Bulletin of Applied Economics, 2022, 9(2), 1-13
https://doi.org/10.47260/bae/921

On the Heteroskedastic-Autoregressive Specification of the Linear Regression Model

Eftychia Lola* Spyridon D. Symeonides†

Abstract

In this paper we examine, from a theoretical viewpoint, the generalized normal linear regression model with disturbances that are simultaneously heteroskedastic and autoregressive. In particular, the error specification of the model is a mixture of Amemiya's linear heteroskedasticity structure with a stationary first-order autoregressive process. Given that the heteroskedastic variances are functions of the first-order autocorrelation coefficient, the estimators used in applied research cannot properly distinguish the estimations of the heteroskedastic and autoregressive parameters of the model. To avoid this problem, we introduce a multi-step estimation procedure, which has mainly theoretical interest, and is not suggested as an alternative to the well-known heteroskedasticity and autocorrelation consistent estimation used in applied econometric research. This estimation procedure facilitates the derivation of two distinct, theoretically important, generalized linear models, one with heteroskedastic and another with first-order autoregressive error terms. These two distinct models can be used for the theoretical examination of the finite-sample distributional properties of the estimators of the heteroskedastic and autoregressive parameters.

Key words: Linear regression model; autoregression; heteroskedasticity; consistent estimation.

J.E.L. classification: C01, C13, C22.

---

*Corresponding Author, Department of Economics, University of Ioannina. E-mail address: eftehialola@yahoo.gr. Part of this paper is a part of miss Lola's Ph.D. Thesis. The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the HFRI Ph.D. Fellowship grant (Fellowship Number: 645).

†Department of Economics, University of Ioannina, Ioannina 455 00, Greece. E-mail address: ssymeoni@uoi.gr

1 Introduction

Most of the single-equation econometric specifications in applied research can be expressed in the form of the generalized normal linear regression model, provided that certain assumptions are made about the structure of the disturbance covariance matrix. Some of the disturbance specifications, most frequently used in applied econometrics, are the AR(1), the heteroskedastic, and the seemingly-unrelated-regressions structures of the error terms. This fact explains the volume of the theoretical and applied work published in those areas. Also, due to the need to cope with more complex economic phenomena, in many cases, econometricians have brought into focus models with random errors which are generated by a mixture of various disturbance specifications, such as models of seemingly unrelated regressions with autocorrelated errors (see, e.g., Symeonides et al. (2016)), or models with mixed heteroskedastic-autoregressive disturbances, which can be estimated by using the heteroskedasticity consistent (HC) and the heteroskedasticity-autocorrelation consistent (HAC) estimators of the error covariance matrix (see, *inter alia*, White (1980), MacKinnon and White (1985), and Newey and West (1987)).

This paper is concerned with the theoretical investigation of the normal linear regression model in which the disturbances are specified as a mixed heteroskedastic-autoregressive process. In particular, we examine the mixture of a stationary first-order autoregressive process with autocorrelation coefficient ρ, and a linear heteroskedastic specification of the form $\mathrm{var}(u_t) = \boldsymbol{z}_{t\bullet}^{\top}\boldsymbol{\varsigma} \quad (t = 1, \ldots, T)$, where $\boldsymbol{\varsigma}$ is a vector of heteroskedasticity parameters (Amemiya 1977).

From the viewpoint of theoretical econometrics, a lot of effort has been devoted, up till now, to the construction of estimators of $\boldsymbol{\varsigma}$ and ρ in econometric models with error terms that are either heteroskedastic or autoregressive. Thus, in the linear model with heteroskedastic variances $\mathrm{var}(u_t) = \boldsymbol{z}_{t\bullet}^{\top}\boldsymbol{\varsigma}$ $(t = 1, \ldots, T)$, some of the most frequently used estimators of $\boldsymbol{\varsigma}$, described in Subsection 3.1, are the least squares or Goldfeld-Quandt estimator, the generalized least squares or Amemiya estimator, the iterative Amemiya estimator, and the maximum likelihood estimator. Moreover, in the linear model with AR(1) errors, some of the most frequently used estimators of ρ, described in Subsection 3.2, are the least squares estimator, the Durbin-Watson estimator, the generalized least squares estimator, the Prais-Winsten estimator, and the maximum likelihood estimator.

However, although there are many estimators of $\boldsymbol{\varsigma}$ and ρ in models with exclusively heteroskedastic or exclusively autoregressive disturbances, respectively, according to our knowledge, no procedure has ever been proposed for how the parameters $\boldsymbol{\varsigma}$ and ρ should be estimated in order to facilitate the theoretical investigation of linear models with a mixed heteroskedastic-autoregressive specification of disturbances. Our purpose, in this paper, is to derive such an estimation procedure. In order to do so, we introduce a reparameterization of the model, which can distinguish the two problems, i.e., heteroskedasticity and autocorrelation, simultaneously present in the error terms.

