Journal of Quantitative Economics, Vol. 10 No.2, July 2012

DAMPED TREND EXPONENTIAL SMOOTHING: PREDICTION AND CONTROL

GIACOMO SBRANA[1]

Abstract

Damped trend exponential smoothing models have gained importance in empirical studies due to their remarkable forecasting performance. This paper derives their theoretical forecast error variance, based on the implied ARIMA model, as algebraic function of the structural parameters. As a consequence, the minimum mean squared error (MMSE) forecasts as well as the h-step ahead theoretical forecast error variances can also be expressed as algebraic (and unique) functions of the structural parameters. Analytical results are provided for the random coefficient state space model, as introduced by McKenzie and Gardner (2010) "Damped trend exponential smoothing: A modeling viewpoint", International Journal of Forecasting, 26 (4), 661-665, in the single source of error context. Moreover, algebraic results are also given for standard Holt-Winters (damped) trend models in the multiple sources of errors context.

Keywords: Exponential smoothing, State-space models, Time Series, ARIMA models, Forecasting

JEL Classification: C22, C53

1. Introduction

The use of trend exponential smoothing models has been widely employed in forecasting time series. Those models, also known as Holt-Winters methods of exponential smoothing, have been further developed in the last twenty years. In particular, during the eighties, the pioneering contributions of Everette S. Gardner & Eddie Mckenzie suggested to damp the trend as the forecast horizon increases (see for example Gardner &Mckenzie 1985, 1988,1989). More specifically, given that forecasting a series into the future using a straight line is not necessarily appropriate, the authors suggested to add an autoregressive (damping) parameter. Subsequently, damped trend exponential smoothing models gained importance in empirical studies due to their remarkable forecasting properties. This is confirmed by Armstrong(2006), who recommended these models as they improve forecasting accuracy. In addition, Fildes, Nikolopous, Crone & Syntetos (2008) identify the damped trend as a benchmark model to beat. The importance of using damped trend is also recognized by Hyndman, Koehler, Ord & Snyder (2008). The choice of adding a damping parameter is particularly important from a modeling point of view. This is because the damping term makes the model extremely flexible in fitting the

[1] NEOMA Business School, 1 rue du Maréchal Juin, BP 215, 76130 Mont Saint Aignan, France. E-mail address: gsb@rouenbs.fr

dynamics of the series. More specifically, as described in Gardner and McKenzie(2011), the damped parameter can be interpreted as a measure of the persistence of different linear trends. Recently, McKenzie & Gardner (MG)(2010) suggest a random coefficient state space model. This model, as described below, has the same ARIMA(1,1,2) reduced form of the standard damped trend but it introduces further flexibility.

Yet, despite the increasing development of the parametrization of those models, we still know little on the theoretical forecast error variance of (damped) trend exponential smoothing models. In other words, the lack of knowledge of the linkage between the ARIMA reduced form parameters and those belonging the data generation process (structural process) represents a clear gap. This paper fills this gap. In what follow an algebraic procedure that allows deriving the theoretical variance of different damped trend exponential smoothing models is shown. These algebraic results are relevant since they provide full control of the forecasting properties of the underlying models. The paper is organized as follows: Section 2 shows the algebraic results when the single source of error (SSOE) framework is adopted. More specifically, the linkages between the reduced form ARIMA parameters and those of the random coefficient state space models are fully described and discussed. Section 3 provides algebraic linkages when the multiple sources of errors (MSOE) framework is employed. In particular, the algebraic relations among the theoretical forecast error variances and the structural parameters of both trend and damped trend exponential smoothing are shown.

2. SSOE

This section reconsiders and extends the results as in MG. More specifically, the theoretical forecast error variance, based on the implied ARIMA model, is derived for the random coefficient state space model. Accordingly, the same notation as in MG is fully adopted. Here, results are shown in the context of SSOE. Yet, in the next section this assumption will be relaxed by allowing for the more general MSOE context.

