Conditional Heteroskedasticity in some Common Count Data Models for Financial Time Series Data*

Kurt Brännäs
Department of Economics & USBE, Umeå University
SE-901 87 Umeå, Sweden
email: kurt.brannas@econ.umu.se

Abstract

Conditional heteroskedasticity properties are derived for some common count data regression and time series models. New extensions are suggested and discussed.

Key Words: Conditional variance, time series, finance, traded stocks, Poisson.

JEL Code: C25, G12, G14

Umeå Economic Studies 592, 2002

1. Introduction

This note studies the conditional variance or heteroskedasticity properties of some common count data models for time series data and discusses some new extensions. Recent financial research applies Poisson or other count data models to the number of traded stocks (e.g., Gourieroux and Jasiak, 2001, ch. 14). As conditional heteroskedasticity is an important ingredient in other time series models for financial markets, the presence of this property in count data therefore appears of potential interest.

Count data models typically have a heteroskedasticity property (e.g., Cameron and Trivedi, 1998), which automatically implies conditional heteroskedasticity. This is in contrast to most continuous variable models for, e.g., the stock price, in which no heteroskedasticity is assumed but conditional heteroskedasticity is a feature of great interest (e.g., Engle, 1982).

The starting point in this note is the Poisson regression model and we mainly consider off-springs of this model. Hence, directly specified semiparametric models (e.g., Fahrmeier and Tutz, 1994, ch. 6) to be estimated by, e.g., GMM are not considered.

We start by studying existing count data models in Section 2. In Section 3 we study ways to expand these specifications to accommodate conditional heteroskedasticity in alternative ways. The final section contains a more general discussion.

2. Models

The basic model for most count data regression modelling is the Poisson model. The Poisson distribution has the property of independent increments which implies that for a count variable y_t at time t

$$\mathrm{E}(y_t) = \mathrm{E}(y_t|F_{t-1}) = \mathrm{V}(y_t) = \mathrm{V}(y_t|F_{t-1}) = \lambda_t,$$

where $F_{t-1} = (Y_{t-1}, X_t)$ is the information set with $Y_t = (y_1, \ldots, y_t)$ and $X_t = (\mathbf{x}_1, \ldots, \mathbf{x}_t)$. Typically,

$$\lambda_t = \exp(\mathbf{x}_t \boldsymbol{\beta}),$$

where $\mathbf{x}_t$ is the vector of exogenous variables and $\boldsymbol{\beta}$ is a vector of parameters.

Hence, in this basic model the unconditional and conditional heteroskedasticities are identical. In addition, the means and variances are equal. This is then a very restrictive specification with respect to financial applications as well as for other time series data.

2.1 Overdispersed Poisson

A common feature of empirical count data is that the variance exceeds the mean. This is usually modelled in terms of an overdispersed Poisson model. Here, y_t is Poisson distributed conditionally on a latent random variable ε_t so that

$$\mathrm{E}(y_t|\varepsilon_t) = \mathrm{V}(y_t|\varepsilon_t) = \varepsilon_t \lambda_t.$$

*Research financially supported by the Wallander-Hedelius Foundation and USBE. Jörgen Hellström is thanked for comments on a previous version of the paper.

Conventionally one assumes $\{\varepsilon_t\}$ to be an iid sequence with $\mathrm{E}(\varepsilon_t) = 1$ and $\mathrm{V}(\varepsilon_t) = \sigma^2$.

The conditional and unconditional moments are internally equal, i.e.

$$\mathrm{E}(y_t) = \mathrm{E}(y_t|F_{t-1}) = \lambda_t$$
$$\mathrm{V}(y_t) = \mathrm{V}(y_t|F_{t-1}) = \lambda_t + \sigma^2\lambda_t^2,$$

but the means and variances are no longer equal.

When ε_t is assumed gamma distributed the unconditional y_t has a negative binomial, NB2, distribution. To estimate, either such a fully parametric model may be estimated by ML or, e.g., a GMM estimator based on only the given moments may be applied.

Even if the variances increase quadratically with respect to the mean level this is a quite restrictive specification.