When a linear heteroskedastic specification is combined with a stationary first-order autoregressive process in order to generate the disturbances in a generalized normal linear regression model, the heteroskedastic variances, $\mathrm{var}(u_t) = \sigma_t^2/(1-\rho^2)$, are functions of the first-order autocorrelation coefficient, ρ. Because of this, the use of standard estimators results in estimated heteroskedasticity parameters which are functions of the first-order autocorrelation coefficient. This means that, although the parameters $\boldsymbol{\varsigma}$ and ρ are identified from a theoretical point of view, they cannot be properly distinguished by any of the estimators $\hat{\boldsymbol{\varsigma}}$ and $\hat{\rho}$ used in applied research.

To account for this problem, we introduce a reparameterization of the model, in which the heteroskedasticity parameter vector is $\boldsymbol{\varsigma}_* = \boldsymbol{\varsigma}(1-\rho^2)^{1/2}$. The use of this alternative parameterization results in a multi-step estimation procedure that enables us to effectively distinguish, from a theoretical viewpoint, the estimation of the heteroskedasticity parameters from the estimation of the first-order autocorrelation coefficient. Such a distinction is extremely useful whenever the researcher wants to construct an adjusted generalized linear model with disturbances that are exclusively heteroskedastic or exclusively autoregressive, in order to theoretically examine certain distributional properties of the estimators of $\boldsymbol{\varsigma}$ and ρ, respectively.

The structure of this paper is as follows: The model and some notational conventions are presented in Section 2. In particular, Subsection 2.1 describes the AR(1) component of the error specification, Subsection 2.2 describes the heteroskedastic component of the error specification, and Subsection 2.3 describes the mixed heteroskedastic-autoregressive specification of the error covariance matrix. Some commonly used estimators of $\boldsymbol{\varsigma}$ and ρ are given in Subsections 3.1, and 3.2, respectively. Section 4 presents an alternative model specification, which facilitates the derivation of the suggested estimation procedure in Section 5. Section 6 contains some concluding remarks.

2 The heteroskedastic-autoregressive Linear Model

The notation proposed by Abadir and Magnus (2002) is used throughout this paper, with minor modifications properly clarified. For any two indices i and j, δ_{ij} denotes Kronecker's delta; $\boldsymbol{L} = [(l_{ij})_{i=1,\ldots,n;\, j=1,\ldots,m}]$ denotes any $n \times m$ matrix with elements l_{ij}; and $\boldsymbol{l} = [(l_i)_{i=1,\ldots,n}]$, $\boldsymbol{l}^\top = [(l_i)_{i=1,\ldots,n}]^\top$ denote any $n \times 1$ and $1 \times n$ vectors, respectively, with elements l_i.

Consider the generalized linear regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}; \quad \mathrm{rk}(\boldsymbol{X}) = K; \quad \boldsymbol{u} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Omega}^{-1}), \tag{1}$$

where $\boldsymbol{y}$ is a $T \times 1$ vector of observations on the endogenous variable, $\boldsymbol{X}$ is a $T \times K$ full column-rank matrix of observations on K non-stochastic regressors, and $\boldsymbol{\beta}$ is a $K \times 1$ vector of unknown structural parameters. Moreover, $\boldsymbol{u} = [(u_t)_{t=1,\ldots,T}]$ is a $T \times 1$ vector of normally distributed unobserved stochastic disturbances, and $\boldsymbol{\Omega}$ is a $T \times T$ symmetric, positive definite matrix. The assumption of normally distributed disturbances is used in many theoretical econometric papers (see, e.g., Turkington (2000)).

For the disturbances to be both heteroskedastic and autoregressive, the t th element of vector $\boldsymbol{u}$ must satisfy the relationship

$$u_t = \sigma_t u_{(\mathrm{AR})t} \quad \forall\, t = 1, \ldots, T, \tag{2}$$

where $\sigma_1, \ldots, \sigma_T$ are positive scalars, uncorrelated with the autoregressive elements of the $T \times 1$ random vector $\boldsymbol{u}_{\mathrm{AR}} = [(u_{(\mathrm{AR})t})_{t=1,\ldots,T}]$.

2.1 The autoregressive specification

The elements of vector $\boldsymbol{u}_{\mathrm{AR}}$ are generated by the stationary AR(1) process

$$u_{(\mathrm{AR})t} = \rho u_{(\mathrm{AR})t-1} + \varepsilon_t; \quad |\rho| < 1 \quad (t = 2, \ldots, T), \tag{3}$$

where $u_{(\mathrm{AR})t} \sim \mathrm{N}(0, 1/(1-\rho^2))$, and ε_t are i.i.d. $\mathrm{N}(0,1)$ random errors $\forall\, t = 1, \ldots, T$. The time-series $u_{(\mathrm{AR})t}$ $(t = 1, \ldots, T)$ is a stationary AR(1) process, provided that $u_{(\mathrm{AR})1} = (1-\rho^2)^{-1/2}\varepsilon_1$. Therefore,

$$\mathrm{E}(\boldsymbol{u}_{\mathrm{AR}}) = \mathbf{0}_T, \quad \text{and} \quad \mathrm{var}(\boldsymbol{u}_{\mathrm{AR}}) = \mathrm{E}(\boldsymbol{u}_{\mathrm{AR}}\boldsymbol{u}_{\mathrm{AR}}^{\top}) = (1-\rho^2)^{-1}\boldsymbol{R}, \tag{4}$$

where $\boldsymbol{R}$ is the following $T \times T$ symmetric, positive definite matrix:

$$\boldsymbol{R} = \left[\left(\rho^{|t-t'|}\right)_{t,t'=1,\ldots,T}\right]. \tag{5}$$

