AUTOREGRESSIVE GAMMA PROCESSES

GOURIEROUX, C.,[1] and J., JASIAK[2]

This version: February 2003
(first version June 2000)

[1]CREST, CEPREMAP and University of Toronto.
[2]York University, Canada.

Autoregressive Gamma Processes

We introduce autoregressive gamma processes of order p [ARG(p)] with transition distributions which are noncentral gamma up to a change of scale. The paper provides the stationarity and ergodicity conditions for ARG processes of any autoregressive order p, including long memory. The analytical expressions of the conditional moments and nonlinear autocorrelations for any ARG(p) are also given. Moreover, the nonlinear state space representation of an ARG process is used to derive the filtering, smoothing and forecasting algorithms. Finally, the paper introduces estimation and inference methods which are illustrated by an application to interquote durations series of an infrequently traded stock listed on the TSE.

Keywords : Durations, Gamma Process, CIR Process, Nonlinear Canonical Analysis, High Frequency Data, Intertrade Durations.

JEL Number : C14, C22

Processus gamma autoregressif

Nous introduisons des processus de Markov, dont les transitions sont des distributions gamma décentrées à un facteur d'échelle près. Nous étudions les propriétés dynamiques de ces processus. En particulier nous donnons les expressions explicites des deux premiers moments conditionnels, dérivons les autocorrélogrammes d'ordre un et deux et explicitons la décomposition canonique non linéaire du processus. Cette spécification est utilisée pour analyser la dynamique des durées entre transactions.

Mots clés : Durées, processus gamma, processus de Cox-Ingersoll-Ross, analyse canonique non linéaire, données haute- fréquence, duréés entre transactions.

Numéro JEL C14, C22:

1. Introduction

The liquidity of a financial asset on a given day and market depends on how frequently this asset is exchanged, or equivalently on the times between consecutive trades, called intertrade durations. [see Gourieroux, Jasiak, Le Fol (1999)]. The duration analysis of financial assets, such as stocks, is a relatively new topic of research which relies on dynamic models for series of durations observed sequentially in time. In the literature, there exist a limited number of such models. Engle, Russel (1998) introduced the autoregressive conditional duration (ACD) model, with conditional hazard function corresponding to an accelerated hazard model with lagged durations as explanatory variables. While this model is quite easy to estimate, it requires a strong and unrealistic assumption of path independent overdispersion coefficient. Another limitation of the ACD lies in the difficulty of deriving the stationarity and ergodicity conditions for models with temporal dependence over several lags. Ghysels, Gourieroux, Jasiak (2003) have introduced the stochastic volatility duration (SVD) model which improves upon the ACD in that it allows for independent variation of the conditional mean and variance and yields a time varying over- or underdispersion. The SVD is a nonlinear factor model, with two dynamic heterogeneity factors that determine the mean and overdispersion, respectively. It has a limited use in applied research due to complicated estimation procedure. Finally, linear autoregressive models for gamma distributed processes, called the Gamma Autoregressive (GAR) have also been proposed in the time series literature in application to the ruin theory [see e.g. Gaver, Lewis (1980) and the survey by Grundwald et alii (1997)].

The aim of this paper is to extend an alternative dynamic model introduced by Sim (1990, 1994), based on the ideas developed in Griffiths (1970), Kotz, Adams (1984), and to discuss its application to intertrade duration analysis. It will be called the autoregressive gamma model (ARG) in the paper. Its advantage compared to the ACD is that the conditional distributions at all lags are easy to derive, along with the ergodicity conditions. Moreover, unlike the SVD, ARG does not require complicated estimation procedures. Also, it doesn't have the same dynamic properties as the GAR process, some of which are undesirable for financial applications. The ARG-based forecasting, filtering and smoothing algorithms are available and given in the paper. In contrast to the ACD but similar to the SVD, ARG can fit any data that feature time varying over- or underdispersion and accommodate complex nonlinear dynamics. Finally, an ARG of even a small autoregressive order can represent processes with either short, or long memory.

The paper is organized as follows. In section 2, we introduce the autoregressive gamma process (ARG) of order one, with the conditional distribution which is as a noncentered gamma, up to a scale factor, and with a noncentrality parameter written as a linear function of lagged durations. We derive the forecasting formula valid at any horizon and write the stationarity condition. The time dependence is analyzed in section 3 by investigating the conditional moments, the first and second order autocorrelograms and the nonlinear canonical decomposition. The autoregressive gamma model can be viewed as a nonlinear state space model, with an integer valued state variable. In section 4, we study dynamic properties of the state variable which represents a latent counting process. We show how the ARG process is related to the continuous time Cox-Ingersoll-Ross process in section 5, and extend the model to include more lags in the autoregressive part and long memory in section 6. Section 7 compares the ARG with other autoregressive gamma processes proposed in the literature. Statistical inference is discussed in section 8. An application to intertrade durations is presented in Section 9. Contrary to the applied literature on intertrade durations which has been focused on very frequently traded assets and time deformation, we consider an infrequently traded stock for which the liquidity risk is substantial. Section 10 concludes the paper. Proofs are gathered in Appendices.

2. The Autoregressive Gamma Process

We are interested in examining Markov processes, with conditional noncentral gamma distributions up to a change of scale.

Let us recall that a variable Y follows the distribution $\tilde{\gamma}(\delta, \beta, c)$, if and only if Y/c follows the noncentral gamma distribution $\gamma(\delta, \beta)$. This family depends on three nonnegative parameters: the degree of freedom δ, the noncentrality parameter β and the change of scale parameter c.

It is well-known that the noncentral gamma distribution arises as a Poisson mixture of gamma distributions : Y follows the distribution $\gamma(\delta, \beta)$ if and only if there exists a Poisson variable $Z \sim \mathcal{P}(\beta)$, such that the conditional distribution of Y given Z is $\gamma(\delta + Z)$. The expression of the p.d.f. of the $\tilde{\gamma}(\delta, \beta, c)$ distribution is:

$$f_{\delta,\beta,c}(y) = \mathbb{1}_{y>o} \exp(-\frac{y}{c}) \sum_{k=0}^{\infty} \left[\frac{y^{\delta+k-1}}{c^{\delta+k}\Gamma(\delta+k)} \frac{\exp -\beta \ \beta^k}{k!} \right], \tag{2.1}$$

and its two first moments are:

$$EY = c\delta + c\beta, VY = c^2\delta + 2c^2\beta. \tag{2.2}$$

2.1 The autoregressive gamma process

Definition 1 : The process (Y_t) is an autoregressive gamma (ARG) process if and only if the conditional distribution of Y_t given Y_{t-1} is $\tilde{\gamma}(\delta, \beta_t Y_{t-1}, c_t)$, where β_t, c_t are deterministic functions of time.

Definition 1 implies that the degree of freedom is time invariant, and the change of scale parameter is path independent whereas the noncentrality parameter is a linear function of the lagged value. The parameter c_t is nonnegative. We denote : $\rho_t = \beta_t c_t$.

The dynamics of the ARG process can be written as:

$$Y_t = \sum_{i=1}^{Z_t} W_{i,t} + \epsilon_t,$$

where $\epsilon_t, Z_t, W_{i,t}$ are independent conditionally on Y_{t-1} with distributions $\gamma(\nu, c_t), \mathcal{P}[\beta_t y_{t-1}], \gamma(1, c_t)$, respectively [see Sim (1990)]. This specification is useful for simulation studies.

The conditional distribution of an autoregressive gamma process Y_t given Y_{t-h} can be derived explicitly for any h.

Proposition 1 : For an autoregressive gamma process the conditional distribution of Y_t given Y_{t-h} is still a noncentral gamma distribution up to a change of scale : $\tilde{\gamma}(\delta, \beta_{t|t-h} Y_{t-h}, c_{t|t-h})$, where :

$$\begin{aligned}
\rho_{t|t-h} &= \beta_{t|t-h} c_{t|t-h} = \textstyle\prod_{\tau=t-h+1}^{t} \rho_\tau, \\
c_{t|t-h} &= \sum_{\tau=t-h+1}^{t} (c_\tau \rho_{\tau+1} \ldots \rho_t).
\end{aligned}$$

Proof : see Appendix 1.