Consider the following error-correction form of a linear trend with additive errors:

$$\begin{aligned} y_t &= l_{t-1}+b_{t-1}+\epsilon_t \\ l_t &= l_{t-1}+b_{t-1}+(1-\alpha)\epsilon_t \\ b_t &= b_{t-1}+(1-\beta)\epsilon_t \end{aligned} \quad \text{... (1)}$$

Note that the coefficients of the innovations in the level and gradient equations are written slightly different compared with the standard notation adopted by the exponential smoothing literature. Yet, considering that this paper is an extension of MG, the same notation is fully adopted here. The damped version of the previous model can be expressed suchthat:

$$\begin{aligned} y_t &= l_{t-1}+\phi b_{t-1}+\epsilon_t \\ l_t &= l_{t-1}+\phi b_{t-1}+(1-\alpha)\epsilon_t \\ b_t &= \phi b_{t-1}+(1-\beta)\epsilon_t \end{aligned} \quad \text{... (2)}$$

MG show both the parameters of the implied ARIMA(0,2,2) belonging to (1) and those of the ARIMA(1,1,2) belonging to (2) are trivial to derive. In addition, in both cases the error term of the ARIMA model corresponds to ϵ_t as in the error-correctionform. The random coefficient state-space model can be expressed as follows:

$$y_t = l_{t-1} + A_t b_{t-1} + \epsilon_t$$
$$l_t = l_{t-1} + A_t b_{t-1} + (1-\alpha^*)\epsilon_t \qquad \text{... (3)}$$
$$b_t = A_t b_{t-1} + (1-\beta^*)\epsilon_t$$

where A_t is a sequence of independent, identically distributed binary random variates with $P(A_t=1)=\phi$ and $P(A_t=0)=1-\phi$. Note that (3) is a variant of the standard damped trend model. Yet, this model is particularly flexible since it is a stochastic mixture of standard linear trend and simple exponential smoothing (i.e. random walk plus noise) models. This feature is not only appealing from a modeling viewpoint, but it is also in line with Brown (1963) who argued that the parameters of the model may change from one segment to another one as processes are thought to be *locally constant*.

The reduced form of (3) is a random coefficient ARIMA as follows:

$$(1-A_t B)z_t = \epsilon_t - (\alpha^* + A_t\beta^*)\epsilon_{t-1} + A_t\alpha^*\epsilon_{t-2} \qquad \text{... (4)}$$

where B is the backshift operator and $z_t=(1-B)y_t$. In addition, after some algebra, the following autocovariances can be obtained:

$$E(z_t z_t) = \sigma_\epsilon^2\left((1+\alpha^{*2}) + \phi\left((1-\beta^*)^2 - 2\alpha^*(1-\beta^*)\right) + \phi^2\left(\frac{(1-\beta^*)^2}{(1-\phi)}\right)\right)$$
$$E(z_t z_{t-1}) = \sigma_\epsilon^2\left(-\alpha^* + \phi(1-\beta^*) - \alpha^*\phi^2(1-\beta^*) + \phi^2\left(\frac{(1-\beta^*)^2}{(1-\phi)}\right)\right) \qquad \text{... (5)}$$
$$E(z_t z_{t-2}) = \sigma_\epsilon^2\left(\phi^2(1-\beta^*)(1-\phi\alpha^*) + \phi^3\left(\frac{(1-\beta^*)^2}{(1-\phi)}\right)\right)$$
$$E(z_t z_{t-3}) = \phi E(z_t z_{t-2})$$

Therefore, as shown by MG, z_t can be generated by the following stochastic difference equation (i.e. ARIMA(1,1,2)):

$$(1-\phi B)z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \qquad \text{... (6)}$$

Note that the variance of a_t is crucial since it is the theoretical forecast error variance of the random coefficient state space model. MG claim that the parameters of (6) are complicated functions of the four parameters of (3). Yet, using the algebraic procedure provided in the appendix, it can be shown that the moving average parameters of (6) are exact functions of α^*; β^*; ϕ; σ_ϵ^2.To see this, note that the autocovariance function of $(1-\phi B)z_t$ are:

$$\gamma_0 = (1+\phi^2)E(z_t z_t) - 2\phi E(z_t z_{t-1})$$
$$\gamma_1 = (1+\phi^2)E(z_t z_{t-1}) - \phi E(z_t z_{t-2}) - \phi E(z_{t-1} z_{t-1}) \qquad \text{... (7)}$$
$$\gamma_2 = E(z_t z_{t-2}) - \phi E(z_{t-1} z_{t-2}) = \alpha^*\phi\sigma_\epsilon^2$$
$$\gamma_k = 0 \text{ for } k \geq 3$$

At this stage,[2] given the results in the Appendix, the autocovariances are sufficient to derive the ARIMA parameters. Thus:

[2] For the sake of clarity, note that $\gamma_2=(1+\phi^2)E(z_t z_{t-2}) - \phi E(z_t z_{t-3}) - \phi E(z_{t-1} z_{t-2})$, but given that $E(z_t z_{t-3}) = \phi E(z_t z_{t-2})$, it collapses to $\sigma_\epsilon^2\phi\alpha^*$.

$$\theta_2=\frac{\alpha^*\phi\sigma_\epsilon^2}{\sigma_a^2}$$
$$\theta_1=\frac{(1+\phi^2)E\left(z_tz_{t-1}\right)-\phi E\left(z_tz_{t-2}\right)-\phi E(z_{t-1}z_{t-1})}{\alpha^*\phi\sigma_\epsilon^2+\sigma_a^2} \quad \text{... (8)}$$

with:

$$\sigma_a^2=\frac{1}{4}(\gamma_0-2\gamma_0+G+\sqrt{2}\sqrt{\gamma_0^2+\gamma_0G-2\left(\gamma_1^2+\gamma_2\left(2\gamma_2+G\right)\right)} \quad \text{... (9)}$$

And

$$G=\sqrt{(\gamma_0-2\gamma_1+2\gamma_2)(\gamma_0+2\gamma_1+2\gamma_2)} \quad \text{... (10)}$$

where $E(z_tz_t)$, $E(z_tz_{t-1})$ and $E(z_tz_{t-2})$ are those provided in (5). These expressions give full control of the theoretical variance of the random coefficient state space model.[3] As a consequence, the minimum mean squared error (MMSE) forecasts as well as the *h-step ahead* theoretical forecast error variance are also algebraic functions of the error-correction parameters.

Interestingly, the algebraic expression of σ_a^2can be used to derive the set of parameters that allows to obtain the following equality $\sigma_a^2=\sigma_\epsilon^2$. This set is shown in Figure 1. More specifically, each point shown in Figure 1 is a combination of α^*, $\beta^*$and ϕ that yields $\sigma_a^2=\sigma_\epsilon^2$. This finding implies that it is not necessarily true that, as claimed by MG, the moving average parameters in (6) must be different from those of the error-correction specification as in (3).

In empirical applications, one might be interested in comparing the implied parameters as in (8) and (9) with the ARMA estimated parameters. A similar exercise was carried out for example in Morley, Nelson & Zivot (2003) to compare the Beveridge-Nelson decomposition with unobserved components models.

3. MSOE

The previous section has focused on exponential smoothing processes assuming the single source of error framework. This section shows that, even in the context of multiple sources of errors, it is possible to derive the theoretical variance of standard (damped) trend exponential smoothing processes. First, algebraic results are provided for the damped trend exponential smoothing. Subsequently, results are given for the standard Holt-Winters methods.

Consider the following MSOE version of the damped trend exponential smoothing:

$$y_t=l_{t-1}+\phi b_{t-1}+\epsilon_t$$
$$l_t=l_{t-1}+\phi b_{t-1}+(1+\alpha)\xi_t \quad \text{... (11)}$$
$$b_t=\phi b_{t-1}+(1+\beta)\eta_t$$

This specification differs from (2) since the model is characterized by three different shocks. The reduced form of (11) can be written as:

$$\left(1-\phi B\right)\left(1-B\right)y_t=\epsilon_t-(1+\phi)\epsilon_{t-1}+\phi\epsilon_{t-2}+(1+\alpha)\left(\xi_{t-1}-\phi\xi_{t-2}\right)+\phi(1+\beta)\eta_{t-1} \quad \text{... (12)}$$

[3] An Eviews 7 program containing a simulation running the whole procedure can be provided by the author upon request. In addition, an Excel file computing the autocovariances as in (5) as well as the moving average parameters, given the error-correction parameters, is also available upon request.