2.2 The heteroskedastic specification

Let $\boldsymbol{z}_{t\cdot}^{\top} = [(z_{tj})_{j=1,\ldots,m}]^{\top}$ be the t th row of the $T \times m$ full column-rank matrix $\boldsymbol{Z}$, where $z_{t1} \equiv 1\ \forall t = 1, \ldots, T$. Matrix $\boldsymbol{Z}$ contains the observations on m non-stochastic variables, some of which may be regressors too. Also, let

$$\boldsymbol{\varsigma} = [(\varsigma_j)_{j=1,\ldots,m}] \in \mathbb{R}^m \setminus \{\mathbf{0}_m\}, \tag{6}$$

be a $m \times 1$ non-zero vector of unknown heteroskedasticity parameters. Then, following Amemiya (1977), we assume that the standard deviations σ_t in equation (2) are generated by the linear functions

$$\sigma_t^2 = \boldsymbol{z}_{t\cdot}^{\top}\boldsymbol{\varsigma} \quad \forall\, t = 1, \ldots, T. \tag{7}$$

We define the $T \times T$ diagonal matrices

$$\boldsymbol{\Sigma}^{1/2} = \mathrm{diag}(\sigma_1, \ldots, \sigma_T) \quad \text{and} \quad \boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_T^2) = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}. \tag{8}$$

2.3 Mixed autoregressive-heteroskedastic specification of $\boldsymbol{\Omega}$

The elements of the $T \times T$ matrix $\boldsymbol{\Omega}$ are functions of the $(m+1) \times 1$ vector

$$\boldsymbol{\gamma} = (\rho, \boldsymbol{\varsigma}^\top)^\top, \tag{9}$$

where ρ is the autocorrelation coefficient in the AR(1) process (3) and $\boldsymbol{\varsigma}$ is the vector of coefficients in the linear heteroskedasticity model (7). Then, equations (2), (3), (4), and (5) imply that

$$\mathrm{E}(u_t) = 0, \;\; \mathrm{var}(u_t) = \sigma_t^2/(1-\rho^2), \;\; \mathrm{cov}(u_t, u_{t'}) = \sigma_t \sigma_{t'} \rho^{|t-t'|}/(1-\rho^2). \tag{10}$$

Hence, the $T \times T$ matrix $\boldsymbol{\Omega}^{-1}$ can be written as

$$\boldsymbol{\Omega}^{-1} = \left[\left(\sigma_t \sigma_{t'} \rho^{|t-t'|}/(1-\rho^2) \right)_{t,t'=1,\ldots,T} \right], \tag{11}$$

and using equations (5), (8), and (11), we can write

$$\boldsymbol{\Omega}^{-1} = \boldsymbol{\Sigma}^{1/2} [(1-\rho^2)^{-1} \boldsymbol{R}] \boldsymbol{\Sigma}^{1/2}. \tag{12}$$

Let $\boldsymbol{D}$ be a $T \times T$ band matrix whose (t, t') th element is 1 if $|t - t'| = 1$ and 0 elsewhere. Also, let $\boldsymbol{\Delta}$ be a $T \times T$ matrix with 1 in the $(1, 1)$ st and (T, T) th positions and 0's elsewhere. Then, the inverse of matrix $(1-\rho^2)^{-1}\boldsymbol{R}$ can be written as

$$[(1-\rho^2)^{-1} \boldsymbol{R}]^{-1} = (1+\rho^2) \boldsymbol{I}_T - \rho \boldsymbol{D} - \rho^2 \boldsymbol{\Delta}. \tag{13}$$

Equations (12) and (13) imply that

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1/2} [(1+\rho^2) \boldsymbol{I}_T - \rho \boldsymbol{D} - \rho^2 \boldsymbol{\Delta}] \boldsymbol{\Sigma}^{-1/2}, \tag{14}$$

where

$$\boldsymbol{\Sigma}^{-1/2} = \mathrm{diag}(1/\sigma_1, \ldots, 1/\sigma_T). \tag{15}$$

3 Asymptotically efficient estimators of $\boldsymbol{\gamma} = (\rho, \boldsymbol{\varsigma}^\top)^\top$

From the various econometric estimators of $\boldsymbol{\varsigma}$ and ρ, the most frequently used in applied research are reported in the following subsections.