2.2 Stationary ARG process

Let us consider the homogenous Markov process, for time independent $\beta_t = \beta, c_t = c$ and $\rho_t = \beta c = \rho$. We get the following parameters for the predictive h step ahead distribution:

$$\rho_{t|t-h} = \rho^h,$$

$$c_{t|t-h} = c(1+\rho+\ldots+\rho^{h-1}) = c\frac{1-\rho^h}{1-\rho},$$

with the limiting values :

$$\lim_{h\to\infty} \rho_{t|t-h} = 0, \lim_{h\to\infty} c_{t|t-h} = \frac{c}{1-\rho}, \text{ if } \rho < 1.$$

It follows that:

Proposition 2 :

i) An ARG process is stationary, when the conditional distribution of Y_t given Y_{t-1} is $\tilde{\gamma}(\delta, \beta Y_{t-1}, c)$ with $\rho = \beta c < 1$.

ii) The conditional distribution of Y_t given Y_{t-h} is :

$$\tilde{\gamma}\left[\delta, \frac{\rho^h(1-\rho)}{c(1-\rho^h)} Y_{t-h}, c\frac{1-\rho^h}{1-\rho}\right], \quad \forall h.$$

iii) The marginal invariant distribution is gamma, up to a change of scale :

$$\tilde{\gamma}(\delta, 0, \frac{c}{1-\rho}).$$

We will discuss in Section 6.3 the limiting case of a unitary autoregressive coefficient.

3. Serial Dependence

There exist various methods of examining serial dependence in stationary ARG processes. In this section we consider the first and second order conditional moments, autocorrelograms, and the nonlinear canonical decomposition.

3.1 Conditional moments

The proposition below follows directly from Proposition 2:

Proposition 3 : Let us consider a stationary ARG process. We have :

$$\begin{aligned} E(Y_t|Y_{t-1}) &= c\delta + \rho Y_{t-1}, \\ V(Y_t|Y_{t-1}) &= c^2\delta + 2\rho c Y_{t-1}, \end{aligned}$$

and more generally :

$$\begin{aligned} E(Y_t|Y_{t-h}) &= \frac{c(1-\rho^h)}{1-\rho}\delta + \rho^h Y_{t-h}, \\ V(Y_t|Y_{t-h}) &= \frac{c^2(1-\rho^h)^2}{(1-\rho)^2}\delta + 2\rho^h \frac{c(1-\rho^h)}{1-\rho} Y_{t-h}. \end{aligned}$$

We find that the stationary ARG process is a weak AR(1) with conditional heteroscedasticity represented by a linear function of the lagged value [3].

The expressions of conditional moments reveal the possibility of the process being under or overdispersed. The conditional overdispersion exists if and only if :

$$V(Y_t|Y_{t-1}) > E(Y_t|Y_{t-1})^2$$

$$\iff 0 > c^2\delta(\delta-1) + 2\rho c Y_{t-1}(\delta-1) + \rho^2 Y_{t-1}^2.$$

The discriminant of the polynomial is :

$$\Delta' = -\rho^2 c^2 (\delta - 1).$$

This expression is negative for $\delta > 1$, and positive for $\delta < 1$. When $\delta > 1$, we have $V(Y_t|Y_{t-1}) < E(Y_t|Y_{t-1})^2, \forall Y_{t-1}$. When $\delta < 1$, the product of the roots $\frac{c^2\delta(\delta-1)}{\rho^2}$ is negative, and the two roots are real. These results are summarized in the following proposition.

Proposition 4 : When $\delta > 1$, the stationary ARG process features conditional and marginal underdispersion. When $\delta < 1$, the stationary ARG

[3] However, the process doesn't admit a strong AR(1) representation with i.i.d. innovations as in the general class of models considered by Grunwald et alii (1997).

process features marginal overdispersion. It may feature either under-, or overdispersion depending on the value of Y_{t-1}.

3.2 First and second order autocorrelograms.

From the weak AR(1) interpretation of an ARG process, it follows that:

$$\rho(h) = \text{ Corr } (Y_t, Y_{t-h}) = \rho^h. \tag{3.1}$$

Let us now consider the second order autocorrelogram. It is proven in Appendix 2 that :

$$\rho^{(2)}(h) = \text{ Corr } (Y_t^2, Y_{t-h}^2) = \rho^h \frac{\rho^h + 2(\delta+1)}{2\delta+3}. \tag{3.2}$$

3.3 Nonlinear canonical decomposition

Nonlinear canonical decomposition [see Lancaster (1963)] is a useful tool for the analysis of nonlinear time dependence in a Markov process. The formula of canonical decomposition is derived in Appendix 3 by considering the Cox-Ingersoll-Ross process [see also Sim (1990, page 329) for an approach based on the joint Laplace transform].

Let us denote by :

$$f(y_t;\delta,c,\rho) = \mathbb{1}_{y_t>0} \exp\left[-\frac{(1-\rho)y_t}{c}\right]\left[\frac{1-\rho}{c}\right]^\delta \frac{y_t^{\delta-1}}{\Gamma(\delta)}, \tag{3.3}$$

the marginal p.d.f. of Y_t, and by

$$f_1(y_t|y_{t-1};\delta,c,\rho) = \mathbb{1}_{y_t>0}\exp(-y_t/c)$$
$$\sum_{k=0}^{\infty}\left[\frac{y_t^{\delta+k-1}}{c^{\delta+k}\Gamma(\delta+k)}\exp(-\rho y_{t-1}/c)\frac{(\rho y_{t-1}/c)^k}{k!}\right], \tag{3.4}$$

the conditional p.d.f. of Y_t given Y_{t-1}.

Proposition 5 : We have :

$$f_1(y_t|y_{t-1};\delta,c,\rho) = f(y_t;\delta,c,\rho)\{1+\sum_{n=1}^{\infty}\rho^n\psi_n(y_t,\delta,c,\rho)\psi_n(y_{t-1},\delta,c,\rho)\},$$

where : $\psi_n(y;\delta,c,\rho) = \left(\frac{\Gamma(\delta)\Gamma(n+1)}{\Gamma(\delta+n)}\right)^{1/2} L_n^{(\delta-1)}\left(\frac{(1-\rho)y}{c}\right)$,

and L_n is a generalized Laguerre polynomial :

$$L_n^{(\delta-1)}(z) = \sum_{k=0}^{n}(-1)^k \frac{\Gamma(\delta+n)}{\Gamma(\delta+k)\Gamma(n-k+1)}\frac{z^k}{k!}.$$

The polynomials ψ_n, n varying, satisfy the conditions :

$$\begin{aligned} E\psi_n(Y_t) &= 0,\ \text{Cov}\,[\psi_n(Y_t),\psi_m(Y_t)] = 0, \forall n \neq m, \\ V\psi_n(Y_t) &= 1. \end{aligned}$$

Proposition 5 directly implies the following corollaries.

Corollary 1 : The conditional p.d.f of Y_t given Y_{t-h} admits the nonlinear canonical decomposition :

$$\begin{aligned} & f_h(y_t|y_{t-h};\delta,c,\rho) \\ & = f(y_t;\delta,c,\rho)\{1+\sum_{n=1}^{\infty}\rho^{hn}\psi_n(y_t;\delta,c,\rho)\psi_n(y_{t-h};\delta,c,\rho)\}. \end{aligned}$$

This expression can be used to predict any polynomial transformation of Y_t.

Corollary 2 : We get : $E[\psi_n(Y_t)|Y_{t-1}] = \rho^n\psi_n(Y_{t-1}), \forall n$.

Thus, there exist several nonlinear transformations of the ARG process that follow a weak AR(1) process. For instance,
$\psi_1(Y_t) = \frac{1}{\delta^{1/2}}[\delta - \frac{(1-\rho)Y_t}{c}]$ is AR(1) with the autoregressive coefficient ρ,

$$\psi_2(Y_t) = \left[\frac{2}{\delta(\delta+1)}\right]^{1/2}\left(\frac{\delta(\delta+1)}{2} - (\delta+1)\frac{(1-\rho)Y_t}{c} + \frac{(1-\rho)^2Y_t^2}{2c^2}\right)$$

is AR(1) with the autoregressive coefficient ρ^2.

Corollary 3 : The joint distribution of (Y_t, Y_{t-1}) is symmetric, which means that the process is time reversible.

4. State Space Representation

The dynamic specification for (Y_t) has a nonlinear state space representation, where the state variable is the mixing variable (Z_t), with the transition equation :

$$Z_t|Y_{t-1} \sim \mathcal{P}[\beta Y_{t-1}].$$

The measurement equation is :

$$Y_t|Z_t \sim \gamma[\delta + Z_t, c].$$

The latent variable Z_t can have interesting interpretations in financial applications. For example, if we consider liquidity risk measured by means of intertrade durations (which are the waiting times to trade), the integer valued Z_t can represent a liquidity rating of a stock, such as $Z_t = 0 = AAA, Z_t = 1 = AA$, etc. The model specifies the new rating given the lagged observed liquidity Y_{t-1} and provides a prediction of the next liquidity given the rating.

In the next subsection, we focus on the dynamics of the latent mixing variable (Z_t). In the first part, our analysis is conditioned on the observable process (Y_t) (smoothing), and in the second part, on the past values of (Z_t).

4.1 Smoothing

Let us show how the path of the latent process (Z_t) can be recovered.

