} + \sigma_{\varepsilon v}\alpha\left[\frac{\phi}{1-\phi c} + \gamma\vartheta\lambda^{(0)}\right], \\
\mathbb{C}ov(y_{t-1}, h_t^{\frac{\delta}{2}}) &= \frac{\sigma_\varepsilon^2\gamma}{(1-\phi c)(1-\phi^2)} + \frac{\sigma_v^2\vartheta\alpha^2 c}{(1-c\phi)(1-c^2)} + \sigma_{\varepsilon v}\alpha\left[\frac{1}{1-\phi c} + \gamma\vartheta\lambda^{(1)}\right].
\end{aligned}$$

Further, if $\gamma = \sigma_{\varepsilon v} = 0$, then $\mathbb{C}ov(y_{t-k}, h_t^{\frac{\delta}{2}}) = \frac{\sigma_v^2\vartheta\alpha^2c^k}{(1-c\phi)(1-c^2)}$.

Next define

$$\zeta_1 = \mathbb{E}(e_t^2)\gamma^2\lambda^{(0)},\ \zeta_2 = \widetilde{\kappa}\alpha^2\lambda_\phi^{(0)},\ \zeta_3 = 2\overline{k}\gamma\alpha\widetilde{\lambda}_\phi^{(0)},\ \zeta = \zeta_1 + \zeta_2 + \zeta_3.$$

Proposition 9 *Let Assumption (A1) be satisfied and asssume $\delta > 1$. If $\mu_{1+1/\delta} < \infty$ and $\zeta_2 < 1$ then μ_2 exists and it is given by*

$$\mu_2 = \frac{\mu_1^2 + \mu_{2/\delta}\zeta_1 + \mu_{1+\frac{1}{\delta}}\zeta_3}{1-\zeta_2}. \tag{21}$$

Moreover, if $\delta = 1$ and $\zeta < 1$, then μ_2 exists and it is given by

$$\mu_2 = \frac{\mu_1^2}{1-\zeta}. \tag{22}$$

Lemma 5 *When $\gamma = \varsigma = 0$ and e_t are normally distributed then equation (21) reduces to*

$$\mu_2 = \frac{\mu_1^2(1-c^2)}{(1-3\alpha^2-\beta^2-2\alpha\beta)}. \tag{23}$$

A.2 Measures of persistence for $h_t^{\frac{\delta}{2}}$

Denote $P_\infty(h_t^{\frac{\delta}{2}}|\varepsilon_t) = \sum_{k=0}^\infty \left(\psi_{h\varepsilon}^{(k)}\right)^2$, $P_\infty(h_t^{\frac{\delta}{2}}|v_t) = \sum_{k=0}^\infty \left(\psi_{hv}^{(k)}\right)^2$ and $P_\infty(h_t^{\frac{\delta}{2}}|\sqrt{\varepsilon_t v_t}) = \sum_{k=0}^\infty \psi_{h\varepsilon}^{(k)}\psi_{hv}^{(k)}$. We can then define the persistence of a shock to $\widetilde{\varepsilon}_t$ and $\widetilde{v}_t$ on the $h_t^{\frac{\delta}{2}}$ variable as

$$P_\infty(h_t^{\frac{\delta}{2}}|\widetilde{\varepsilon}_t) = \sum_{k=0}^\infty \left(\widetilde{\psi}_{h\varepsilon}^{(k)}\right)^2 = \widetilde{\sigma}_\varepsilon^2 P_\infty(h_t^{\frac{\delta}{2}}|\varepsilon_t) + \widetilde{\sigma}_{\varepsilon v} P_\infty(h_t^{\frac{\delta}{2}}|v_t) + 2\widetilde{\sigma}_\varepsilon\sqrt{\widetilde{\sigma}_{\varepsilon v}} P_\infty(h_t^{\frac{\delta}{2}}|\sqrt{\varepsilon_t v_t}), \tag{24}$$

$$P_\infty(h_t^{\frac{\delta}{2}}|\widetilde{v}_t) = \sum_{k=0}^\infty \left(\widetilde{\psi}_{hv}^{(k)}\right)^2 = \widetilde{\sigma}_{\varepsilon v} P_\infty(h_t^{\frac{\delta}{2}}|\varepsilon_t) + \widetilde{\sigma}_v^2 P_\infty(h_t^{\frac{\delta}{2}}|v_t) + 2\widetilde{\sigma}_v\sqrt{\widetilde{\sigma}_{\varepsilon v}} P_\infty(h_t^{\frac{\delta}{2}}|\sqrt{\varepsilon_t v_t}). \tag{25}$$