3.1 Estimators of $\boldsymbol{\varsigma} = (\varsigma_1, \ldots, \varsigma_m)^\top$

Let $\boldsymbol{x}_{t\boldsymbol{\cdot}}^\top$ be the t th row of the $T \times K$ matrix $\boldsymbol{X}$. Then, model (1) can be written as

$$y_t = \boldsymbol{x}_{t\boldsymbol{\cdot}}^\top \boldsymbol{\beta} + u_t, \;\; \text{where } \mathrm{E}(u_t) = 0 \text{ and } \mathrm{var}(u_t) = \boldsymbol{z}_{t\boldsymbol{\cdot}}^\top \boldsymbol{\varsigma} \;\; \forall\, t = 1, \ldots, T. \tag{16}$$

The most frequently used estimators of $\boldsymbol{\varsigma}$ in (16) are given below:

(i) The least squares or Goldfeld-Quandt (GQ) estimator (Goldfeld and Quandt 1965)

$$\hat{\boldsymbol{\varsigma}}_{\mathrm{GQ}} = \left[\sum_{t=1}^{T} \boldsymbol{z}_{t\boldsymbol{\cdot}} \boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top}\right]^{-1} \sum_{t=1}^{T} \boldsymbol{z}_{t\boldsymbol{\cdot}} (y_t - \boldsymbol{x}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\beta}}_{\mathrm{LS}})^2, \tag{17}$$

where $\hat{\boldsymbol{\beta}}_{\mathrm{LS}}$ is the least squares estimator of $\boldsymbol{\beta}$.

(ii) The generalized least squares or Amemiya (A) estimator (Amemiya 1977)

$$\hat{\boldsymbol{\varsigma}}_{\mathrm{A}} = \left[\sum_{t=1}^{T} (\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\varsigma}}_{\mathrm{GQ}})^{-2} \boldsymbol{z}_{t\boldsymbol{\cdot}} \boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top}\right]^{-1} \sum_{t=1}^{T} (\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\varsigma}}_{\mathrm{GQ}})^{-2} \boldsymbol{z}_{t\boldsymbol{\cdot}} (y_t - \boldsymbol{x}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\beta}}_{\mathrm{LS}})^2, \tag{18}$$

where $\hat{\boldsymbol{\beta}}_{\mathrm{LS}}$ is the least squares estimator of $\boldsymbol{\beta}$.

(iii) The iterative Amemiya (IA) estimator

$$\hat{\boldsymbol{\varsigma}}_{i} = \left[\sum_{t=1}^{T} (\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\varsigma}}_{i-1})^{-2} \boldsymbol{z}_{t\boldsymbol{\cdot}} \boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top}\right]^{-1} \sum_{t=1}^{T} (\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\varsigma}}_{i-1})^{-2} \boldsymbol{z}_{t\boldsymbol{\cdot}} (y_t - \boldsymbol{x}_{t\boldsymbol{\cdot}}^{\top} \hat{\boldsymbol{\beta}}_{i-1})^2, \tag{19}$$

where index $i = 2, 3, \ldots$ denotes the iteration number, and $\hat{\boldsymbol{\varsigma}}_{i-1}$, $\hat{\boldsymbol{\beta}}_{i-1}$ are the estimator of $\boldsymbol{\varsigma}$ and the feasible GL estimator of $\boldsymbol{\beta}$ from the previous iteration. For $i = 1$, the estimator of $\boldsymbol{\varsigma}$ from the first iteration is $\hat{\boldsymbol{\varsigma}}_1 = \hat{\boldsymbol{\varsigma}}_{\mathrm{A}}$.

(iv) The maximum likelihood (ML) estimator, $\hat{\boldsymbol{\varsigma}}_{\mathrm{ML}}$, which maximizes the log-likelihood function

$$\ell(\boldsymbol{\beta}, \boldsymbol{\varsigma}) = -1/2 \sum_{t=1}^{T} \ln(\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \boldsymbol{\varsigma}) - 1/2 \sum_{t=1}^{T} (y_t - \boldsymbol{x}_{t\boldsymbol{\cdot}}^{\top} \boldsymbol{\beta})^2 / (\boldsymbol{z}_{t\boldsymbol{\cdot}}^{\top} \boldsymbol{\varsigma}). \tag{20}$$

3.2 Estimators of ρ

In a linear regression model with autoregressive disturbances, the most frequently used estimators of ρ are given below:

(i) The least squares (LS) estimator

$$\hat{\rho}_{\mathrm{LS}} = \sum_{t=2}^{T} \hat{u}_{t(\mathrm{LS})} \hat{u}_{t-1(\mathrm{LS})} \Big/ \sum_{t=1}^{T} \hat{u}_{t(\mathrm{LS})}^2, \tag{21}$$

where $\hat{u}_{t(\mathrm{LS})}$ are the least squares residuals.

(ii) The Durbin-Watson (DW) estimator

$$\hat{\rho}_{\text{DW}} = 1 - DW/2, \tag{22}$$

where DW is the Durbin-Watson statistic (Durbin and Watson 1950, 1951).

(iii) The generalized least squares (GL) estimator

$$\hat{\rho}_{\text{GL}} = \sum_{t=2}^{T} \hat{u}_{t(\text{GL})}\hat{u}_{t-1(\text{GL})} / \sum_{t=1}^{T} \hat{u}^2_{t(\text{GL})}, \tag{23}$$

where $\hat{u}_{t(\text{GL})}$ are the generalized least squares residuals.

(iv) The Prais-Winsten estimator, $\hat{\rho}_{\text{PW}}$, which, together with $\hat{\beta}_{\text{PW}}$, minimizes the sum of squared GL residuals (Prais and Winsten 1954).

(v) The maximum likelihood (ML) estimator, $\hat{\rho}_{\text{ML}}$, which satisfies a cubic equation with coefficients defined in terms of the ML residuals in the regression model (Beach and MacKinnon 1978).