Proposition 7 :

i) The values $Z_1, \dots, Z_T$ of the state variable are independent conditional on $Y_o, Y_1, \dots, Y_T$.

ii) The conditional distribution of Z_t given $Y_o, Y_1, \dots, Y_T$ depends on the observable variables Y_{t-1}, Y_t only.

iii) This conditional p.d.f. is given by :

$$l(z_t|y_{t-1}, y_t) = \frac{\dfrac{1}{\Gamma(\delta + z_t)\Gamma(z_t + 1)} \left[\dfrac{\beta y_t y_{t-1}}{c}\right]^{z_t}}{\sum_{z=0}^{\infty} \left\{ \dfrac{1}{\Gamma(\delta + z)\Gamma(z+1)} \left(\dfrac{\beta y_t y_{t-1}}{c}\right)^{z} \right\}}.$$

Proof : See Appendix 4.

4.2 Marginal distribution of the latent process

Proposition 8 : The latent process (Z_t) is a Markov process such that the transition probability :

$$l(z_t|z_{t-1}) = \frac{(\beta c)^{z_t}}{(\beta c+1)^{\delta+z_t+z_{t-1}}} \frac{\Gamma(\delta+z_t+z_{t-1})}{\Gamma(z_t+1)\Gamma(\delta+z_{t-1})}$$

is a negative binomial distribution with parameters :
$n = \delta + z_{t-1}, p = \dfrac{1}{1+\beta c}$.

Proof : See Appendix 5. This distribution describes the migration between liquidity rating classes considered in financial theory.

5. Continuous Time Analogue

It is interesting to see that the continuous time analogue of the ARG process is the Cox-Ingersoll-Ross (CIR) model.

Let us consider the infinitesimal drift and volatility functions, that is, the behavior of a prediction of Y_t at horizon h, when h tends to zero. We get :

$$
\begin{aligned}
& \lim_{h \to 0} \frac{1}{h}[E(Y_t | Y_{t-h} = y) - y] \\
= & \lim_{h \to 0} \frac{1}{h}\left[\frac{c(1-\rho^h)}{1-\rho}\delta + \rho^h y - y\right] \\
= & \lim_{h \to 0} \frac{1-\rho^h}{h}\left[\frac{c\delta}{1-\rho} - y\right] \\
= & -\log \rho \left(\frac{c\delta}{1-\rho} - y\right).
\end{aligned}
$$

$$
\begin{aligned}
& \lim_{h \to 0} \frac{1}{h} V(Y_t | Y_{t-h} = y) \\
= & \lim_{h \to 0} \left\{\frac{1}{2}c^2 \frac{(1-\rho^h)^2}{(1-\rho)^2}\delta + 2\frac{\rho^h}{h}\frac{c(1-\rho^h)}{1-\rho}y\right\} \\
= & -\frac{2c\log\rho}{1-\rho}y.
\end{aligned}
$$

We infer that the limiting continuous time process associated with the ARG(1) process is the Cox-Ingersoll-Ross process [Cox, Ingersoll, Ross (1985)]:

$$
dY_t = a(b - Y_t)dt + \sigma\sqrt{Y_t}dW_t, \tag{5.1}
$$

where : $a = -\log \rho > 0, b = \dfrac{c\delta}{1-\rho}, \sigma^2 = \dfrac{-2\log\rho}{1-\rho}c$. In fact we have the following more general result [see e.g. Lamberton, Lapeyre (1992)].

Proposition 9 : The stationary ARG process is a discretized version of the CIR process.

The link between the ARG and CIR processes points to the importance of the ARG model for applications to intertrade durations. Indeed, it is well-known that volatility and liquidity are strongly related. Thus, a model close to CIR, which is the basic model for stochastic volatility, seems most appropriate for intertrade durations.

6. The Autoregressive Gamma Process of Order p

The ARG specification can be extended to include more lags in the autoregressive term. The advantage of the extension given below is that ARG(p) based prediction formulas and ergodicity conditions are simple and easy to derive.

6.1 Definition

Definition 2 : The process (Y_t) is an autoregressive gamma process of order p [ARG (p)], if and only if the conditional distribution of Y_t given $\underline{Y_{t-1}} = (Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p})'$ is $\tilde{\gamma}(\delta, \beta'\underline{Y_{t-1}}, c)$, where $\beta'\underline{Y_{t-1}} = \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p}$, and $\beta_j \geq 0, j = 1, \ldots, p$.

The noncentrality parameter is now a linear function of the lagged values of the process up to and including the lag p. In this setup the p-dimensional process $\underline{Y_t} = (Y_t, \ldots, Y_{t-p+1})'$ is a Markov process of order one, with the conditional Laplace transform given by :

$$\begin{aligned}
& E\left[\exp -u'\underline{Y_t} | \underline{Y_{t-1}}\right] \\
= \; & E\left[\exp(-u_1 Y_t \ldots - u_p Y_{t-p+1}) | \underline{Y_{t-1}}\right] \\
= \; & \exp\left[-u_2 Y_{t-1} \ldots - u_p Y_{t-p+1}\right] E\left[\exp -u_1 Y_t | \underline{Y_{t-1}}\right] \\
= \; & \left[1 + cu_1\right]^{-\delta} \exp\left(-u_2 Y_{t-1} \ldots - u_p Y_{t-p+1}\right) \exp\left[\frac{-cu_1}{1+cu_1}\beta'\underline{Y_{t-1}}\right] \\
& \text{(from Lemma 1 in Appendix 1)} \\
= \; & \left[1 + cu_1\right]^{-\delta} \exp\left[-A(u)'\underline{Y_{t-1}}\right], \qquad (6.1)
\end{aligned}$$

where :

$$A(u) = \left[\frac{cu_1}{1+cu_1}\beta_1 + u_2, \frac{cu_1}{1+cu_1}\beta_2 + u_3; \ldots, \frac{cu_1}{1+cu_1}\beta_{p-1} + u_p, \frac{cu_1}{1+cu_1}\beta_p\right]. \qquad (6.2)$$

6.2 Stationarity condition

It is shown in Darolles, Gourieroux, Jasiak (2002) that the stationarity condition depends on the function A.

Proposition 9 : The ARG (p) process is stationary if $\lim_{h\to 0} A^{oh}(u) = 0$, for any u with nonnegative components where A^{oh} denotes the function A compounded h times with itself.

This proposition has the following corollary.

Corollary 3 : The stationarity condition is : $c(\beta_1 + \beta_2 + \ldots + \beta_p) < 1$.

Proof : See Appendix 6.

6.3 Stochastic Autoregression and Long Memory

Finally note that long memory can easily be accommodated in the autoregressive gamma framework. As for a standard gaussian autoregressive process, two approaches can be followed. The first one consists in studying the limiting case of a unitary autoregressive coefficient. The second one considers ARG processes with stochastic autoregressive coefficients.

i) **The limiting case** $\rho = 1$.
Let us fix the scale parameter at $c = 1$. When $\rho = \beta = 1$, the conditional Laplace transform of the process becomes [see Appendix 1, Lemma 1]:

$$E[\exp -uY_t | Y_{t-1}] = \exp[-a(u)Y_{t-1} + b(u)],$$

where $a(u) = u/(1+u), b(u) = -\delta \log(1+u)$.

The Laplace transform at horizon h is:

$$E[\exp -uY_{t+h-1} | Y_{t-1}] = \exp[-a^{\circ h}(u)Y_{t-1} + b_h(u)], \text{(say)},$$

where the compounded function $a^{\circ h}$ satisfies:

$$a^{\circ h}(u) = \frac{u}{1+hu}.$$

We see that $\lim_{h\to\infty} a^{\circ h}(u) = 0$, which implies the weak ergodicity of the process [Darolles et alii (2002)]. Moreover, the conditional expectations $E[\exp(-uY_{t+h-1})|Y_{t-1}]$ tend to their limiting value at an hyperbolic rate in h. This implies that the stationary Markov process will have an autocorrelation function with a hyperbolic rate of decay, that is will feature long memory.