Proposition 10 *Assume that $\mathbf{\Sigma}$ exists. Then $P_\infty(h_t^{\frac{\delta}{2}}|\widetilde{\varepsilon}_t)$ and $P_\infty(h_t^{\frac{\delta}{2}}|\widetilde{v}_t)$ are given by equations (24) and (25) respectively, where*

$$\begin{aligned}
P_\infty(h_t^{\frac{\delta}{2}}|\varepsilon_t) &= \frac{\gamma^2(1+\lambda_1\lambda_2)}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}, && (26)\\
P_\infty(h_t^{\frac{\delta}{2}}|v_t) &= \alpha^2\left\{1 + \frac{1}{(\lambda_1-\lambda_2)^2}\left[\frac{\lambda_1^2(\lambda_1-\phi)^2}{(1-\lambda_1^2)} + \right.\right. && (27)\\
&\qquad \left.\left. \frac{\lambda_2^2(\lambda_2-\phi)^2}{(1-\lambda_2^2)} - \frac{2\lambda_1\lambda_2(\lambda_1-\phi)(\lambda_2-\phi)}{(1-\lambda_1\lambda_2)}\right]\right\}, \\
P_\infty(h_t^{\frac{\delta}{2}}|\sqrt{\varepsilon_t v_t}) &= \gamma\alpha[1 + \frac{(\lambda_1-\phi)+\lambda_2(1-\phi\lambda_1)}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}].
\end{aligned}$$

Proof. The desired result is obtained straightforwardly from Proposition 3. ■

Lemma 6 *Interestingly if $\lambda_1 = \phi$ and $\lambda_2 = c$, then*

$$\begin{aligned} P_\infty(h_t|\varepsilon_t) &= \frac{\gamma^2(1+\phi c)}{(1-\phi^2)(1-c^2)(1-\phi c)}, \; P_\infty(h_t|v_t) = \frac{\alpha^2}{1-c^2}, \\ P_\infty(h_t|\sqrt{\varepsilon_t v_t}) &= \gamma\alpha[1 + \frac{c}{(1-\phi c)(1-c^2)}]. \end{aligned}$$

Obviously, in this case the expression for $P_\infty(h_t|v_t)$ is the one for the GARCH($1,1$) process (see Fiorentini and Sentana, 1999). Further, notice that if $\gamma = 0$ $P_\infty(h_t|\varepsilon_t) = P_\infty(h_t|\sqrt{\varepsilon_t v_t}) = 0$.

A.3 Proofs

Proof of Proposition 1. Multiplying equation (1) by $(1-cL)$ and substituting (3) into equation (1) gives (5). Similarly, multiplying equation (3) by $(1-\phi L)$ and substituting (1) into equation (3) gives (6). ■

Proof of Proposition 2. Taking expectations from both sides of equation (6) yields (7). ■

Proof of Proposition 3. On account of (5) and equation (A.1) in Karanasos (2007), we obtain equations (8)-(9) by straightforward manipulation. ■

Proof of Proposition 4. The proof follows from the ARMA representation of y_t in equation (5) and Proposition 2 in Karanasos (2007). ■

Proof of Proposition 5. The desired result is obtained straightforwardly from Proposition 3. ■

Proof of Proposition 6. The proof follows from the ARMA representation of y_t in equation (5) and equation (A.10) in Karanasos (2001). ■

Proof of Proposition 7. The proof follows from the ARMA representation of $h_t^{\frac{\delta}{2}}$ in equation (6) and Proposition 2 in Karanasos (2007). ■

Proof of Proposition 8. The proof follows from the ARMA representations of y_t and $h_t^{\frac{\delta}{2}}$ in Lemma 1 and Proposition 3 in Karanasos (2007). ■