3.3 Comment on the estimation of $\boldsymbol{\varsigma}$

Given that $\text{var}(u_t) = \sigma_t^2/(1-\rho^2)$, any estimator $\hat{\boldsymbol{\varsigma}}$, of the heteroskedasticity parameters, is a function of the first-order autocorrelation coefficient ρ. Because of this, the parameters $\boldsymbol{\varsigma}$ and ρ cannot be readily distinguished by any standard single-step estimation procedure, which is based upon the estimators $\hat{\boldsymbol{\varsigma}}$ and $\hat{\rho}$ reported in Subsections 3.1 and 3.2, respectively.

The fact that the heteroskedastic variances $\text{var}(u_t) = \sigma_t^2/(1-\rho^2)$ are functions of ρ, is accounted for by the alternative model specification introduced in Section 4, which results in a very useful reparameterization of the model.

4 Alternative model specification

Let $\sigma^2_{u_t}$ and $\sigma_{u_t u_{t'}}$ $(t, t' = 1, \ldots, T)$ denote the variances and covariances, respectively, of the disturbances in (1). Then, equation (10) implies that

$$\sigma^2_{u_t} = \sigma_t^2/(1-\rho^2) \implies \sigma_{u_t} = \sigma_t/(1-\rho^2)^{1/2}, \tag{24}$$

and equations (10) and (24) imply that

$$\sigma_{u_t u_{t'}} = \sigma_t \sigma_{t'} \rho^{|t-t'|}/(1-\rho^2) = \sigma_{u_t}\sigma_{u_{t'}}\rho^{|t-t'|}. \tag{25}$$

Notice that, by incorporating the factor $1/(1-\rho^2)^{1/2}$ into the standard deviations σ_{u_t}, equations (24) and (25) express the heteroskedastic variances and covariances explicitly as functions of the autocorrelation coefficient ρ. This results in the reparameterized heteroskedasticity parameter vector $\boldsymbol{\varsigma}_*$ proposed in the following subsection.

4.1 Alternative heteroskedastic specification

Define the $m \times 1$ non-zero vector

$$\boldsymbol{\varsigma}_* = (1-\rho^2)^{-1}\boldsymbol{\varsigma} = [(\varsigma_{*j})_{j=1,\dots,m}] \in \mathbb{R}^m \setminus \{\mathbf{0}_m\}, \tag{26}$$

with elements $\varsigma_{*j} = (1-\rho^2)^{-1}\varsigma_j$ $(j = 1,\dots,m)$. Then, by combining (7), (24) and (26) we find that

$$\sigma^2_{u_t} = \left(\boldsymbol{z}_{t\bullet}^\top\boldsymbol{\varsigma}\right)/(1-\rho^2) = \boldsymbol{z}_{t\bullet}^\top\left[(1-\rho^2)^{-1}\boldsymbol{\varsigma}\right] = \boldsymbol{z}_{t\bullet}^\top\boldsymbol{\varsigma}_* \quad \forall\, t = 1,\dots,T. \tag{27}$$

Dividing the t th observation, $y_t = \boldsymbol{x}_{t\bullet}^\top\boldsymbol{\beta}+u_t$, of model (1) by $\sigma_{u_t} = (\boldsymbol{z}_{t\bullet}^\top\boldsymbol{\varsigma}_*)^{1/2}$, we find the alternative autoregressive specification given in the following subsection.

4.2 Alternative autoregressive specification

Define the $T \times 1$ random vector $\boldsymbol{u}_* = [(u_{*t})_{t=1,\dots,T}]$ with elements

$$u_{*t} = u_t/\sigma_{u_t}. \tag{28}$$

Then, equations (2), (10), (24), and (28) imply that

$$u_{*t} = (1-\rho^2)^{1/2}u_{(\mathrm{AR})t}. \tag{29}$$

Then, $u_{(\mathrm{AR})t} \sim \mathrm{N}(0, 1/(1-\rho^2)) \implies u_{*t} \sim \mathrm{N}(0,1)\ \forall\, t$. Therefore,

$$\mathrm{E}(\boldsymbol{u}_*) = \mathbf{0}_T \quad \text{and} \quad \mathrm{var}(\boldsymbol{u}_*) = \mathrm{E}(\boldsymbol{u}_*\boldsymbol{u}_*^\top) = \boldsymbol{R}. \tag{30}$$

Equations (5), (10), (29) and (30) imply that

$$\mathrm{E}(u_{*t}) = 0, \;\; \mathrm{var}(u_{*t}) = 1, \;\; \mathrm{cov}(u_{*t}u_{*t'}) = \rho^{|t-t'|}. \tag{31}$$

Finally, equations (3), (28), and (29) imply that

$$\begin{aligned} u_{*t} &= (1-\rho^2)^{1/2}\left(\rho u_{(\mathrm{AR})t-1} + \varepsilon_t\right) \\ &= \rho\left[(1-\rho^2)^{1/2}u_{(\mathrm{AR})t-1}\right] + (1-\rho^2)^{1/2}\varepsilon_t \\ &= \rho u_{*t-1} + \varepsilon_{*t}, \end{aligned} \tag{32}$$

where $\varepsilon_{*t} = (1-\rho^2)^{1/2}\varepsilon_t$ are i.i.d. $\mathrm{N}(0,(1-\rho^2))$ random errors $\forall t = 1,\dots,T$. Equation (32) implies that the elements of the random vector $\boldsymbol{u}_*$ are generated by a stationary AR(1) process with autocorrelation coefficient ρ.