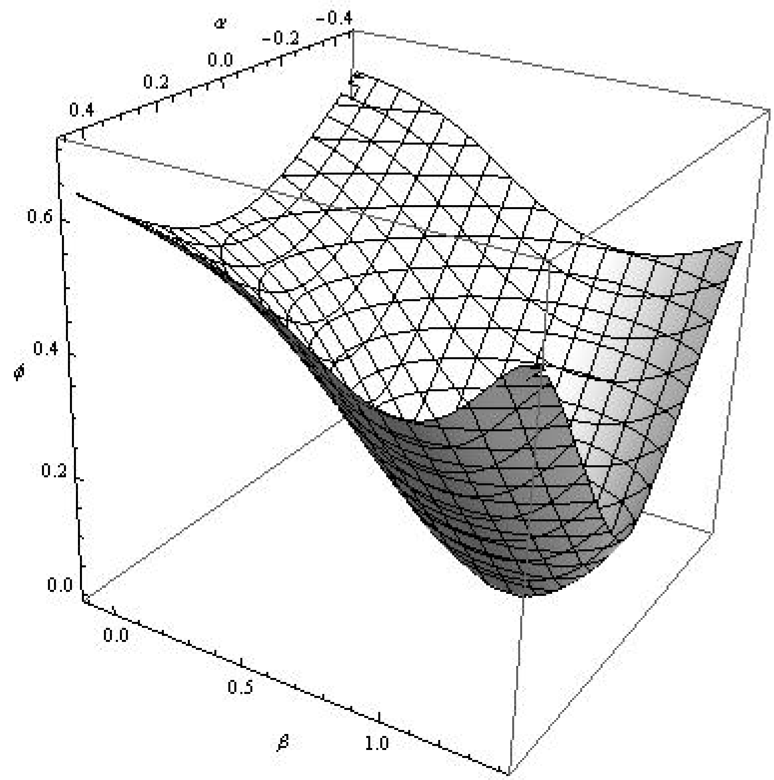

Figure 1. Region of parameters (ϕ; α^*;β^*) allowing $\sigma_\epsilon^2=\sigma_a^2$

having the following autocovariance functions:

$$\begin{aligned} \gamma_0 &= \left(1+(1+\phi)^2+\phi^2\right)\sigma_\epsilon^2+\left((1+\alpha)^2+\left(\phi(1+\alpha)\right)^2\right)\sigma_\xi^2+(\phi(1+\beta))^2\sigma_\eta^2 \\ \gamma_1 &= -(1+\phi)\sigma_\varepsilon^2-\phi(1+\phi)\sigma_\varepsilon^2-\phi(1+\alpha)^2\sigma_\xi^2 \\ \gamma_2 &= \phi\sigma_\epsilon^2 \\ \gamma_k &= 0 \text{ for } k\geq 3 \end{aligned} \qquad \text{... (13)}$$

Therefore:

$$\begin{aligned} \theta_2 &= \frac{\phi\sigma_\epsilon^2}{\sigma_a^2} \\ \theta_1 &= \frac{-(1+\phi)\sigma_\epsilon^2-\phi(1+\phi)\sigma_\epsilon^2-\phi(1+\alpha)^2\sigma_\xi^2}{\phi\sigma_\epsilon^2+\sigma_a^2} \end{aligned} \qquad \text{... (14)}$$

with:

$$\sigma_a^2=\frac{1}{4}[2\sigma_\epsilon^2+2\sigma_\epsilon^2\phi^2+\sigma_\eta^2\phi^2+2\sigma_\eta^2\beta^2\phi^2+\sigma_\eta^2\beta\phi^2+(1+\alpha)^2\left(1+\phi^2\right)\sigma_\xi^2+H+$$

$$\frac{1}{2}[-\sigma_\epsilon^4-6\sigma_\epsilon^4\phi^2-4\sigma_\epsilon^2\sigma_\xi^2\phi^2-8\alpha\sigma_\epsilon^2\sigma_\xi^2\phi^2-4\alpha^2\sigma_\epsilon^2\sigma_\xi^2\phi^2-\sigma_\epsilon^4\phi^4+\frac{1}{2}H$$

$$-\sigma_\xi^4\phi^2\left(1+4\alpha+6\alpha^2+4\alpha^3+\alpha^4\right)+2\sigma_\epsilon^2\sigma_\eta^2\phi^3\left(1+2\beta+2\beta^2\right)+M$$

$$+\frac{1}{2}M^2]^{\frac{1}{2}}] \qquad \text{... (15)}$$

and

$$H=[(1+\alpha)^2\sigma_\xi^2(-1+\phi)^2+(1+\beta)^2\sigma_\eta^2\phi^2]^{\frac{1}{2}}[(1+\beta)^2\sigma_\eta^2\phi^2+$$

$$4\sigma_\epsilon^2(1+\phi)^2+(1+\alpha)^2\sigma_\xi^2(1+\phi)^2]^{\frac{1}{2}} \qquad \text{... (16)}$$

$$M=(1+\beta)^2\sigma_\eta^2\phi^2+2\sigma_\epsilon^2(1+\phi^2)+(1+\alpha)^2\sigma_\xi^2(1+\phi^2) \qquad \text{... (17)}$$