This property can be relevant for applications to intertrade durations of frequently traded stocks, in which a long range of persistence is generally observed. The long range persistence can be captured by an ARG model with a small number of lags, whereas it requires a large number of lags in the competing ACD or SVD models [see Jasiak (1998)].

ii) **stochastic autoregressive coefficient**

Let us consider an ARG(1) process with parameters $\delta, \rho = \beta c, c$. When ρ is fixed, the bivariate density of y_t, y_{t-h} can be written as [see Corollary 1]:

$$\begin{aligned} f_h(y_t, y_{t-h}) &= f[y_t; \delta, c/(1-\rho)]\, f[y_{t-h}; \delta, c/(1-\rho)] \\ &\quad \{1 + \sum_{h=1}^{\infty} \rho^{hn} \Psi_n[y_t; \delta, c/(1-\rho)] \Psi_n(y_{t-h}; \delta, c/(1-\rho)]\}. \end{aligned}$$

In this decomposition, the transformed parameters δ and $\gamma = c/(1-\rho)$ appear in the marginal distribution and in the expressions of canonical variates whereas the parameter ρ is a measure of temporal dependence. In particular, the first and second order marginal moments are:

$$\begin{aligned} E(Y_t|\delta, \rho, \gamma) &= m(\delta, \gamma), \\ V(Y_t|\delta, \rho, \gamma) &= \sigma^2(\delta, \gamma), \end{aligned}$$

and depend on δ and $\gamma = c/(1-\rho)$ only, whereas the covariance function is given by:

$$Cov(Y_t, Y_{t-h}|\delta, \rho, \gamma) = \rho^h \sigma^2(\delta, \gamma).$$

Let us now assume that the autoregressive parameter ρ is stochastic and takes values from the interval [0,1] with probability density π, while δ and $\gamma = c/(1-\rho)$ remain fixed. By integrating the stochastic parameter ρ out of the first and second order moments, we obtain:

$$\begin{aligned} E(Y_t|\delta, \pi, \gamma) &= m(\delta, \gamma), \\ V(Y_t|\delta, \pi, \gamma) &= \sigma^2(\delta, \gamma) \end{aligned}$$

$$\begin{aligned}
Cov(Y_t, Y_{t-h}) &= E[Cov(Y_t, Y_{t-h}|\delta, \rho, \gamma)|\delta, \pi, \gamma)] \\
&\quad + Cov[E(Y_t|\delta, \rho, \gamma), E(Y_{t-h}|\delta, \rho, \gamma)|\delta, \pi, \gamma)] \\
&= E[Cov(Y_t, Y_{t-h}|\delta, \rho, \gamma)|\delta, \pi, \gamma)] \\
&= E[\rho^h \sigma^2(\delta, \gamma)|\delta, \pi, \gamma)] \\
&= \sigma^2 E_\pi(\rho^h).
\end{aligned}$$

We find that the ACF is now defined by:

$$Corr(Y_t, Y_{t-h}|\delta, \pi, \gamma) = E_\pi(\rho^h).$$

Consequently, the autocorrelation function features hyperbolic decay when the heterogeneity distribution of the coefficient ρ assigns sufficiently large probabilities to values close to one. For example, this is the case when the distribution of ρ is a beta distribution [see Granger, Joyeux (1980)].

7. Other Dynamic Specifications for Gamma Processes

In the time series literature, there exist different dynamic models for gamma processes. None of them have been used in applications to intertrade durations. In this section we discuss the main model that exists in the literature, called the Gamma Autoregressive process (GAR) and introduced by Gaver, Lewis (1980). They considered the linear autoregression:

$$Y_t = \rho Y_{t-1} + \epsilon_t,$$

where the innovations are i.i.d. random variables. They proved that for $0 \leq \rho < 1$, there exists a distribution of the error term ϵ_t such that Y_t has a marginal exponential distribution (or more generally a gamma distribution). When the marginal distribution of Y_t is exponential $\gamma(1, \lambda)$, the innovation distribution is a mixture of a point mass at zero with probability ρ and an exponential distribution $\gamma(1, \lambda)$ with probability $1 - \rho$.

Let us point out that the GAR model differs in three respects from the ARG process:

i) By construction, the GAR is conditionally homoscedastic and doesn't allow for independent variation of the conditional mean and variance of Y_t. In application to intertrade durations, the assumption of constant duration volatility is a very strong restriction on the form of liquidity risk.

ii) The transition distribution admits a discrete component, which implies deterministic dynamics $Y_t = \rho Y_{t-1}$ with probability ρ.

iii) The dependence between serial correlation and innovation distribution is difficult to interpret. For example, in the unit root case $\rho = 1$, the GAR process (Y_t) takes value zero, while remaining stationary.

In fact, ARG and GAR processes are special cases of CAR (Compound Autoregressive) processes. Indeed, autoregressive processes with marginal gamma distributions are easy to construct by considering a conditional Laplace transform of the type:

$$E[\exp -uY_t|Y_{t-1}] = \exp[-a(u)Y_{t-1} + c(u) - c[a(u)]],$$

where c is the log-Laplace transform of the marginal gamma distribution. The models differ essentially by the form of function a, which determines serial dependence.

Table 1: Autoregressive Models for Gamma Processes

Gaver, Lewis (1980): GAR	$a(u) = \rho u$
Gourieroux, Jasiak (2003): ARG Sim (1990), (1994)	$a(u) = \rho u/(1 + cu)$

8. Statistical Inference

Let us consider a sample of discrete time observations $y_1, \ldots, y_T$, of size T. We briefly review standard estimation methods such as GMM and the maximum likelihood in application to ARG processes.

8.1 Moment estimators

The parameters of the ARG (p) process can be estimated from the first and second order conditional moments, since :

$$\begin{aligned} E\left(Y_t|\underline{Y_{t-1}}\right) &= c\delta + c\beta'\underline{Y_{t-1}}, \\ V\left(Y_t|\underline{Y_{t-1}}\right) &= c^2\delta + 2c^2\beta'\underline{Y_{t-1}}. \end{aligned}$$

Consistent estimators of $c\delta$ and $c\beta_j$ $j = 1, \dots, p$ are obtained by regressing Y_t on $1, Y_{t-1}, \dots, Y_{t-p}$. The efficiency of these estimators can be improved by taking into account the conditional hereroscedasticity. For instance, we can apply a pseudo maximum likelihood method based on a Gaussian pseudo family. The PML estimators are the solutions of :

$$\left(\hat{c}, \hat{\delta}, \hat{\beta}\right)' = \arg\max_{c,\delta,\beta} \sum_{t=p+1}^{T} \left\{ -\frac{1}{2} \log\left(c^2\delta + 2c^2\beta'\underline{y_{t-1}}\right) - \frac{1}{2} \frac{\left(y_t - c\delta - c\beta'\underline{y_{t-1}}\right)^2}{c^2\delta + 2c^2\beta'\underline{y_{t-1}}} \right\}.$$

8.2 Maximum likelihood estimators

The log-likelihood function :

$$L_T(\delta, c, \beta) = \sum_{t=1}^{T} \log f_1(y_t|\underline{y_{t-1}}; \delta, c, \beta), \tag{8.1}$$

does not admit a simple form and in practice has to be approximated. The possible approximations for $p = 1$ are discussed below.

i) We can truncate the series expansion of the conditional density, which leads to :

$$L_T^{(1)}(\delta, c, \rho) = -\sum_{t=1}^{T} \frac{1}{c}(y_t + \rho y_{t-1}) + \sum_{t=1}^{T} \log\left[\sum_{k=0}^{K} \frac{y_t^{\delta+k-1}}{c^{\delta+k}\Gamma(\delta+k)} \frac{(\rho y_{t-1}/c)^k}{k!}\right]. \tag{8.2}$$

ii) We can also truncate the nonlinear canonical decomposition :

$$L_T^{(2)}(\delta, c, \rho) = \sum_{t=1}^{T} \log f(y_t; \delta, c, \rho) + \sum_{t=1}^{T} \log[1 + \sum_{n=1}^{N} \rho^n \psi_n(y_t; \delta, c, \rho)\psi_n(y_{t-1}; \delta, c, \rho)].$$

Intuitively ii) is a method of moments involving all marginal and cross moments $E(Y_t^m Y_{t-1}^m)$, for $m, n, \leq N$. A drawback of this approach is the possibility of getting negative values of the likelihood due to truncation.

iii) Finally we can apply a simulated maximum likelihood method based on artificial drawings of the latent Poisson state variable :

$$L_T^{(3)}(\delta, c, \rho) = -\sum_{t=1}^{T} \frac{y_t}{c} + \sum_{t=1}^{T} \log \left[\sum_{s=1}^{S} \frac{y_t^{\delta + z_t^s - 1}}{c^{\delta + z_t^s} \Gamma(\delta + z_t^s)} \right],$$

where S is the number of drawings and $(z_t^1, \ldots, z_t^S)$ are independently drawn in the Poisson distribution $\mathcal{P}(\rho y_{t-1}/c) = \mathcal{P}(\beta y_{t-1})$.

The advantage of the last approach is that it can be easily extended to any autoregressive order p, since this would instead require drawing the Poisson state variable in the Poisson distribution $\mathcal{P}[\beta' \underline{y_{t-1}}]$.

9. Application to interquote durations

In the applied literature, the analysis of intertrade durations typically concerns very frequently traded stocks which feature long range duration persistence. However, liquidity risk is much more important to be taken into account for infrequently traded (illiquid) stocks. The aim of this section is to examine an example of an infrequently traded stock with long waiting times to the next upcoming transaction. In consequence of the dynamics, the number of observations available in our sample is considerably lower that in other applications concerning a sampling period of similar length.