Proof of Proposition 9. From Proposition 7 and the fact that $\mathbb{V}(h_t^{\frac{\delta}{2}}) = \mu_2 - {\mu_1}^2$, $\sigma_\varepsilon^2 = \mu_{2/\delta}\mathbb{E}(e_t^2)$, $\sigma_v^2 = \mu_2\widetilde{\kappa}$ and $\sigma_{\varepsilon v} = \mu_{1+\frac{1}{\delta}}\overline{k}$ we obtain

$$\mu_2 = {\mu_1}^2 + \mu_{2/\delta}\zeta_1 + \mu_2\zeta_2 + \mu_{1+\frac{1}{\delta}}\zeta_3.$$

Notice that if $\delta < 1$, both $2/\delta$ and $1+\frac{1}{\delta}$ are greater than 2. When $\delta \geq 1$ the desired result is obtained straightforwardly from the above equation. ■

A.4 Figures

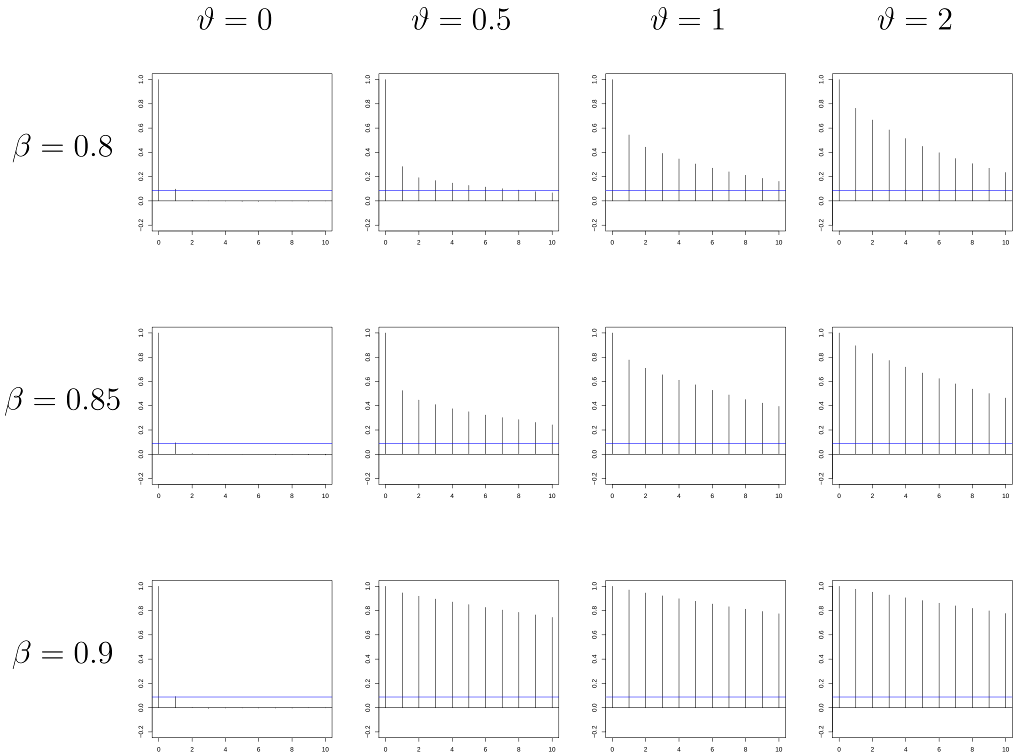

Figure 1: ACF of y_t for $\phi = 0.1$.

