Economics Letters 150 (2017) 130–134

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Robustness of binary choice models to conditional heteroscedasticity

Tim Ginker, Offer Lieberman *

Bar-Ilan University, Israel

HIGHLIGHTS

- Probit/Logit MLE with misspecified conditional heteroscedasticity is analyzed.
- The MLE's probability limit is a positive scalar multiple of the true parameter.
- Bounds are given on the scalar.
- Predictions based on the estimator are unaffected by the misspecification.
- t-tests can be used inspite of the misspecification.

ARTICLE INFO

Article history:
Received 25 September 2016
Received in revised form 17 November 2016
Accepted 18 November 2016
Available online 24 November 2016

JEL classification:
C22

Keywords:
Conditional heteroscedasticity
Misspecified models
Probit

ABSTRACT

We show that when the true data generating process of a large class of binary choice models contains conditional heteroscedasticity, predictions based on the misspecified MLE in which conditional heteroscedasticity is ignored, are unaffected by the misspecification.

1. Introduction

In recent years qualitative response models have become popular in time series analysis. In the financial context, in particular, they have been applied in connection with the sign of asset return forecasts, as these may lead to profitable speculative positions and correct hedging decisions. See, among others, Levich (2001). While, *apriori*, this class of models seems to be less informative than continuous response models, Leung et al.'s (2000) detailed comparative study between binary and continuous response models revealed that the former outperform the latter in its ability to generate trading profits.

Sign forecastability has been assumed to be driven mainly by conditional mean dynamics of the underlying process in most recent studies on this issue. For instance, Nyberg (2011) examined the ability of the binary dependent dynamic probit model to predict the direction of monthly excess stock returns, extending Kauppi and Saikkonen's (2008) model. He concluded that in terms of out-of-sample performance, binary models can be useful in asset allocation decisions, especially when the mean dynamics resemble an error correction specification.

Binary choice models were shown to be useful in the context of decomposition—type models. For instance, Anatolyev and Gospodinov (2010) expressed the financial asset return as a product of its sign and its absolute value. The two components were modeled separately as a copula before a joint forecast was constructed. Earlier, Rydberg and Shephard (2003) specified the stock return as a product of two binary variables, defining the returns direction and market activity, and multiplied by a process which defines the size of a price change.

Previous findings on modeling conditional heteroscedasticity in qualitative response models showed that the volatility parameters are statistically significant and may have a good explanatory power. Specifically, Deuker (1999) modeled the discrete changes in the bank prime lending rate by a dynamic ordered probit with Markov-switching conditional heteroscedasticity. His results indicated that conditional heteroscedasticity plays an important role in explaining the data. Broseta (2000) reported a good fit for

* Correspondence to: Department of Economics and Research Institute for Econometrics (RIE), Bar-Ilan University, Ramat Gan 52900, Israel.
E-mail address: offer.lieberman@gmail.com (O. Lieberman).

http://dx.doi.org/10.1016/j.econlet.2016.11.024
0165-1765/

a learning model in which the latent residuals were allowed to follow an ARCH(1) process (Engle, 1982). Hausman and Lo (1992) estimated a model for stock price changes with heteroscedastic ordered probit by dividing the price changes into eight intervals.

In the same line of literature, Christoffersen and Diebold (2006) suggested that the volatility and other high-order conditional moments may produce sign dependence, drawing the theoretical connection between the asset return volatility dynamics and its sign forecastability. From their study it follows that directional forecasts could be inferred from the volatility dynamics even when the conditional mean is constant. There are numerous other studies that have documented the strong dependence of asset returns volatility. For surveys of the empirical evidence and volatility modeling in finance, the reader is referred to Mikosch et al. (2009) and Bollerslev et al. (1992).

To this end, we consider in this paper the model

$$y_t = 1\{x_t'\gamma + \varepsilon_t > 0\},\ t = 1, \ldots, n, \tag{1}$$

where $1\{\cdot\}$ is the indicator function which takes the value of unity if the condition in the brackets is satisfied and zero, otherwise, x_t is a $K \times 1$ vector of explanatory variables which are assumed to be ergodic stationary, γ is a $K \times 1$ vector of unknown parameters, and for all t and s, conditional on $\mathcal{F}_{t-1}$ and x_s, $\varepsilon_t \sim F(0, \sigma_t^2)$, where $\mathcal{F}_t$ is the increasing sequence of σ-fields generated by $\left\{\varepsilon_j\right\}_{j=1}^{t}$ and F is a symmetric CDF on which conditions are given in Assumption A below. The conditional variance, σ_t^2, is merely assumed to satisfy some very mild regularity conditions so that the class of heteroscedasticity models allowed is very general and includes in it, as a special case, the prominent GARCH(p, q) specification (Bollerslev, 1986). For this model, under the classical assumptions including a fixed σ_t^2, $\forall t$, the main workhorse for estimating this model is undoubtedly the probit maximum likelihood estimator (MLE), if F is normal, or by Logit, if F is logistic, although other alternatives exist, such as Horowitz's (1993) semiparametric estimator. It is well known that under these restrictive assumptions (i.e., which include homoscedasticity of ε_t), the MLE is consistent and asymptotically efficient. However, when the true data generating process follows (1) but is misspecified to have homoscedastic ε_t, the MLE will no longer be consistent. See for instance, Greene (2012, p. 693) and Yatchew and Griliches (1985)—the latter developed an approximation for the probit MLE bias in the presence of a simple heteroscedasticity form in a cross sectional context. We show in this paper that this misspecification will result in a positive scaling effect on the asymptotic mean of the MLE. This form of inconsistency under the general setting has not been known hitherto. The implication is that, surprisingly, the MLE—based predictions will be unaffected by the misspecification. In other words, even if conditional heteroscedasticity of a general form will be ignored and the model will be estimated by the MLE, the predictions based on the (wrong) estimator will be unaffected. This result is of importance and practical relevance because it has been widely acknowledged that the volatility of asset returns varies across time.