Further, by combining (24) and (27) we find that

$$\sigma_t^2 = \sigma^2_{u_t}(1-\rho^2) = \boldsymbol{z}_{t\bullet}^\top\boldsymbol{\varsigma}_*(1-\rho^2) \;\; \forall\, t = 1,\dots,T. \tag{33}$$

Also, equations (6) and (26) imply that

$$\boldsymbol{\varsigma} = \boldsymbol{\varsigma}_*(1-\rho^2) \implies \varsigma_i = \varsigma_{*i}(1-\rho^2) \;\; (i = 1,\dots,m). \tag{34}$$

4.3 Alternative representation of $\boldsymbol{\Omega}$

Equations (11), (24), and (25) imply that

$$\boldsymbol{\Omega}^{-1} = \left[\left(\sigma_{u_t} \sigma_{u_{t'}} \rho^{|t-t'|} \right)_{t,t'=1,\dots,T} \right]. \tag{35}$$

Define the $T \times T$ diagonal matrices $\boldsymbol{\Sigma}_*^{1/2} = \mathrm{diag}(\sigma_{u_1}, \dots, \sigma_{u_T})$ and $\boldsymbol{\Sigma}_*^{-1/2} = \mathrm{diag}(1/\sigma_{u_1}, \dots, 1/\sigma_{u_T})$. Then, equation (24) implies that

$$\boldsymbol{\Sigma}_*^{1/2} = (1-\rho^2)^{-1/2}\, \mathrm{diag}(\sigma_1, \dots, \sigma_T) = (1-\rho^2)^{-1/2} \boldsymbol{\Sigma}^{1/2}, \tag{36}$$

and

$$\boldsymbol{\Sigma}_*^{-1/2} = (1-\rho^2)^{1/2} \boldsymbol{\Sigma}^{-1/2}. \tag{37}$$

From equations (12) and (36) we take

$$\boldsymbol{\Omega}^{-1} = \left[(1-\rho^2)^{-1/2} \boldsymbol{\Sigma}^{1/2} \right] \boldsymbol{R} \left[(1-\rho^2)^{-1/2} \boldsymbol{\Sigma}^{1/2} \right] = \boldsymbol{\Sigma}_*^{1/2} \boldsymbol{R} \boldsymbol{\Sigma}_*^{1/2}, \tag{38}$$

and form equations (14) and (37) we take

$$\begin{aligned} \boldsymbol{\Omega} &= \left[(1-\rho^2)^{-1/2} \boldsymbol{\Sigma}_*^{-1/2} \right] \left[(1+\rho^2) \boldsymbol{I}_T - \rho \boldsymbol{D} - \rho^2 \boldsymbol{\Delta} \right] \left[(1-\rho^2)^{-1/2} \boldsymbol{\Sigma}_*^{-1/2} \right] \\ &= (1-\rho^2)^{-1} \boldsymbol{\Sigma}_*^{-1/2} [(1+\rho^2) \boldsymbol{I}_T - \rho \boldsymbol{D} - \rho^2 \boldsymbol{\Delta}] \boldsymbol{\Sigma}_*^{-1/2}. \end{aligned} \tag{39}$$

5 Suggested estimation procedure

Let LS, GL, IG, and ML denote the least squares, generalized least squares, iterative generalized least squares, and maximum likelihood estimators, respectively. Also, let $\hat{\boldsymbol{\beta}}_{\mathrm{I}}$ denote any consistent estimator of $\boldsymbol{\beta}$ indexed by $\mathrm{I} = \mathrm{LS}, \mathrm{GL}, \mathrm{IG}, \mathrm{ML}$. The discussion in Sections 2 and 4 suggests the following 7-step procedure for the estimation of model (1):

Step 1: Estimate model (1) using the $\hat{\boldsymbol{\beta}}_{\mathrm{I}}$ estimator. Then, the corresponding residual vector:

$$\hat{\boldsymbol{u}}_{\mathrm{I}} = \boldsymbol{y} - \hat{\boldsymbol{\beta}}_{\mathrm{I}} \boldsymbol{X} = \left[(\hat{u}_{t(\mathrm{I})})_{t=1,\dots,T} \right] \tag{40}$$

is a consistent predictor of the disturbance vector $\boldsymbol{u}$.