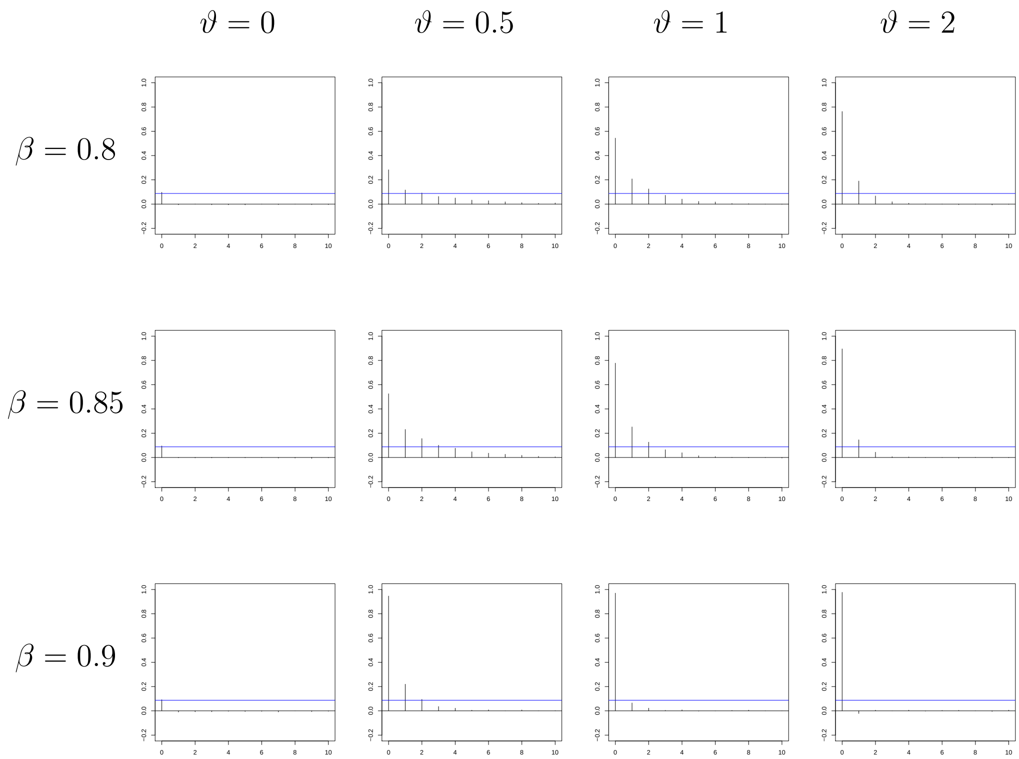

Figure 2: PACF of y_t for $\phi = 0.1$.

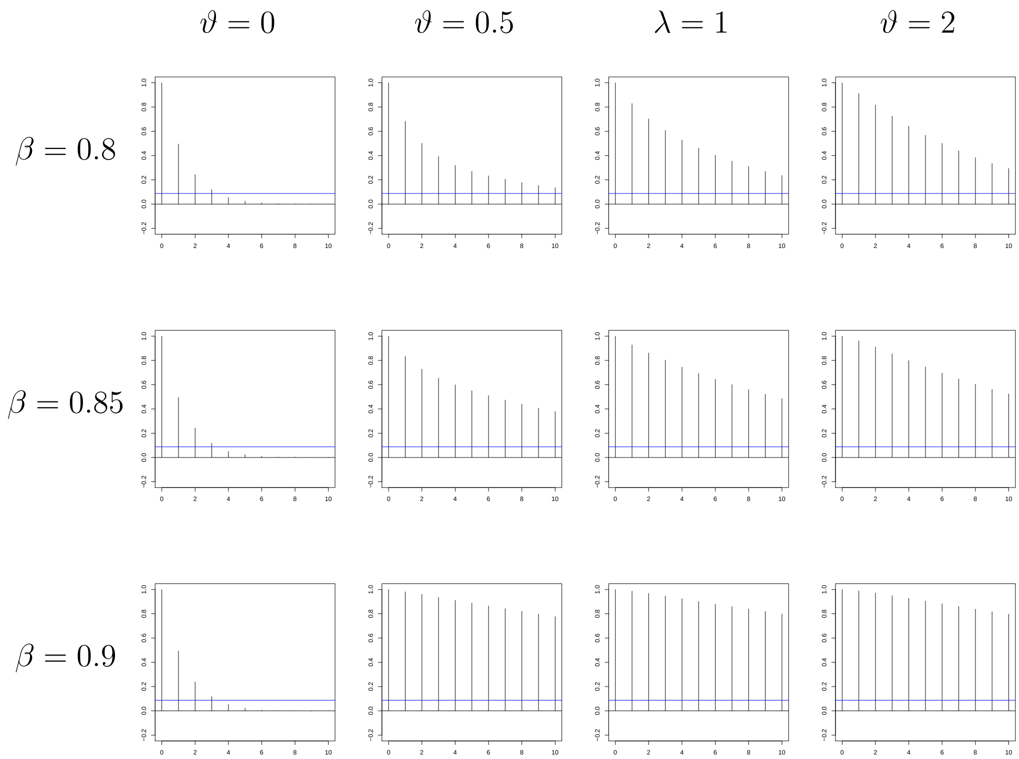

Figure 3: ACF of y_t for $\phi = 0.5$.