The ARG model has been estimated from data on interquote durations of the Dayton Mining stock traded on the Toronto Stock Exchange in October 1998. The choice of a sotck from the mining industry is not arbitrary. In the past, a false announcement of a gold mine discovery has undermined the credibility of the whole sector and diminished the frequency at which those stocks are traded.

Prior to estimation, durations between quotes with the same time stamp were aggregated [4]. The times between market openings and closures were deleted from the sample. We also removed from the sample all observations on quotes recorded during market preopenings. Thus, we are mainly concerned with the intraday dynamics of interquote durations. Dayton Mining belongs to infrequently traded stocks, and is not frequently quoted either. The sample consists of 381 observations, with mean 1162.5 sec. [approx 20 min] [5] and variance 4762851 [standard error of about 35 min]. The kernel

[4] In particular, we disregard simultaneous trades at either market opening or during the day, when a large buy (sell) order is filled by several sell (buy) orders.

[5] To compare with an average intertrade duration less than one minute for a frequently

smoothed marginal density of the Dayton interquote durations is displayed in Figure 1. It features the typical decreasing shape of a gamma distribution with a small degree of freedom.

Insert Figure 1: Dayton Mining Durations

In the first approach, several ARG models with different numbers of lags are estimated by OLS. The Dayton Mining series displays a rather short range of serial dependence in interquote durations, typical for infrequently traded stocks. This differs from the long range persistence observed for very frequently traded assets. Since the coefficients on lags equal and greater than two were found nonsignificant, we retained for further analysis the ARG(1) and ARG(2) models. Both models were estimated by QMLE. Table 2 below summarizes the results obtained from fitting the ARG(1) and ARG(2) models to the sample of Dayton Mining using the regression approach and QMLE.

Table 2: Estimation of ARG(1) and ARG(2)

PARAMETER	ARG(1)		ARG(2)	
	OLS	QMLE	OLS	QMLE
$c\delta$	10.024E2	10.527E2	8.925E2	9.396E2
	(1.25E2)	(0.26E2)	(1.35E2)	(0.24E2)
ρ_1	0.132	0.0794	0.1196	0.0201
	(0.050)	(0.057)	(0.051)	(1.262)
ρ_2	-	-	0.1104	0.1612
	-	-	(0.051)	(0.393)

We observe that the estimates of coefficients ρ_1 and ρ_2 are positive, and their sum is less than one, so that the stationarity condition is satisfied for the estimated models. Both models fit the data fairly well, and have all significant coefficients. While the mean log likelihood of ARG(2), equal to (-8.154), shows little improvement compared to that of ARG(1), equal to (-8.175), the R-squared has doubled when we added the second lag (0.030 compared to 0.018 [6]).

traded stock, such as the IBM.

[6] Note that the value of the R^2 are rather small since they measure linear serial dependence in a nonlinear dynamic framework.

In the next step, the parameters of the ARG(2) are used in a simulation experiment. It consists in fixing the vector of initial durations at two consecutive values observed at the end of the sample: $y(T) = 5226$ and $y(T-1) = 7512$, and simulating the next four observations $y(T+1), y(T+2), y(T+3), y(T+4)$. We replicate this experiment 1000 times. Next we compute the average predicted waiting times for 1, 2, 3, and 4 quotes, given by $y(T+1), [y(T+1)+y(T+2)]/2$, $[y(T+1)+y(T+2)+y(T+3)]/3$, and $[y(T+1)+y(T+2)+y(T+3)+y(T+4)]/4$, consecutively. They provide a natural measure of future liquidity. More precisely, if we consider that the quotes correspond to trades with the same traded volume V, say, these measures represent the average time necessary to trade the basic quantity V of assets when we want to trade $V, 2V, 3V, 4V$, etc [see Gourieroux, Jasiak, Lefol (1999)]. By comparing the distributions, we see how the liquidity cost measured by the time necessary to trade without causing a price change, depends on the traded quantity. Their means and variances are reported below, while their distributions are plotted in Figure 2.

Table 3: One-step Predictions from ARG(2)

	$y(T+1)$	two quotes	three quotes	four quotes
mean	1793.93	1985.58	1815.43	1940.20
variance	8929084.7	10497642	12056015	13455554

Insert Figure 2: Comparison of Distributions of Predicted Durations

10. Conclusion

This paper provided an overview of properties of autoregressive gamma processes. In practice, ARG models can represent dynamics of various positively valued time series such as durations, and of other positive variables related to durations. For example, the squared returns, which are the proxies for return volatility, are directly linked to the intertrade durations, which measure liquidity. Also, an ARG model can fit the volumes per trade which are considered as the second liquidity component. The ARG specification is sufficiently flexible to accommodate short as well as long memory processes for which the stationarity and ergodicity conditions are available and derived in this paper. The estimation relies on a variety of methods, ranging from a simple regression to more sophisticated simulation-based estimators. The fit of the model was tested on interquote duration data, and proven to be satisfactory.

Figure 1: Dayton Mining, durations

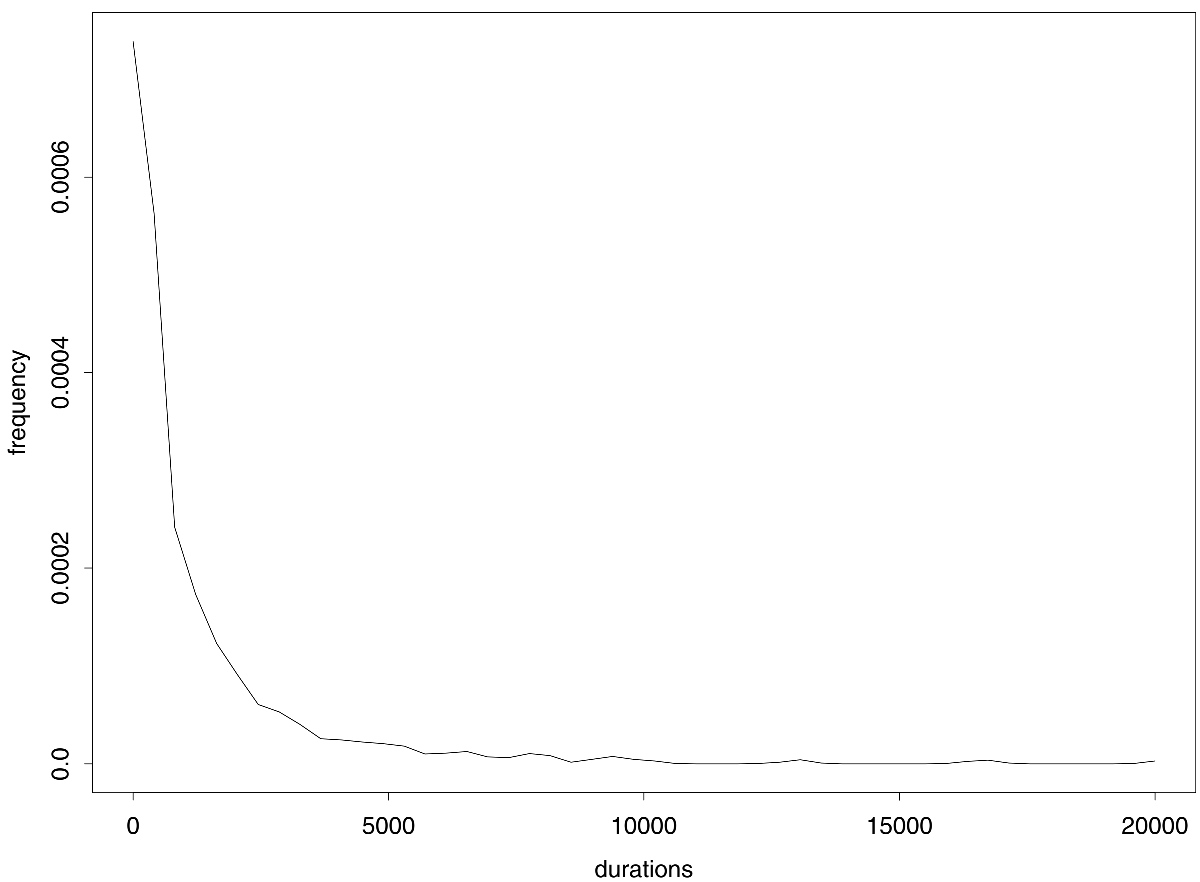
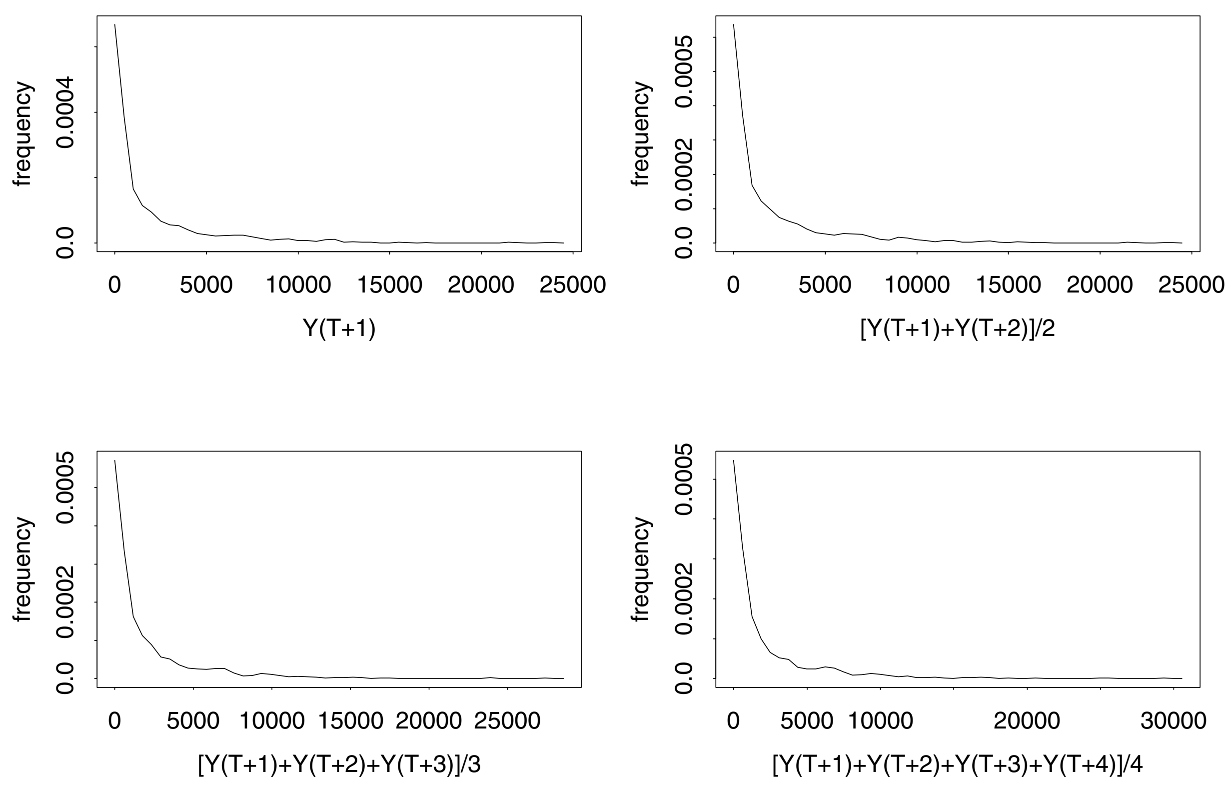