Note that when ϕ=1 the damped trend (11) collapses to the standard Holt-Winters trend model whose reduced form is:

$$(1-B)^2y_t=\epsilon_t(1-B)^2+(1+\alpha)\left(\xi_{t-1}-\xi_{t-2}\right)+(1+\beta)\eta_{t-1} \qquad \text{... (18)}$$

With autocovariance functions:

$$\begin{aligned}\gamma_0&=6\sigma_\epsilon^2+2(1+\alpha)^2\sigma_\xi^2+(1+\beta)^2\sigma_\eta^2\\ \gamma_1&=-4\sigma_\epsilon^2-(1+\alpha)^2\sigma_\xi^2\\ \gamma_2&=\sigma_\epsilon^2\\ \gamma_k&=0 \text{ for } k\geq 3\end{aligned} \qquad \text{... (19)}$$

Such that:

$$\theta_2=\frac{\sigma_\epsilon^2}{\sigma_a^2}, \qquad \theta_1=\frac{-4\sigma_\epsilon^2-(1+\alpha)^2\sigma_\xi^2}{\sigma_\epsilon^2+\sigma_a^2} \qquad \text{... (20)}$$

with:

$$\sigma_a^2=\sigma_\epsilon^2+\frac{1}{2}\sigma_\xi^2(1+\alpha)^2+\frac{1}{4}\sigma_\eta(1+\beta)J+\frac{1}{4}\sigma_\eta^2(1+\beta)^2+\frac{\sqrt{2}}{4}[\sigma_\eta^4(1+\beta)^4+4(1+\alpha)^2(1+\beta)^2\sigma_\eta^2\sigma_\xi^2+$$

$$2(1+\alpha)^4\sigma_\xi^4+(1+\beta)^3\sigma_\eta^3J+2(1+\alpha)^2(1+\beta)\sigma_\eta\sigma_\xi^2J$$

$$+4\sigma_\epsilon^2\left(2(1+\alpha)^2\sigma_\xi^2+\sigma_\eta(1+\beta)\left(3\sigma_\eta(1+\beta)+J\right)\right)]^{\frac{1}{2}} \qquad \text{... (21)}$$

and

$$J=\sqrt{16\sigma_\epsilon^2+(1+\beta)^2\sigma_\eta^2+4(1+\alpha)^2\sigma_\xi^2} \qquad \text{... (22)}$$

Therefore, once the parameters of the error-correction form are obtained, the forecasting performance of the model can be computed instantaneously. To conclude, these results shed light on the forecasting properties of damped trend exponential smoothing models. More specifically, even in the MSOE context we have full control of the reduced form parameters and thus on the MMSE forecasts relative to the data generation process.

4. Conclusions

As recognized in Gardner (2006), the damped trend exponential smoothing has gained importance in empirical studies due to its remarkable forecasting performance. This paper derives its theoretical forecast error variance based on the implied ARIMA model. Algebraic results are provided for the standard Holt-Winters trend models, the damped trend exponential smoothing and for the random coefficient state space model. These algebraic results are relevant since they provide full control of the ARIMA reduced form parameters as exact functions of the structural parameters of the underlying models.

Acknowledgements

I wish to thank Jeyananthan Sivakumaran for his editorial assistance.

Bibliography

Armstrong, J. S. (2006) "Findings from evidence-based forecasting: Methods for reducing forecast error", *International Journal of Forecasting*,22,583-598.

Brown, R.G. (1963) *Smoothing, forecasting and prediction of discrete time series,* Englewood Cliffs, NJ: Prentice-Hall, Inc

Fildes, R., Nikolopous, K., Crone, S., and Syntetos, A. (2008) "Forecasting and operational research: a review", *Journal of Operational Research Society*,59,1-23

Gardner, E. S., Jr. (2006) "Exponential smoothing: The state of art-Part II", *International Journal of Forecasting*,22,637-666.