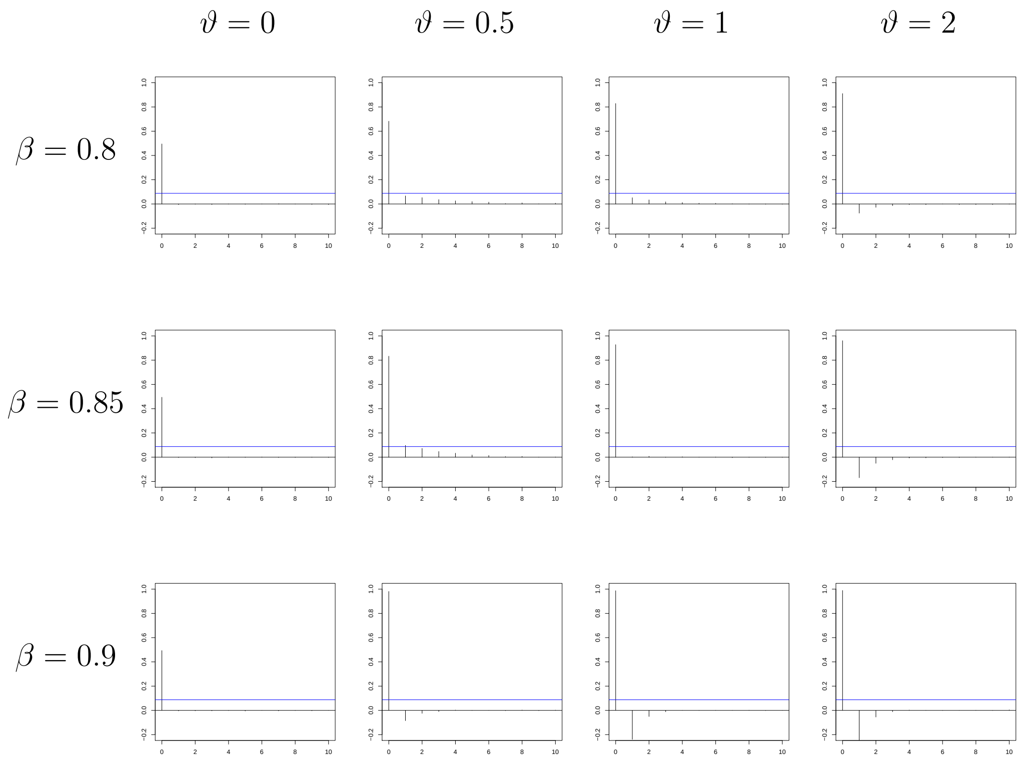

Figure 4: PACF of y_t for $\phi = 0.5$.