Figure 2: Distributions of predicted durations



Appendix 1
Conditional distribution at any horizon

Let us prove the result for $h = 2$ by using the conditional Laplace transforms. The argument is the same for any horizon.

Lemma 1 : $E[\exp -uY_t|Y_{t-1}]$

$$= (1 + c_t u)^{-\delta} \exp\{-Y_{t-1}\frac{\beta_t c_t u}{1 + c_t u}\}.$$

Proof : We have :

$$\begin{aligned}
& E[\exp -uY_t|Y_{t-1}] \\
= \ & E[[E\exp -uY_t|Z_t]|Y_{t-1}) \\
= \ & E[(1 + c_t u)^{-(\delta + Z_t)}|Y_{t-1}], \text{ where} : Z_t|Y_{t-1} \sim \mathcal{P}[\beta_t Y_{t-1}], \\
= \ & (1 + c_t u)^{-\delta} \exp\{-\beta_t Y_{t-1}(1 - \frac{1}{1 + c_t u})\} \\
= \ & (1 + c_t u)^{-\delta} \exp\{-Y_{t-1}\frac{\beta_t c_t u}{1 + c_t u}\}.
\end{aligned}$$

QED

Let us now consider the conditional Laplace transform at horizon $h = 2$. We get :

$$
\begin{aligned}
& E[\exp -uY_t|Y_{t-2}] \\
= \quad & E[E[\exp -uY_t|Y_{t-1}]|Y_{t-2}] \\
= \quad & (1+c_t u)^{-\delta} E|\exp -Y_{t-1}\frac{\beta_t c_t u}{1+c_t u}|Y_{t-2}] \\
= \quad & (1+c_t u)^{-\delta}(1+c_{t-1}\frac{\beta_t c_t u}{1+c_t u})^{-\delta}\exp\{-Y_{t-2}\frac{\beta_{t-1}c_{t-1}\frac{\beta_t c_t u}{1+c_t u}}{1+c_{t-1}\frac{\beta_t c_t u}{1+c_t u}}\} \\
= \quad & [1+(c_t+c_{t-1}\beta_t c_t)u]^{-\delta}\exp\{-Y_{t-2}\frac{\beta_t c_t \beta_{t-1}c_{t-1}u}{1+(c_t+c_{t-1}\beta_t c_t)u}\} \\
= \quad & (1+c_{t|t-2}u)^{-\delta}\exp\{-Y_{t-2}\frac{\beta_{t|t-2}c_{t|t-2}u}{1+c_{t|t-2}u}\}.
\end{aligned}
$$

Appendix 2
Second order autocorrelogram

We have :

$$\begin{aligned}
\gamma^{(2)}(h) &= \text{Cov}\,(Y_t^2, Y_{t-h}^2) \\
&= E(Y_t^2 Y_{t-h}^2) - (EY_t^2)^2 \\
&= E[E(Y_t^2|Y_{t-h})Y_{t-h}^2] - (EY_t^2)^2 \\
&= E[V(Y_t|Y_{t-h})Y_{t-h}^2] + E[E(Y_t|Y_{t-h})^2 Y_{t-h}^2] - (EY_t^2)^2 \\
&= E((c_{t|t-h}^2\delta + \rho_{t|t-h}c_{t|t-h}Y_{t-h})^2 Y_{t-h}^2) + E[(c_{t|t-h}\delta + \rho_{t|t-h}Y_{t-h})^2 Y_{t-h}^2] \\
&\quad -(EY_t^2)^2 \\
&= c_{t|t-h}^2\delta(1+\delta)m_2 + 2\rho_{t|t-h}c_{t|t-h}(1+\delta)m_3 \\
&+ \rho_{t|t-h}^2 m_4 - m_2^2,
\end{aligned}$$

$$\text{where : } m_j = E(Y_t^j) = \frac{c^j}{(1-\rho)^j}\delta(\delta+1)\ldots(\delta+j-1).$$

We find:

$$\begin{aligned}
\gamma^{(2)}(h) &= c^2\frac{(1-\rho^h)^2}{(1-\rho)^2}\delta^2(\delta+1)^2\frac{c^2}{(1-\rho)^2} \\
&+ 2\rho^h c\frac{1-\rho^h}{1-\rho}(1+\delta)\delta(\delta+1)(\delta+2)\frac{c^3}{(1-\rho)^3} \\
&+ \rho^{2h}\frac{c^4}{(1-\rho)^4}\delta(\delta+1)(\delta+2)(\delta+3) - \frac{c^4}{(1-\rho)^4}\delta^2(\delta+1)^2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{c^4}{(1-\rho)^4}\delta(\delta+1)\Big[(1-\rho^h)^2\delta(\delta+1)+2\rho^h(1-\rho^h)(\delta+1)(\delta+2) \\
&\quad +\rho^{2h}(\delta+2)(\delta+3)-\delta(\delta+1)\Big] \\
&= \frac{c^4}{(1-\rho)^4}\delta(\delta+1)\Big[\rho^{2h}[\delta(\delta+1)-2(\delta+1)(\delta+2)+(\delta+2)(\delta+3)] \\
&\quad +\rho^h(-2\delta(\delta+1)+2(\delta+1)(\delta+2)]\Big] \\
&= \frac{c^4}{(1-\rho)^4}\delta(\delta+1)[2\rho^{2h}+4(\delta+1)\rho^h] \\
&= \frac{c^4}{(1-\rho)^4}2\rho^h\delta(\delta+1)[\rho^h+2(\delta+1)] \\
&= \gamma^{(2)}(0)\rho^h\frac{\rho^h+2(\delta+1)}{2\delta+3}.
\end{aligned}
$$

Appendix 3
The nonlinear canonical analysis

Since an ARG process is a discretized version of the CIR process (see Proposition 9), we can derive the nonlinear canonical decomposition of the conditional distribution of Y_t given Y_{t-1} from the spectral analysis of the infinitesimal generator associated with the CIR process.

