Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 141 (2007) 383–415
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Generalized R-estimators under conditional heteroscedasticity

Kanchan Mukherjee

Department of Mathematical Sciences, The University of Liverpool, Liverpool L69 7ZL, UK

Available online 20 November 2006

Abstract

In this paper, we extend the classical idea of Rank estimation of parameters from homoscedastic problems to heteroscedastic problems. In particular, we define a class of rank estimators of the parameters associated with the conditional mean function of an autoregressive model through a three-steps procedure and then derive their asymptotic distributions. The class of models considered includes Engel's ARCH model and the threshold heteroscedastic model. The class of estimators includes an extension of Wilcoxon-type rank estimator. The derivation of the asymptotic distributions depends on the uniform approximation of a randomly weighted empirical process by a perturbed empirical process through a very general weight-dependent partitioning argument.

JEL classification: C14; C22

Keywords: Rank estimation; Heteroscedastic model; Weighted empirical process; Uniform approximation

1. Introduction

Since the introduction of the autoregressive conditional heteroscedastic (ARCH) time series model of Engle (1982), there have been huge developments on the theory and application of this model and its various generalizations to economics and finance. ARCH models have been used to represent the volatility, i.e., the strong dependence of the instantaneous variability of a time series on its own past, in numerous economic and financial data sets. For a literature review, see Bollerslev et al. (1992), Shephard (1996), and Gouriéroux (1997), among others. Most of the existing methodological literature have

E-mail address: k.mukherjee@liverpool.ac.uk.

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2006.10.002

focused on developing estimation procedures for the parameters associated with the conditional variability using pseudo-likelihood methods. However, development of the estimation methods associated with the conditional mean component of a heteroscedastic problem is also important from the application point of view and this has been largely overlooked. In this paper, we aim to fill that gap by developing a rank-based robust procedure for estimating the mean parameter of an autoregressive model with conditional heteroscedastic errors.

In a parametric formulation, linearity of regression, independence and normality of errors, homoscedasticity or form of heteroscedasticity, etc. are typically assumed for drawing conclusions about parameters of interest. However, there is no guarantee that such regularity assumptions will be valid in a given situation and therefore it is natural to investigate alternative procedures that can perform well under probable departures from model assumptions. Among different types of such robust procedures, estimators based on ranks or the so-called R-estimators are sometimes preferable to their other competitors for their global robustness property as they generally demand much less restrictive assumptions on the underlying distributions; see, for example, Jurečková and Sen (1996, Section 3.4) for a discussion on this. The need for using such robust estimators is even more for financial data due to the empirical finding that 'outliers' appear more often in asset returns than that implied by white noises having normal distribution. For more on this, see Tsay (2002, Section 3.3) and Engle and Gonzalez-Rivera (1991) who quantified the loss of efficiency resulting from the use of estimators arising from the first-order conditions for the normal MLE (called the quasi-maximum likelihood estimator or the QMLE) on non-normal distributions and concluded that 'it is worthwhile searching for estimators that can improve on QMLE'.

There is a vast literature on the R-estimation of parameters in homoscedastic regression and autoregression models. For a glimpse, see Koul (1992, Section 4.4), Jurečková and Sen (1996, Section 3.4, Chapter 6) and Hájek et al. (1999, Section 10.3), among others. In linear regression model with i.i.d. or homoscedastic long memory errors, R-estimators are known to have highly desirable efficiency; see, e.g., Jurečková (1971), Koul (1971), Jaeckel (1972), and Koul and Mukherjee (1993). In the homoscedastic autoregressive time series model (1.1) with $\sigma \equiv 1$, analogs of the R-estimators are known to have similar efficiency and robustness properties as investigated by Koul and Ossiander (1994) and Mukherjee and Bai (2002). It is thus natural to investigate their behavior in the heteroscedastic set up.

Accordingly, consider the following autoregressive model with heteroscedastic error where for known integers s, p, and r, $\{X_i,\ 1-s \leqslant i \leqslant n\}$ is an observable time series. Set $\boldsymbol{W}_{i-1} := (X_{i-1}, X_{i-2}, \ldots, X_{i-s})'$ and $\boldsymbol{Y}_{i-1} = c(\boldsymbol{W}_{i-1})$, $1 \leqslant i \leqslant n$, where $c : \mathbb{R}^s \to \mathbb{R}^p$ is a known function. Let Ω_j, $j = 1, 2$, be open subsets of $\mathbb{R}^p$, $\mathbb{R}^r$, respectively with $\Omega := \Omega_1 \times \Omega_2 \subset \mathbb{R}^m$, where $m = p + r$. Let σ be a known function from $\mathbb{R}^p \times \Omega_2$ to $\mathbb{R}^+ := (0, \infty)$, differentiable in its second argument. Consider the model

$$X_i = \boldsymbol{Y}_{i-1}'\boldsymbol{\alpha} + \sigma(\boldsymbol{Y}_{i-1}, \boldsymbol{\beta})\eta_i, \quad 1 \leqslant i \leqslant n, \tag{1.1}$$

where $\boldsymbol{\alpha} \in \Omega_1$, $\boldsymbol{\beta} \in \Omega_2$ are the unknown parameters, and the unobservable errors $\{\eta_i,\ i \geqslant 1\}$ are i.i.d. with zero mean and finite variance having a distribution function (d.f.) G and probability density function (p.d.f.) g. Throughout, we also assume that $\{\eta_i,\ i \geqslant 1\}$ are independent of $\boldsymbol{W}_0 := (X_0, X_{-1}, \ldots, X_{1-s})'$ and hence independent of $\boldsymbol{Y}_0$; for each $y \in \mathbb{R}^p$, $\dot{\sigma}(y, \boldsymbol{t})$ is the derivative of $\sigma(y, \boldsymbol{t})$ with respect to $\boldsymbol{t}$; and $\{X_i\}$ is strictly stationary and ergodic. All of these assumptions will be referred to as the model assumptions in the sequel.