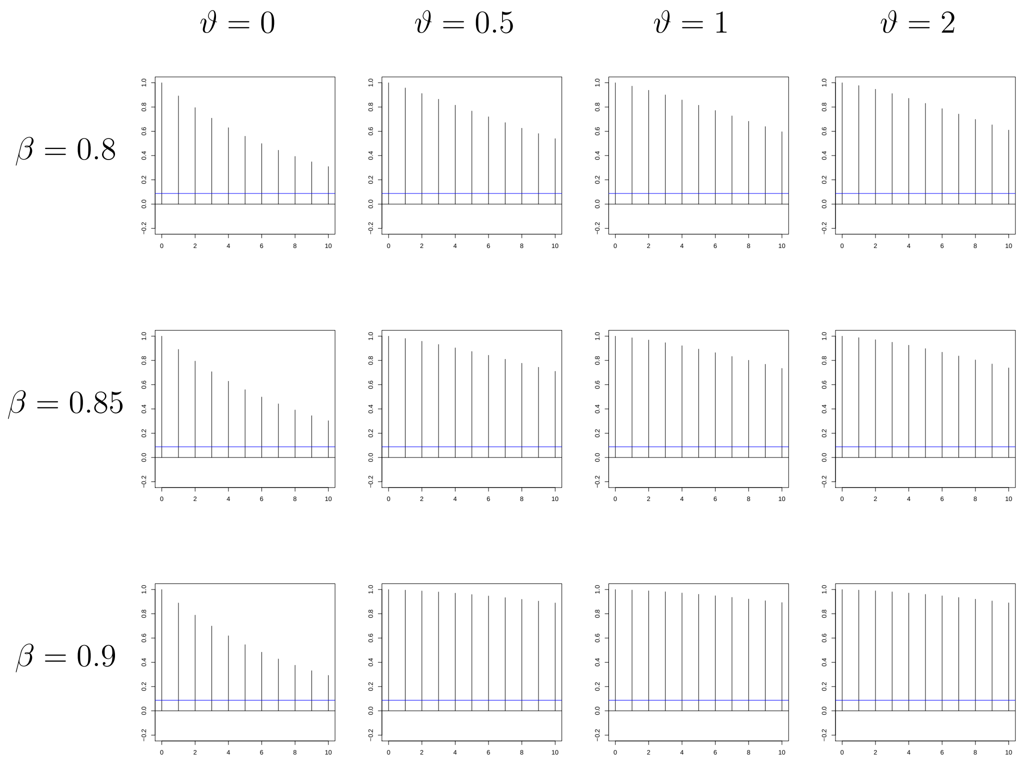

Figure 5: ACF of y_t for $\phi = 0.9$.

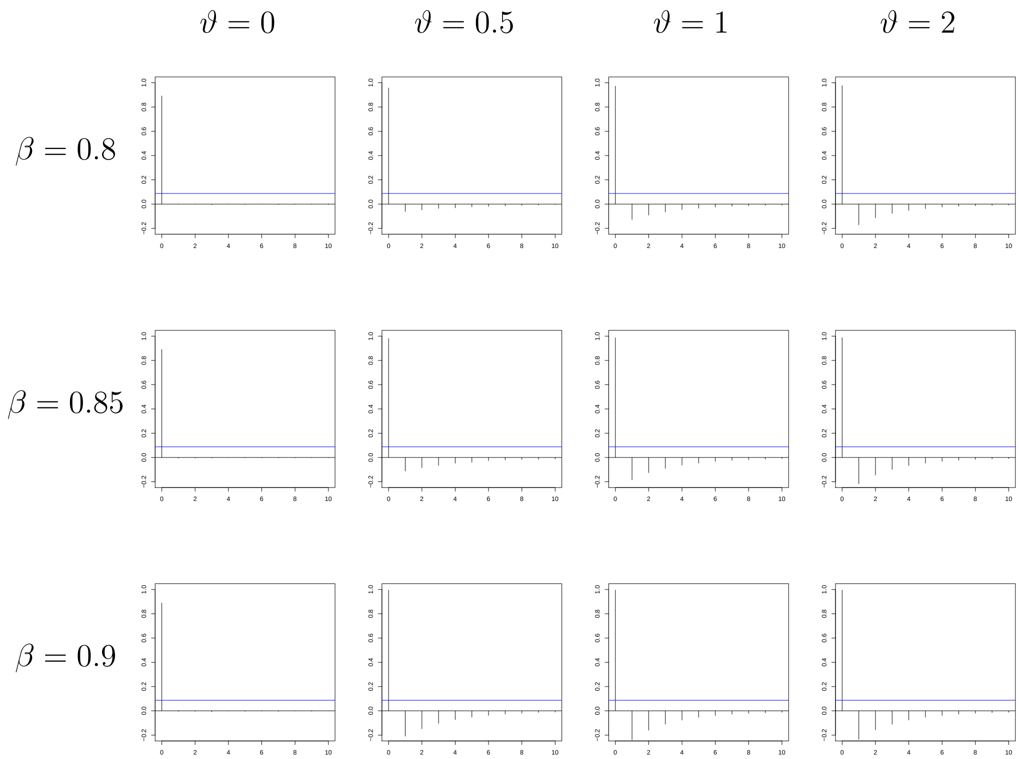

Figure 6: PACF of y_t for $\phi = 0.9$.