Gardner, E. S., Jr. and McKenzie, E. (1985) "Forecasting trends in time series", *Management Science*, 31,1237-1246

Gardner, E. S., Jr. and McKenzie, E. (1988) "Forecasting trends in time series", *Journal of the Operational Research Society*, 39,863-867

Gardner, E. S., Jr. and McKenzie, E. (1989) "Seasonal exponential smoothing with damped trends", *Management Science*, 35,372-376

Gardner, E. S., Jr. and McKenzie, E. (2011) "Why the damped trend works", *Journal of Operational Research Society*,62, 1177-1180

Hyndman, R., Koehler, A., Ord, J. K., and Snyder, R. D. (2008)*Forecasting with exponential smoothing: the state space approach*, Berlin: Springer-Verlag.

McKenzie, E. and Gardner, E. S., Jr. (2010) "Damped trend exponential smoothing: A modeling viewpoint", *International Journal of Forecasting*, 26 (4), 661-665

Morley, J. C., Nelson, C. and E. Zivot (2003) "Why are the Beveridge-Nelson and unobserved-components decomposition of GDP so different?"*Review of Economics & Statistics*, LXXXV, 2, 235-243

Sbrana, G. (2011), Forecasting Aggregated Moving Average Processes with an Application to the Euro-Area real interest rate, *Journal of Forecasting*, 31(1), 85-98

Sbrana, G. (2012), Structural time series models and aggregation: some analytical results, *Journal of Time Series Analysis*, 32(3), 315-316

APPENDIX

Let z_t be an invertible moving average processes of order two such that:

$z_t=e_t+\theta_1 e_{t-1}+\theta_2 e_{t-2}$

where e_t is white noise process. Considering the following autocovariance functions: Where $\gamma_0=E(z_t z_t)$, $\gamma_1=E(z_t z_{t-1})$ and $\gamma_2=E(z_t z_{t-2})$. The moving average parameters can be recovered solving the following system of three equations (autocorrelations):

$$\frac{\gamma_2}{\gamma_1}=\frac{\theta_2}{(1+\theta_1^2+\theta_2^2)}$$

$$\frac{\gamma_1}{\gamma_0}=\frac{\theta_1+\theta_1\theta_2}{(1+\theta_1^2+\theta_2^2)}$$

$$\sigma_e^2=\frac{\gamma_0}{(1+\theta_1^2+\theta_2^2)}$$

These equations represent respectively the second and first order autocorrelations of the process and the variance of e_t.

First, the following analytical solutions can be easily obtained:

$$\theta_2=\frac{\gamma_2}{\sigma_e^2}$$

$$\theta_1=\frac{\gamma_1}{(\sigma_e^2+\gamma_2)}$$

Secondly, substituting these solutions in the last equation of the system, the following quartic equation in x (with $x=\sigma_e^2$) can be obtained:

$$\frac{x^4+(2\gamma_2-\gamma_0)x^3+(2\gamma_2^2-2\gamma_2\gamma_0+\gamma_1^2)x^2+(2\gamma_2^3-\gamma_2^2\gamma_0)x+\gamma_2^4}{x(x+\gamma_2)^2}=0$$

This equation has four different solutions. Yet, the only solution leading to the invertible process[4] is:

$$\sigma_e^2=\frac{1}{4}\left(\gamma_0-2\gamma_2+G+\sqrt{2}\sqrt{\gamma_0^2+\gamma_0 G-2\left(\gamma_1^2+\gamma_2\left(2\gamma_2+G\right)\right)}\right) \quad \text{... (23)}$$

with:

$$G=\sqrt{(\gamma_0-2\gamma_1+2\gamma_2)(\gamma_0+2\gamma_1+2\gamma_2)} \quad \text{... (24)}$$

and:

$$\theta_2=\frac{4\gamma_2}{\left(\gamma_0-2\gamma_2+G+\sqrt{2}\sqrt{\gamma_0^2+\gamma_0 G-2\left(\gamma_1^2+\gamma_2\left(2\gamma_2+G\right)\right)}\right)}$$

$$\theta_1=\frac{4\gamma_1}{\left(\gamma_0+2\gamma_2+G+\sqrt{2}\sqrt{\gamma_0^2+\gamma_0 G-2\left(\gamma_1^2+\gamma_2\left(2\gamma_2+G\right)\right)}\right)} \quad \text{... (25)}$$

Similar results can be found in Sbrana (2011, 2012).

♦♦♦

[4] A process with roots of the characteristic function that lie outside the unit circle (i.e. $\theta_1+\theta_2<1$; $\theta_2-\theta_2<1$; $-1<\theta_2<1$