2.2 Zeger's Model

Zeger (1988) suggested an extension of the overdispersed Poisson model for time series data. One sets $\mathrm{E}(y_t|\varepsilon_t) = \mathrm{V}(y_t|\varepsilon_t) = \varepsilon_t\lambda_t$ and assumes the stationary $\{\varepsilon_t\}$ sequence to again have $\mathrm{E}(\varepsilon_t) = 1$ and $\mathrm{V}(\varepsilon_t) = \sigma^2$. Besides implying overdispersion this model gives serially correlated counts. However, conditionally on ε_t and ε_s, respectively, y_t and y_s are independent. The unconditional mean and variance are those of the previous model. One may also obtain the time-varying autocovariance function of the $\{y_t\}$ sequence.

To derive the conditional moments we need the following result:

$$\mathrm{E}(y_t^k|F_{t-1}) = \mathrm{E}_{\varepsilon_t}\left[\mathrm{E}(y_t^k|\varepsilon_t)|F_{t-1}\right],$$

which holds for conventional time series models for ε_t. The result holds since

$$\begin{aligned}
\mathrm{E}(y_t^k|F_{t-1}) &= \sum_{y=0}^{\infty} y_t^k \frac{\mathrm{Pr}(y_t, F_{t-1})}{\mathrm{Pr}(F_{t-1})} \\
&= \sum_{y=0}^{\infty} y_t^k \frac{\int_0^\infty \mathrm{Pr}(y_t, \varepsilon_t, F_{t-1})\ d\varepsilon_t}{\mathrm{Pr}(F_{t-1})} \\
&= \int_0^\infty \sum_{y=0}^{\infty} y_t^k\, \mathrm{Pr}(y_t|\varepsilon_t) f(\varepsilon_t|F_{t-1})\ d\varepsilon_t \\
&= \int_0^\infty \mathrm{E}(y_t^k|\varepsilon_t) f(\varepsilon_t|F_{t-1})\ d\boldsymbol{\varepsilon}_t \\
&= \mathrm{E}_{\varepsilon_t}\left[\mathrm{E}(y_t^k|\varepsilon_t)|F_{t-1}\right].
\end{aligned}$$

It follows then that

$$\begin{aligned}
\mathrm{E}(y_t|F_{t-1}) &= \lambda_t \mathrm{E}(\varepsilon_t|F_{t-1}) \\
\mathrm{V}(y_t|F_{t-1}) &= \mathrm{E}_{\varepsilon_t}\left[\mathrm{E}(y_t^2|\varepsilon_t)|F_{t-1}\right] \\
&\quad -\mathrm{E}^2_{\varepsilon_t}\left[\mathrm{E}(y_t|\varepsilon_t)|F_{t-1}\right] \\
&= \lambda_t \mathrm{E}(\varepsilon_t|F_{t-1}) + \lambda_t^2 \mathrm{V}(\varepsilon_t|F_{t-1}).
\end{aligned}$$

Consider as an example the AR(1) model $\varepsilon_t = \theta\varepsilon_{t-1} + (1-\theta) + u_t$, where $\{u_t\}$ is a zero mean random sequence with variance σ_u^2. The parametrization is such that $\mathrm{E}(\varepsilon_t) = 1$. Then $\mathrm{E}(\varepsilon_t|F_{t-1}) = \theta\varepsilon_{t-1} + (1-\theta)$ and $\mathrm{V}(\varepsilon_t|F_{t-1}) = \mathrm{V}(\varepsilon_t) = \sigma_u^2$. Hence

$$\begin{aligned}
\mathrm{E}(y_t|F_{t-1}) &= [\theta\varepsilon_{t-1} + (1-\theta)]\,\lambda_t \\
\mathrm{V}(y_t|F_{t-1}) &= [\theta\varepsilon_{t-1} + (1-\theta)]\,\lambda_t + \sigma_u^2\lambda_t^2.
\end{aligned}$$

In this case the conditional mean is affected in the same way as the conditional variance is.

Consider as another example the MA(1) model $\varepsilon_t = u_t + \theta u_{t-1}$, where $\{u_t\}$ is a random sequence with mean $1/(1+\theta)$ and variance σ_u^2. Then $\mathrm{E}(\varepsilon_t|F_{t-1}) = 1/(1+\theta) + \theta u_{t-1}$ and $\mathrm{V}(\varepsilon_t|F_{t-1}) = \sigma_u^2$. Hence

$$\begin{aligned}
\mathrm{E}(y_t|F_{t-1}) &= [1/(1+\theta) + \theta u_{t-1}]\,\lambda_t \\
\mathrm{V}(y_t|F_{t-1}) &= [1/(1+\theta) + \theta u_{t-1}]\,\lambda_t + \sigma_u^2\lambda_t^2.
\end{aligned}$$

It can also be shown that these results hold when y_t is not only conditional on ε_t but on F_{t-1} as well.