Step 2: Use one of the consistent estimators given in Subsection 3.1 in order to estimate the parameter vector $\boldsymbol{\varsigma}_*$. Then, estimate matrix $\boldsymbol{\Sigma}_*^{-1/2}$ as

$$\hat{\boldsymbol{\Sigma}}_*^{-1/2} = \mathrm{diag}(1/\hat{\sigma}_{u_1}, \dots, 1/\hat{\sigma}_{u_T}), \tag{41}$$

where

$$\hat{\sigma}_{u_t} = \left(\boldsymbol{z}_{t\bullet}^{\top} \hat{\boldsymbol{\varsigma}}_* \right)^{1/2} \quad \forall\, t = 1, \dots, T. \tag{42}$$

Step 3: Estimate the heteroskedasticity-corrected residuals

$$\hat{\boldsymbol{u}}_{*\,\mathrm{I}} = \hat{\boldsymbol{\Sigma}}_*^{-1/2}\hat{\boldsymbol{u}}_{\mathrm{I}} = \left[(\hat{u}_{*t(\mathrm{I})})_{t=1,\ldots,T}\right], \tag{43}$$

where

$$\hat{u}_{*t(\mathrm{I})} = \frac{\hat{u}_{t(\mathrm{I})}}{\hat{\sigma}_{u_t}} \quad \forall\, t = 1, \ldots, T, \tag{44}$$

and $\hat{\boldsymbol{u}}_{\mathrm{I}}$ is the predictor of $\boldsymbol{u}$ estimated by (40).

Step 4: Use one of the consistent estimators given in Subsection 3.2 in order to calculate an initial estimate $\hat{\rho}_*$ of the autocorrelation coefficient ρ.

Step 5: Use (34) and the consistent estimators $\hat{\boldsymbol{\varsigma}}_*$ and $\hat{\rho}_*$ in order to estimate the parameter vector $\boldsymbol{\varsigma}$ as

$$\hat{\boldsymbol{\varsigma}} = \hat{\boldsymbol{\varsigma}}_*(1 - \hat{\rho}_*^2) \implies \hat{\varsigma}_i = \hat{\varsigma}_{*i}(1 - \hat{\rho}_*^2) \quad \forall\, i = 1, \ldots, m. \tag{45}$$

Then, estimate matrix $\boldsymbol{\Sigma}^{-1/2}$ as

$$\hat{\boldsymbol{\Sigma}}^{-1/2} = \mathrm{diag}(1/\hat{\sigma}_1, \ldots, 1/\hat{\sigma}_T), \tag{46}$$

where

$$\hat{\sigma}_t = \left(\boldsymbol{z}_{t\bullet}^\top \hat{\boldsymbol{\varsigma}}\right)^{1/2} \quad \forall\, t = 1, \ldots, T. \tag{47}$$

Alternatively, $\boldsymbol{\varsigma}$ can be estimated via the following asymptotically equivalent process:

(i) Use the initial estimator $\hat{\rho}_*$ in order to transform model (1) into the autoregression-corrected model

$$\boldsymbol{y}_{\mathrm{H}} = \boldsymbol{X}_{\mathrm{H}}\boldsymbol{\beta} + \boldsymbol{u}_{\mathrm{H}}, \tag{48}$$

where the elements of vector $\boldsymbol{u}_{\mathrm{H}} = \left[(u_{(\mathrm{H})t})_{t=1,\ldots,T}\right]$ are purely heteroskedastic disturbances, given by the following formulae:

$$u_{(\mathrm{H})1} = (1 - \hat{\rho}_*^2)^{1/2} u_1, \quad u_{(\mathrm{H})t} = u_t - \hat{\rho}_* u_{t-1} \quad \forall\, t = 2, \ldots, T. \tag{49}$$

(ii) use one of the consistent estimators given in Subsection 3.1 in order to estimate the parameter vector $\boldsymbol{\varsigma}$, and then estimate matrix $\boldsymbol{\Sigma}^{-1/2}$ via (46) and (47).

Although from the estimation viewpoint (45) is perfectly adequate as a consistent estimator of $\boldsymbol{\varsigma}$, the estimator $\hat{\boldsymbol{\varsigma}}$ based on the residuals of model (48) enables the researcher to find the finite-sample distributional properties of any consistent estimator of $\boldsymbol{\varsigma}$ in Subsection 3.1.

Step 6: Premultiply model (1) by $\hat{\boldsymbol{\Sigma}}^{-1/2}$ given in (46), in order to derive heteroskedasticity-corrected model

$$\boldsymbol{y}_{\mathrm{AR}} = \boldsymbol{X}_{\mathrm{AR}}\boldsymbol{\beta} + \boldsymbol{u}_{\mathrm{AR}}, \tag{50}$$

where the elements of vector $\boldsymbol{u}_{\mathrm{AR}} = \left[(u_{(\mathrm{AR})t})_{t=1,\dots,T}\right]$ are purely autoregressive disturbances, given by the following formula:

$$u_{(\mathrm{AR})t} = u_t/\hat{\sigma}_t \quad \forall\, t = 1, \dots, T, \tag{51}$$

where $\hat{\sigma}_t$ are given in (47). Then, use one of the consistent estimators given in Subsection 3.2 in order to estimate the autocorrelation coefficient ρ. The estimator $\hat{\rho}$ based on the residuals of model (50) enables the researcher to find the finite-sample distributional properties of any consistent estimator of ρ in Subsection 3.2.

Step 7: Use the estimators $\hat{\boldsymbol{\Sigma}}^{-1/2}$ and $\hat{\rho}$ from Steps 5 and 6, respectively, in order to calculate the estimator

$$\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Sigma}}^{-1/2}[(1+\hat{\rho}^2)\boldsymbol{I}_T - \hat{\rho}\boldsymbol{D} - \hat{\rho}^2\boldsymbol{\Delta}]\hat{\boldsymbol{\Sigma}}^{-1/2}, \tag{52}$$

which can be used for the feasible generalized least squares estimation of model (1).