Our main Theorems corroborate some of the simulation results reported by Munizaga et al. (2000), which revealed the remarkable robustness of the misspecified Probit and Logit model—based predictions to conditional heteroscedasticity. Moreover, we show that t-tests can be based on the MLE with reference to the standard normal distribution in spite of the misspecification.

The main results of the paper are given in the following Section. Simulations are reported in Section 3 and final remarks are provided in Section 4.

2. Main results

By $\mathcal{F}_{t-1}$ we denote the σ-field generated by $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$. We shall make the following assumption.

Assumption A.

(1) The data generating function is given by $y_t = 1\{x_t'\gamma + \varepsilon_t > 0\}$.
(2) For all t and s, conditional on $\mathcal{F}_{t-1}$ and x_s, ε_t has a zero mean, conditional variance σ_t^2, $0 < \sigma_t < \infty$ and cumulative distribution function (CDF) F. The CDF F is smooth and strictly monotonic with a bounded density f which has $\mathbb{R}$ as its support and which is symmetric. In addition, $F(\nu)$ is concave for $\nu > 0$.
(3) The true parameter vector γ_0 is an element of the interior of a convex parameter space $\Gamma \subset \mathbb{R}^K$.
(4) The $K \times 1$ vector x_t is finite, strictly stationary and ergodic, and is not contained in any linear subspace of $\mathbb{R}^K$, $\forall t$.
(5) The process $\{\sigma_t\}$ is strictly stationary and ergodic and independent of x_s, for all t and s.

We remark that Assumption A(2) holds for the normal and logistic distributions. The misspecified log-likelihood function in which conditional heteroscedasticity is ignored and σ_t is set to unity $\forall t$, is given by

$$\tilde{l}_n(\gamma) = \sum_{t=1}^{n} \tilde{l}_t(\gamma),$$

where

$$\tilde{l}_t(\gamma) = y_t \log\left(F^*(x_t'\gamma)\right) + (1 - y_t)\log\left(1 - F^*(x_t'\gamma)\right) \tag{2}$$

and F^* is the CDF of a random variable with a CDF F after it has been normalized to have mean zero and unit variance. Similarly, by f^* we denote the density corresponding to F^*. Let $\tilde{\gamma}_n = \arg\max_\Gamma \tilde{l}_n(\gamma)$. To emphasize, $\tilde{\gamma}_n$ is the maximizer of a misspecified log-likelihood function. By E_{γ_0} we denote an expectation taken under the true parameter value. The main result follows below.

Theorem 1. *Under Assumption A, there exists a finite and positive ρ which satisfies*

$$0 < \frac{1}{E_{\gamma_0}(\sigma_t)} \leq \rho \leq E_{\gamma_0}\left(\frac{1}{\sigma_t}\right) < \infty, \tag{3}$$

such that $\tilde{\gamma}_n \xrightarrow{p} \rho\gamma_0$.

The MLE of the correct likelihood,[1] say $\hat{\gamma}_n$, is consistent. If, given $x_t = x$, the researcher wishes to base the predictive value, $\hat{y}_t$, of y_t, according to the rule $\hat{y}_t = 1\{F^*(x'\hat{\gamma}_n/\sigma_t) > 0.5\}$, then for large n, the rule is equivalent to $1\left\{x'\gamma_0 > 0\right\}$. Basing the prediction on $\tilde{\gamma}_n$ instead does not affect the result because for large n it is tantamount to

$$1\{\rho x'\gamma_0 > 0\} = 1\left\{x'\gamma_0 > 0\right\}.$$

In other words, the misspecified MLE-based prediction remains unaltered even though $\tilde{\gamma}_n$ is inconsistent. This result corroborates some of the simulation results of Munizaga et al. (2000).

When the classical assumptions including homoscedasticity hold and the usual normalization, $\sigma_t = 1\ \forall t$, is imposed, it follows from (3) that $\rho = 1$, i.e., that the MLE based on the correct specification is consistent. Therefore, Theorem 1 is a generalization of the standard result.

Proof of Theorem 1. The proof can be made by verifying the conditions of Theorem 2.7 of Newey and McFadden (1994), the difference being that instead of the true parameter γ_0, we will show

[1] By 'correct likelihood' it is meant that, among other things, σ_t is correctly specified. In general, it would be a function of a finite dimensional vector of parameters, as in the GARCH(p, q) process, for instance, and its parameters would have had to be estimated jointly with γ.