1) Spectral decomposition of the infinitesimal generator

The infinitesimal generator A associated with the diffusion equation :

$$dY_t = a(b - Y_t)dt + \sigma\sqrt{Y_t}dW_t,$$

is given by :

$$\psi \to A\psi(y) = a(b - y)\frac{d\psi}{dy}(y) + \frac{1}{2}\sigma^2 y\frac{d^2\psi}{dy^2}(y).$$

Let us verify that the eigenfunctions of the infinitesimal generator are polynomials.

i) Let us first consider a polynomial of degree n, ψ_n say, satisfying the condition :

$$a(b - y)\frac{d\psi_n}{dy}(y) + \frac{1}{2}\sigma^2 y\frac{f d^2\psi_n}{dy^2}(y) = \lambda_n\psi_n(y),$$

where λ_n is the corresponding eigenvalue. By identifying the coefficients of y^n, we see that : $\lambda_n = -na = n\log\rho$. This implies the necessary form of the eigenvalue.

ii) Next we need to solve the differential equations:

$$a(b - y)\frac{d\psi_n}{dy}(y) + \frac{1}{2}\sigma^2 y\frac{d^2\psi_n}{dy^2}(y) + na\psi_n(y) = 0,$$

$$\text{or :}\quad (b - y)\frac{d\psi_n}{dy}(y) + \frac{1}{2}\frac{\sigma^2}{a}y\frac{d^2\psi_n}{dy^2}(y) + n\psi_n(y) = 0.$$

Let us introduce the change of variable :

$$z = \frac{2a}{\sigma^2} y, \psi_n(y) = \psi_n^*(\frac{2a}{\sigma^2} y) = \psi_n^*(z), \tag{A.1}$$

where : $2a/\sigma^2 = (1-\rho)/c$.

The differential equation becomes :

$$\frac{2a}{\sigma^2}(b-y)\frac{d\psi_n^*}{dz}(z) + \frac{1}{2}\frac{\sigma^2}{a} y \frac{(2a)^2}{(\sigma^2)^2}\frac{d^2\psi_n^*}{dz^2}(z) + n\psi_n(z) = 0,$$

$$\left(\frac{2ab}{\sigma^2} - z\right)\frac{d\psi_n^*}{dz}(z) + z\frac{d^2\psi_n^*}{dz^2}(z) + n\psi_n(z) = 0, \tag{A.2}$$

where : $\dfrac{2ab}{\sigma^2} = \delta$. It is known that the solution of differential equation (A.2) is proportional to the generalized Laguerre polynomial [Abramowitz, Stegun (1970), formula 22.6.15] :

$$L_n^{(\delta-1)}(z) = \sum_{k=0}^{n} (-1)^k \frac{\Gamma(\delta+n)}{\Gamma(\delta+k)\Gamma(n-k+1)}\frac{z^k}{k!}. \tag{A.3}$$

When integrated with respect to the $\gamma(\delta)$ distribution, these polynomials satisfy [Abramowitz, Stegun (1970), formula 22.2.12] :

$$\int_0^\infty \left[L_n^{(\delta-1)}(z)\right]^2 \frac{1}{\Gamma(\delta)} \exp(-z) z^{\delta-1} dz = \frac{\Gamma(\delta+n)}{\Gamma(\delta)\Gamma(n+1)}.$$

We derive the standardized polynomials :

$$L_n^{*(\delta-1)}(z) = \left[\frac{\Gamma(\delta)\Gamma(n+1)}{\Gamma(\delta+n)}\right]^{1/2} L_n^{(\delta-1)}(z). \tag{A.4}$$

Finally the eigenfunctions of the initial differential equation are derived by applying the change of variable (A.1) :

$$\psi_n(y) = \left[\frac{\Gamma(\delta)\Gamma(n+1)}{\Gamma(\delta+n)}\right]^{1/2} L_n^{(\delta-1)}\left(\frac{(1-\rho)y}{c}\right). \tag{A.5}$$

Therefore $\psi_n, n = 0, 1, \dots$ is a basis of eigenfunctions of the infinitesimal generator associated with the eigenvalues $\lambda_n = \log \rho^n$.

iv) It is known that for a univariate diffusion equation the functions ψ_n, n varying, are also eigenfunctions of the conditional expectation operator :

$\psi \to E(\psi(Y_t)|Y_{t-1} = y)$, associated with the eigenvalues : $\exp \lambda_n = \rho^n$. Moreover the process is time reversible. Then the conditional distribution of Y_t given Y_{t-1}, $f_1(y_t|y_{t-1}; \delta, \rho, c)$ (say) can be written as (see Lancaster (1963)) :

$$f_1(y_t|y_{t-1}; \delta, \rho, c) = f(y_t; \delta, \rho, c)\{1 + \sum_{n=1}^{\infty} \rho^n \psi_n(y_t)\psi_n(y_{t-1})\}, \tag{A.6}$$

where $f(y_t; \delta, \rho, c)$ is the marginal distribution of Y_t, i.e. the distribution $\tilde{\gamma}(\delta, 0, \frac{c}{1-\rho})$.

Appendix 4
Proof of Proposition 7

Let us denote by $g(y_t|z_t)$ the conditional p.d.f of Y_t given Z_t and by $h(z_t|y_{t-1})$ the conditional p.d.f. of Z_t given Y_{t-1}. Then the conditional distribution of $Y_1, \ldots, Y_T, Z_1, \ldots, Z_T$ given Y_0 is :

$$l(y_1, \ldots, y_T, z_1, \ldots, z_T|y_o) = \sum_{\tau=1}^{t}[g(y_\tau|z_\tau)h(z_\tau|y_{\tau-1})].$$

We find that :

$$\begin{aligned}
& l(z_1, \ldots, z_T|y_o, y_1, \ldots, y_T) \\
= \; & \frac{l(y_1, \ldots, y_T, z_1, \ldots, z_T|y_o)}{l(y_1, \ldots, y_T|y_o)} \\
= \; & \frac{\prod_{\tau=1}^{t}[g(y_\tau|z_\tau)h(z_\tau|y_{\tau-1})]}{\prod_{\tau=1}^{t}\sum_{z=0}^{\infty}[g(y_\tau|z)h(z|y_{\tau-1})]} \\
= \; & \prod_{\tau=1}^{t}\left[\frac{g(y_\tau|z_\tau)h(z_\tau|y_{\tau-1})}{\sum_{z=0}^{\infty} g(y_\tau|z)h(z|y_{\tau-1})}\right] = \prod_{\tau=1}^{t} l(z_\tau|y_\tau, y_{\tau-1}), (\text{say}).
\end{aligned}$$

Therefore we note that :

i) $Z_1, \ldots, Z_T$ are independent, conditionally to $Y_o, Y_1, \ldots, Y_T$;
ii) the conditional distribution :

$$l(z_t|y_o, y_1, \ldots, y_T) = l(z_t|y_t, y_{t-1}),$$

depends on y_{t-1}, y_t only ;

$$
\text{iii)}\quad l(z_t|y_t,y_{t-1}) = \frac{g(y_t|z_t)h(z_t|y_{t-1})}{\sum_{z=0}^{\infty}\{g(y_t|z)h(z|y_{t-1})\}}
$$

$$
= \frac{\frac{1}{c^{\delta+z_t}}\exp\left(-\frac{y_t}{c}\right)\frac{y^{\delta+z_t-1}}{\Gamma(\delta+z_t)}\exp(-\beta y_{t-1})\frac{(\beta y_{t-1})^{z_t}}{\Gamma(z_t+1)}}{\sum_{z=0}^{\infty}\left\{\frac{1}{c^{\delta+z}}\exp\left(-\frac{y_t}{c}\right)\frac{y_t^{\delta+z-1}}{\Gamma(\delta+z)}\exp(-\beta y_{t-1})\frac{(\beta y_{t-1})^{z}}{\Gamma(z+1)}\right\}}
$$

$$
= \frac{\frac{1}{\Gamma(\delta+z_t)\Gamma(z_t+1)}\left[\frac{\beta y_t y_{t-1}}{c}\right]^{z_t}}{\sum_{z=0}^{\infty}\left\{\frac{1}{\Gamma(\delta+z)\Gamma(z+1)}\left(\frac{\beta y_t y_{t-1}}{c}\right)^{z}\right\}}.
$$

Appendix 5
Transition distribution of the factor process

We get :

$$
\begin{aligned}
& l(z_t|z_{t-1}) \\
= & \int l(z_t, y_{t-1}|z_{t-1}) dy_{t-1} \\
= & \int h(z_t|y_{t-1}) g(y_{t-1}|z_{t-1}) dy_{t-1} \\
= & \int \exp(-\beta y_{t-1}) \frac{\beta^{z_t} y_{t-1}^{z_t}}{\Gamma(z_t+1)} \frac{y_{t-1}^{\delta+z_{t-1}-1}}{c^{\delta+z_{t-1}}\Gamma(\delta+z_{t-1})} \exp\left(-\frac{y_{t-1}}{c}\right) dy_{t-1} \\
= & \frac{\beta^{z_t}}{\Gamma(z_t+1)\Gamma(\delta+z_{t-1})} \frac{1}{c^{\delta+z_{t-1}}} \int \exp[-y_{t-1}(\beta+\frac{1}{c})] y_{t-1}^{\delta+z_t+z_{t-1}-1} dy_{t-1} \\
= & \frac{\beta^{z_t}}{c^{\delta+z_{t-1}}} \frac{1}{(\beta+\frac{1}{c})^{\delta+z_t+z_{t-1}}} \frac{\Gamma(\delta+z_t+z_{t-1})}{\Gamma(z_t+1)\Gamma(\delta+z_{t-1})} \\
= & \frac{(\beta c)^{z_t}}{(\beta c+1)^{\delta+z_t+z_{t-1}}} \frac{\Gamma(\delta+z_t+z_{t-1})}{\Gamma(z_t+1)\Gamma(\delta+z_{t-1})}.
\end{aligned}
$$