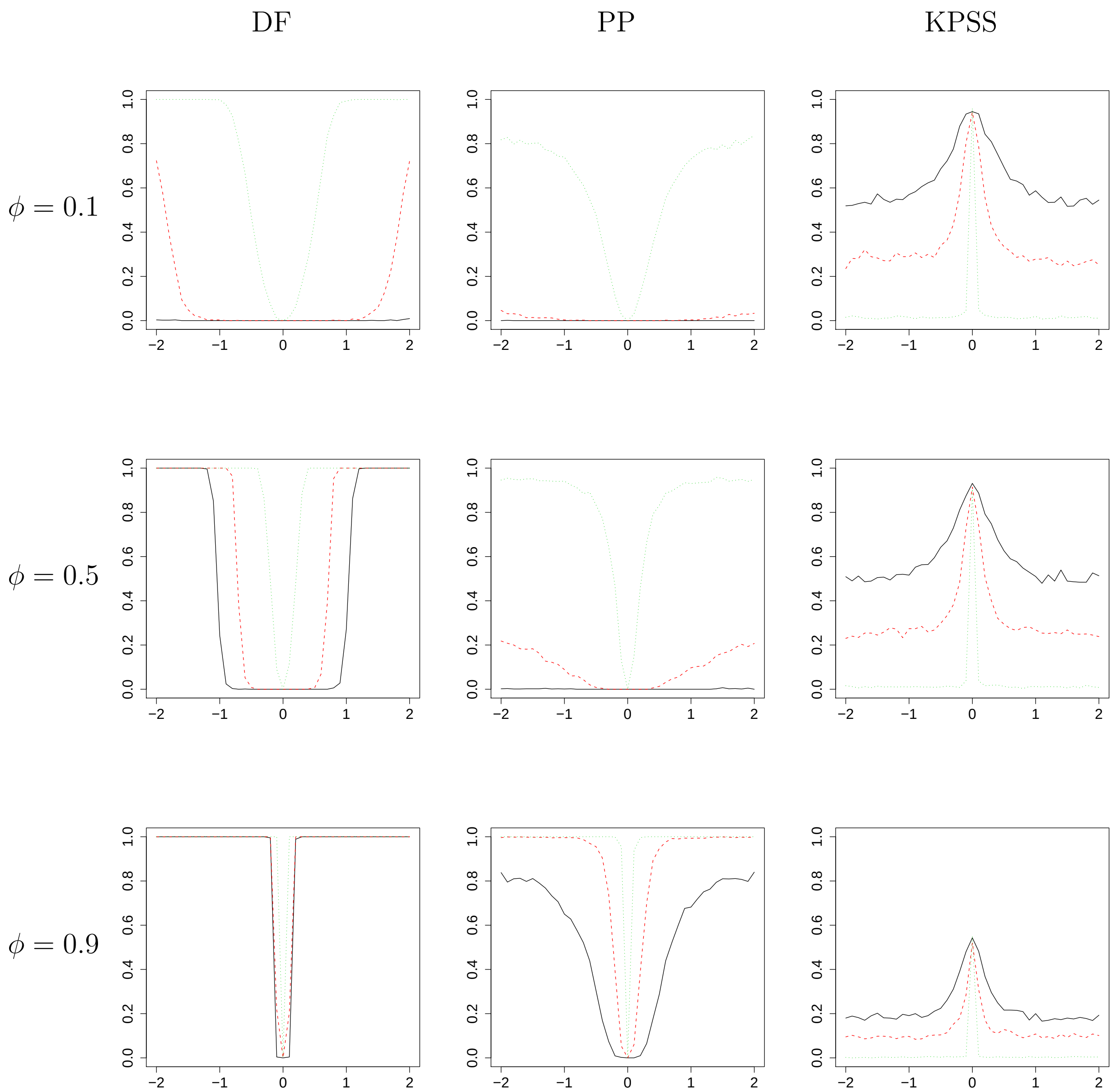

Figure 7: Left: Accepted Nullhypothesis of the DF, PP and KPSS tests (from the left to the right) for $\phi = 0.1, 0.5, 0.9$ (top down), black: $\beta = 0.8$, red: $\beta = 0.85$, green: $\beta = 0.9$, T=500