6 Concluding remarks

In this paper, we examined the normal linear regression model with a disturbance specification, which is the combination of the Amemiya's (1977) linear heteroskedasticity structure with a time-series of stationary first-order autoregressive error terms.

Since the heteroskedastic variances, $\mathrm{var}(u_t) = \sigma_t^2/(1-\rho^2)$, are functions of the first-order autocorrelation coefficient ρ, the estimators $\hat{\boldsymbol{\varsigma}}$ and $\hat{\rho}$ used in applied research cannot properly distinguish the estimation of the parameters $\boldsymbol{\varsigma}$ and ρ. To cope with this problem, we introduced an alternative parameterization of the model, from which we derived a multi-step estimation procedure that distinguishes the information conveyed by the estimators $\hat{\boldsymbol{\varsigma}}$ and $\hat{\rho}$. Needless to note that this estimation procedure is not proposed as an alternative to the heteroskedasticity-autocorrelation consistent estimation in applied econometric research, but it is introduced for theoretical purposes only.

The suggested estimation procedure enables the researcher to come up with two distinct adjusted generalized linear models of theoretical importance in econometrics, one of which has heteroskedastic disturbances, and another that has first-order autoregressive disturbances. These two distinct models can be used for the theoretical examination of some crucial finite-sample distributional properties of the estimators $\hat{\varsigma}$ and $\hat{\rho}$. These distributional properties can, in turn, be used for the calculation of finite-sample Edgeworth and Cornish-Fisher size corrected t and F tests in the linear model with heteroskedastic-autoregressive disturbances. This is an interesting topic for further research.

Given the attention paid, in applied research, to the avoidance of erroneous tests in small samples, much endeavor has been devoted to the development of size corrected econometric tests, which are implemented by using either the Edgeworth-corrected critical values, or the corresponding Cornish-Fisher corrected test statistics (see, *inter alia*, Cornish and Fisher (1937), Fisher and Cornish (1960), Hill and Davis (1968), Rothenberg (1984)). Such size corrected t and F tests have already been proposed for the linear model with AR(1) errors (Magdalinos and Symeonides 1995), for the linear model with heteroskedastic errors (Symeonides et al. 2007), and for the S.U.R. model with autocorrelated errors (Symeonides et al. 2016).

References

Abadir, K. M. and Magnus, J. R. (2002). Notation in Econometrics: A Proposal for A Standard. *Econometrics Journal*, 5:76–90.

Amemiya, T. (1977). A Note on a Heteroscedastic Model. *Journal of Econometrics*, 6:365–370.

Beach, C. M. and MacKinnon, J. G. (1978). A Maximum Likelihood Procedure for Regression with Autocorrelated Errors. *Econometrica*, 46(1):51–58.

Cornish, E. A. and Fisher, R. A. (1937). Moments and Cumulants in The Specification of Distributions. *Revue de l' Institute International de Statistic*, 4:1–14.

Durbin, J. and Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression: I. *Biometrika*, 37(3):409–428.

Durbin, J. and Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression: II. *Biometrika*, 38(1):159–177.

Fisher, R. A. and Cornish, E. A. (1960). The Percentile Points of Distributions Having Known Cumulants. *Technometrics*, 2(2):209–225.

Goldfeld, S. M. and Quandt, R. E. (1965). Some Tests for Homoscedasticity. *Journal of the American Statistical Association*, 60(310):539–547.

Hill, G. W. and Davis, A. W. (1968). Generalized Asymptotic Expansions of Cornish-Fisher Type. *The Annals of Mathematical Statistics*, 39(4):1264–1273.

MacKinnon, J. G. and White, H. (1985). Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics*, 29:305–325.

Magdalinos, M. A. and Symeonides, S. D. (1995). Alternative Size Corrections for Some GLS Test Statistics: The Case of the AR(1) Model. *Journal of Econometrics*, 66(1–2):35–59.

Newey, W. K. and West, K. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55:703–708.

Prais, S. J. and Winsten, C. B. (1954). Trend Estimators and Serial Correlation. Cowles Commission Discussion Paper: Statistics No. 383.

Rothenberg, T. J. (1984). Hypothesis Testing in Linear Models when the Error Covariance Matrix is Nonscalar. *Econometrica*, 52(4):827–842.

Symeonides, S. D., Tzavalis, E., and Kandilorou, H. (2007). Cornish-Fisher Size Corrected t and F Statistics for the Linear Regression Model with Heteroskedastic Errors. In Philips, G. D. A. and Tzavalis, E., editors, *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, pages 173–204. Cambridge University Press, Cambridge.

Symeonides, S. D., Tzavalis, E., and Karavias, Y. (2016). Size Corrected Significance Tests in Seemingly Unrelated Regressions with Autocorrelated Errors. *Journal of Time Series Econometrics*, 9(1):1–40.

Turkington, D. A. (2000). Generalised Vec Operators and the Seemingly Unrelated Regression Equations Model with Vector Correlated Disturbances. *Journal of Econometrics*, 99(2):225–53.

White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48(4):817–838.