Appendix 6
Stationarity condition

From Proposition 9, we have to infer the conditions, which ensure that the solution of the p-dimensional recursive system :

$$\begin{aligned} X_{1,t} &= \frac{cX_{1,t-1}}{1+cX_{1,t-1}}\beta_1 + X_{2,t-1}, \\ &\vdots \\ X_{p-1,t} &= \frac{cX_{1,t-1}}{1+cX_{1,t-1}}\beta_{p-1} + X_{p,t-1}, \\ X_{p,t} &= \frac{cX_{1,t-1}}{1+cX_{1,t-1}}\beta_p, \end{aligned}$$

tends to $(0,\dots 0)'$, when t tends to infinity, for any nonnegative initial value $(X_{1,0},\dots,X_{p,0})'$. The system is equivalent to :

$$\begin{aligned} X_{1,t} &= X_{p,t}\frac{\beta_1}{\beta_p} + X_{2,t-1}, \\ &\vdots \\ X_{p-1,t} &= X_{p,t}\frac{\beta_{p-1}}{\beta_p} + X_{p,t-1}, \\ X_{p,t} &= \beta_p - \frac{\beta_p}{1+cX_{1,t-1}}. \end{aligned}$$

It follows from this system that $X_{j,t}, j = 1,\dots,p$ take nonnegative values and that $X_{p,t}$ is always smaller than β_p.

Moreover the sequence $(X_{p,t})$ satisfies the nonlinear recursive equation :

$$X_{p,t} = \beta_p - \frac{\beta_p}{1 + c\left[X_{p,t-1}\frac{\beta_1}{\beta_p} + \ldots + X_{p,t-p+1}\frac{\beta_{p-1}}{\beta_p} + X_{p,t-p}\right]}.$$

A possible limiting value l of this sequence satisfies :

$$l = \beta_p - \frac{\beta_p}{1 + cl\left[\frac{\beta_1}{\beta_p} + \ldots + \frac{\beta_{p-1}}{\beta_p} + 1\right]}.$$

Therefore the admissible values are $l = 0$ and $l = \beta_p \left[1 - \frac{1}{c(\beta_1 + \ldots + \beta_p)}\right]$.

If $c(\beta_1 + \ldots + \beta_p) < 1$, the sequence $(X_{p,t})$ takes values in the compact set $[0, \beta_p]$, with a unique admissible limiting value $l = 0$. We infer its convergence to zero.

If $c(\beta_1 + \ldots + \beta_p) > 1$, we can get the convergence of $(X_{p,t})$ to $\beta_p \left[1 - \frac{1}{c(\beta_1 + \ldots + \beta_p)}\right]$ by selecting appropriately the initial value.

REFERENCES

Abramowitz, M., and I., Stegun (1970) : "Handbook of Mathematical Functions", 9th edition, Dover.

Barret, J., and D., Lampard (1955) : "An Expansion for Some Second Order Probability Distributions and Its Applications to Noise Problems", IRE Trans. Information Theory, IT-1, 10-15.

Bauwens, L., and P., Giot (1997) : "The Logarithmic ACD Model : An Application to the Bid—Ask Quote Process of the NYSE Stocks", CORE DP 9789.

Cherian, K. (1941) : "A Bivariate Gamma Type Distribution Function", J. Indian Math. Soc., 5, 133-144.

Cox, J., Ingersoll, J., and S., Ross (1985) : "A Theory of the Term Structure of Interest Rates", Econometrica, 53, 385-408.

Darolles, S., Gourieroux, C., and J. Jasiak (2002) : "Structural Laplace Transform and Compound Autoregressive Models", CREST DP.

Ding, Z., and C., Granger (1996) : "Modeling Volatility Persistence of Speculative Returns : A New Approach", Journal of Econometrics, 73, 185-215.

Eagleson, G. (1964) : "Polynomial Expansions of Bivariate Distributions", Annals of Mathematical Statistics, 35, 1208-1215.

Engle, R., and J., Lange (1997) : "Measuring, Forecasting and Explaining Time Varying Liquidity in the Stock Market", DP Univ. of California, San Diego.

Engle, R., and J., Russell (1998) : "The Autoregressive Conditional Duration Model: A New Model for Irregularly Spaced Data", Econometrica 66, 1127-1162.

Gaver, D., and P. Lewis (1980): "First Order Autoregressive Gamma Sequences and Point Processes", Adv. Appl. Probab., 12, 727-745.

Ghysels, E., Gourieroux, C., and J., Jasiak (2003) : "Stochastic Volatility Duration Models", Journal of Econometrics, forthcoming

Gourieroux, C., Jasiak, J., and G., Le Fol (1999) : "Intraday Market Activity", Journal of Financial Markets, Vol.2, 193-226.

Granger, C., and R. Joyeux (1980): "An Introduction to Long Memory Time Series Models and Fractional Differencing", Journal of Time Series Analysis, 1, 15-29.

Griffiths, R. (1970): "Infinitely Divisible Multivariate Gamma Distributions", Sankhya, A32, 393-404.

Grunwald, G., Hyndman, R., Tedesco, L., and R. Tweedie (1997): "A Unified View of Linear AR(1) Models", Colorado State University, D.P.

Jacobs, R., and P. Lewis (1978): " Discrete Time Series Generated by Mixtures I: Correlational and Runs Properties", JRSS, B, 40, 94-105.

Jasiak, J. (1998): "Persistence in Intertrade Durations", Finance, 19, 166-195.

Karatzas, I., and S., Shreve (1988) : "Brownian Motion and Stochastic Calculus", New-York, Springer Verlag.

Kotz, S., and J., Adams (1964): "Distribution of a Sum of Identically Distributed Exponentially Correlated Gamma Variables", Annals of Math. Stat., 35, 277-283.

Lamberton, D., and B. Lapeyre (1992): "Introduction au calcul stochastique appliqué à la Finance", Mathematique et Applications, Ellipses-Editions, Paris.

Lampard, D. (1968): " A Stochastic Process whose Intervals Between Events Form a First Order Markov Chain", J. Appl. Prob., 5, 648-668.

Lancaster, H., (1958) : "The Structure of Bivariate Distribution", Annals of Mathematical Statistics, 29, 719-736.

Lancaster, H., (1963) : "Correlations and Canonical Forms of Bivariate

Distributions", Ann. Math. Statist, 34, 532.538.

Lawrance, A., and P. Lewis (1981): "A New Autoregressive Time Series Models in Exponential Variables", Adv. Appl. Probab., 13, 826-845.

Lawrance, A., and P. Lewis (1985): "Modelling and Residual Analysis of Nonlinear Autoregressive Time Series in Exponential Variables", JRSS, B, 47, Lloyd, E., and S. Saleem. (1979): "A Note on Seasonal Markov Chains

with Gamma or Gamma-like Distributions", Journal of Appl. Probab., 16, 117-128.

Lunde, A. (1997) : "A Conjugate Gamma Model for Duration in Transaction Data", DP. Univ. of Aarhus.

Phatarfod, R. (1971): "Some Approximate Results in Renewal and Dam Theories", Journal of Australian Mathematical Society, 12, 425-432.

Sim, C. (1986): "Simulation of Weibull and Gamma Autoregressive Stationary Process", Commun. Statist., B15, 1141-1146.

Sim, C. (1990): "First Order Autoregressive Models for Gamma and Exponential Processes", J. Appl. Probab., 27, 325-332.

Sim, C. (1994): "Modelling Nonnormal First Order Autoregressive Time Series", Journal of Forecasting, 13, 369-381.

Smith, S., and J., Miller (1986) : "A Non-Gaussian State Space Model and Application to Prediction of Records", Journal of the Royal Statistical Society, B, 48, 79-88.

Tong, H. (1990) : "Nonlinear Time Series : A Dynamic System Approach", Oxford University Press.

Wong, E. (1964) : "The Construction of a Class of Stationary Markov Processes", Sixteenth Symposium on Applied Mathematics, American Mathematical Society, 264-276.

Wong, E., and J., Thomas (1962) : "On Polynomial Expansions of Second

Order Distributions", SIAM J. Appl. Math, 10, 507-516.