The Zeger specification is still restrictive in the sense of not allowing for a less tight relationship between the conditional mean and variance.

2.3 The Zeger-Qaqish Model

The Zeger and Qaqish (1988) model contains lagged $y_{t-i}, i > 0$, in the λ_t function and specifies a conditional model for y_t given past observations. This approach can be extended by introducing an ε_t as in either of the two previous subsections. It is quite straightforward to demonstrate that no changes to the conditional moments of these subsections will arise. The only exception is the presence of lagged endogenous variables in λ_t.

3. Modified Models

We consider two types of modifications of the basic models in Section 2. First, we redefine σ^2 to become time dependent and possibly dependent on previous observations. Second, we alter the basic conditional expression.

Consider the overdispersed Poisson model (Section 2.1) and let all assumptions used above remain true, but let the variance of ε_t be a function of past observations, i.e. $\mathrm{V}(\varepsilon_t) = \sigma_t^2(F_{t-1})$.

This time dependence will not imply dependence between successive counts nor will it affect the conditional and unconditional means. However, the conditional variance changes into

$$\mathrm{V}(y_t|F_{t-1}) = \lambda_t + \sigma_t^2(F_{t-1})\lambda_t^2.$$

This then adds flexibility for the model specification, but suitable specifications of $\sigma_t^2(F_{t-1})$ need to be considered. To guarantee that σ_t^2 remains positive an exponential form appears reasonable. Corresponding to an EGARCH(1,1) specification we could specify, say,

$$\sigma_t^2 = \exp\left(\alpha_0 + \alpha_1 \ln \sigma_{t-1}^2 + \alpha_2 u_{t-1}^2\right),$$

where $u_t = y_t - \lambda_t$ is an error term. Given this specification GMM estimation or some type of two-stage estimator of the α_i parameters are feasible. Alternatively with ε_t gamma distributed y_t follows a NB2 distribution and then ML estimation is feasible. Within these estimation frameworks LM-type tests against added conditional heteroskedasticity (i.e. $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$) can be constructed.

Corresponding results hold for the more general Zeger or Zeger and Qaqish models.

If we wish to have identical conditional and unconditional means but with a more variable conditional heteroskedasticity we could also start with

$$\mathrm{E}(y_t|\varepsilon_t, h_t) = \mathrm{V}(y_t|\varepsilon_t, h_t) = \lambda_t + (\varepsilon_t - 1)h_t\lambda_t,$$

where $\{\varepsilon_t\}$ is an iid sequence with unit mean and variance. Then h_t is the conditional standard deviation of ε_t and could, e.g., depend on past observations.

For this model

$$\begin{aligned} \mathrm{E}(y_t) &= \mathrm{E}(y_t|F_{t-1}) = \lambda_t \\ \mathrm{V}(y_t) &= \lambda_t + \mathrm{E}(h_t^2)\lambda_t^2 \\ \mathrm{V}(y_t|F_{t-1}) &= \lambda_t + h_t^2\lambda_t^2. \end{aligned}$$

An obvious drawback with this type of model arises from the requirement that $\lambda_t + (\varepsilon_t - 1)h_t\lambda_t \geq 0$. This is of importance when $\varepsilon_t < 1$. If, for example, $\varepsilon_t = 0$ then $h_t < 1$ must hold.

Approximately, the same moment properties can be obtained from the conditional representation $\lambda_t \exp(\varepsilon_t h_t)$. If $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{V}(\varepsilon_t) = 1$ and $\varepsilon_t h_t$ is small, a first order Taylor expansion gives $\exp(\varepsilon_t h_t) \approx 1 + h_t\varepsilon_t$. Then $\mathrm{E}(\exp(\varepsilon_t h_t)) \approx 1$ and $\mathrm{V}(\exp(\varepsilon_t h_t)) \approx 1 + h_t^2$. For this specification only size restrictions are involved on $\varepsilon_t h_t$. Note that a conditional specification $\lambda_t\varepsilon_t h_t$, which appears closer to the continuous variable specification, would with $\mathrm{E}(\varepsilon_t) = 1$ result in a model where it would be difficult to separate mean and variance effects.