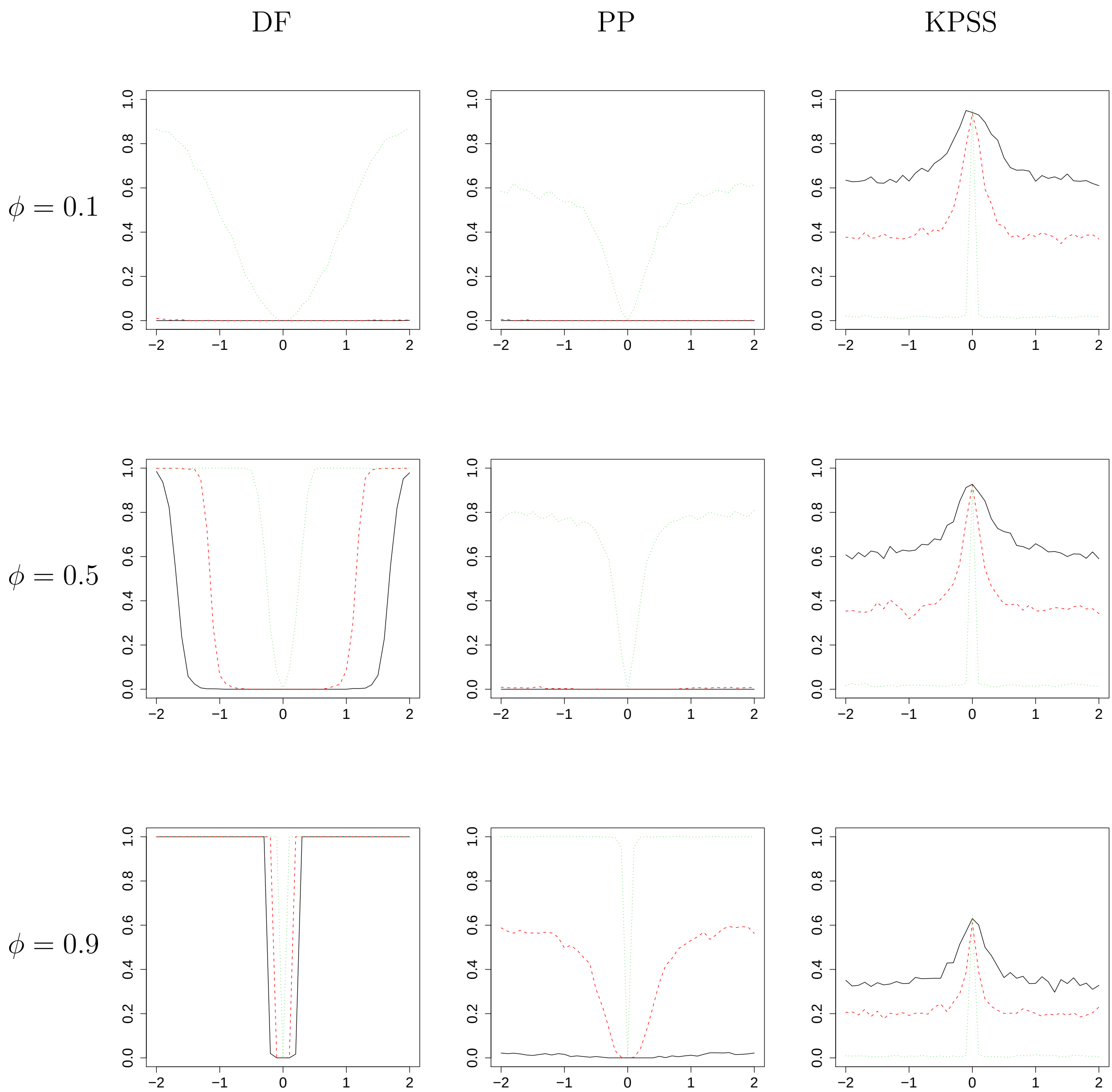

Figure 8: Left: Accepted Nullhypothesis of the DF, PP and KPSS tests (from the left to the right) for $\phi = 0.1, 0.5, 0.9$ (top down), black: $\beta = 0.8$, red: $\beta = 0.85$, green: $\beta = 0.9$, T=1000