We could obviously also express the model on a form closer to the mainstream conditional heteroskedasticity literature. By using $y_t = \mathrm{E}(y_t) + u_t$, where $\mathrm{E}(u_t) = 0$ and $\mathrm{V}(u_t) = \lambda_t$, we get results corresponding to the Poisson model. If we set $u_t = \varepsilon_t\lambda_t$ with $\mathrm{E}(\varepsilon_t) = 0$ and $\mathrm{V}(\varepsilon_t) = \sigma_t^2(F_{t-1})$ we get $\mathrm{V}(y_t|F_{t-1}) = \lambda_t + \sigma_t^2(F_{t-1})\lambda_t^2$. Distributionally this route is far from easy.

While in both this and the extended, overdispersed Poisson model the resulting conditional variances are related, the actual data generating process for the latter makes it a more appealing approach.

4. Discussion

In the mainstream literature on conditional heteroskedasticity the mean function is not affected. In restricted versions there is no conditional heteroskedasticity. The exception, M-ARCH, contains conditional heteroskedasticity as an explanatory variable in the mean function. By contrast all count data models studied above (and other ones as well) always contain conditional heteroskedasticity. In widely used count data models (Sections 2.1-2.3) there are obviously close relationships between conditional mean and variance functions. Attempts to relax these ties imply technical difficulties in terms of size restrictions on conditional variance functions. The extension of the overdispersed Poisson model appears the most reasonable modelling approach.

Another class of models to consider is the integer-valued ARMA or INARMA (e.g., McKenzie, 1986). Brännäs and Hall (2000) gave conditional variance results for a few alternative INMA models. Their INMA(1)-Model 1 has the conditional variance

$$\mathrm{V}(y_t|F_{t-1}) = \sigma^2 + \theta(1-\theta)\varepsilon_{t-1},$$

where $\mathrm{V}(\varepsilon_t) = \sigma^2$ and $\theta \in (0,1]$. Brännäs and Hellström (2001) gave results for generalizations of the basic INAR(1) model. The standard INAR(1) model has conditional variance

$$\mathrm{V}(y_t|F_{t-1}) = \alpha(1-\alpha)y_{t-1} + \sigma^2,$$

where $\mathrm{V}(\varepsilon_t) = \sigma^2$ and $\alpha \in [0,1]$. Obviously, this class can be viewed as an alternative to the extended, overdispersed Poisson model.

When it comes to estimation it is a general result that models (the λ_t part) can be estimated consistently by the Poisson ML estimator (Gourieroux, Monfort, Trognon, 1984) even if there is added conditional heteroskedasticity. One would also expect this pseudo-ML estimator to remain efficient (cf. Brännäs and Johansson, 1996). When parameters characterizing the conditional heteroskedasticity are of interest GMM estimation appears a reasonable approach. In fact, there will be no loss in efficiency even if these parameters are estimated separately in a second stage (Ahn and Schmidt, 1995, Brännäs and Johansson, 1996).

References

Ahn, S.G. and Schmidt, P. (1995). A Separability Result for GMM Estimation with Applications to GLS Prediction and Conditional Moment Tests. *Econometric Reviews* **14**, 19-34.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applied Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K. and Hellström, J. (2001). Generalized Integer-Valued Autoregression. *Econometric Reviews* **20**, 425-443

Brännäs, K. and Johansson, P. (1996). Panel Data Regression for Counts. *Statistical Papers* **37**, 191-213.

Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.

Engle, R. (1982). Autoregressive Heteroskedasticity with Estimates of the Variance of UK Inflation. *Econometrica* **50**, 987-1008.

Fahrmeier, L. and Tutz, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer-Verlag, New York.

Gourieroux, C. and Jasiak, J. (2001). *Financial Econometrics: Problems, Models, and Methods*. Princeton University Press, Princeton.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica* **52**, 701-720.

McKenzie, E. (1986). Autoregressive Moving-Average Processes with Negative-Binomial and Geometric Marginal Distributions. *Advances in Applied Probability* **18**, 679-705.

Zeger, S. (1988). A Regression Model for Time Series of Counts. *Biometrika* **75**, 621-629.

Zeger, S. and Qaqish, B. (1988). Markov Regression Models for Time Series: A Quasi-Likelihood Approach. *Biometrics* **44**, 1019-1031.