

# Conditional Heteroskedasticity in Count Data Regression: Self-Feeding Activity in Fish

Kurt Brännäs<sup>a</sup> and Eva Brännäs<sup>b</sup>

<sup>a</sup> Department of Economics & USBE, Umeå University

<sup>b</sup> Department of Aquaculture, Swedish University of Agricultural Sciences

**Abstract:** The paper introduces a new approach to incorporating time dependent overdispersion for Poisson related regression models. To handle the added flexibility in conditional heteroskedasticity in time series count data some wellknown estimators are adapted and a GMM type estimator is suggested. The estimators are applied to a time series of self-feeding activity in Arctic charr. There is strong support for both a dynamic conditional mean function and a dynamic model for the overdispersion.

**JEL:** C13, C22, C25, C25, C51.

**Key words:** Poisson, Overdispersion, ARCH, Estimation, Self-Feeding, Arctic Charr.

*Umeå Economic Studies* 595, 2002

## 1. INTRODUCTION

The paper introduces a time dependent overdispersion or a conditional heteroskedasticity specification for count data regression models. Empirically we detect overdispersion when an empirical variance exceeds a variance implied by a count data model.

In previous count data regressions (e.g., Cameron & Trivedi, 1998) overdispersion (e.g., Johnson & Kotz, 1969, ch. 5) and overdispersion with serial correlation in the count data sequence (e.g., Zeger, 1988) have been introduced by conditioning the count variable on a random variable or on a random stationary sequence. In these cases the count variable  $y_t$  at time  $t$  given the random variable  $\varepsilon_t$  follows a Poisson distribution (the conventional base case) with parameter  $\lambda_t \varepsilon_t \geq 0$ , where  $\lambda_t$  may be a function of a vector of explanatory variables  $\mathbf{x}_t$  and unknown parameters in a vector  $\boldsymbol{\beta}$ . By assuming that  $E(\varepsilon_t) = 1$ ,  $V(\varepsilon_t) = \sigma^2$  and that  $Corr(\varepsilon_t, \varepsilon_{t-k}) = \rho_k, k = 1, 2, \dots$ , are the autocorrelations of the stationary sequence  $\{\varepsilon_t\}$ , the unconditional first and second order moments of  $y_t$  can be obtained. With these, generalized method of moments (GMM, e.g., Hansen, 1982; Brännäs

& Johansson, 1994) or other moment estimators (e.g., Zeger, 1988) can be obtained without a full distributional assumption about  $\varepsilon_t$ .

In time series of counts, the process can be more volatile than expected under, e.g., an overdispersed Poisson assumption. This could for example happen when the mean level is high. Related phenomena have in recent years been of considerable interest in the financial econometrics literature. Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model and its extensions are based on decomposing an error term of a regression model as  $\varepsilon_t \sigma_t$ , where  $\varepsilon_t$  is an iid sequence with zero mean and unit variance and where  $\sigma_t$  is a conditional (on past observations) standard deviation. In addition, conditional and unconditional independence between  $\varepsilon_t$  and  $\sigma_t$  are required. With count (i.e. integer-valued) data Engle's decomposition is not a convenient one. Here, we introduce a new mechanism in terms of moments, which is better suited for count data models.

Section 2 presents the modelling approach and gives some of its implied properties. In Section 3, we consider ML and GMM approaches to the estimation of unknown model parameters. We revisit old estimators and adapt them to the new needs of the model.

Section 4 reports the results of a study of self-feeding activity in fish. The number of hourly bites on self-feeders is the manifest variable for the self-feeding activity, and a 137 days long period is examined. The current examination of the data utilizes a different modelling approach than the original study of Brännäs, Berglund and Eriksson (2001). The final section concludes.

## 2. MODEL APPROACH

The basic building block for most count data regression modelling is the Poisson model. The Poisson distribution is based on independent increments which implies that for a count variable  $y_t$  at time  $t$

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

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

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

where  $\mathbf{x}_t$  is a vector of explanatory variables and  $\boldsymbol{\beta}$  is a vector of parameters. In this model the unconditional and conditional heteroskedasticities are identical and observations are serially uncorrelated. The means and variances are equal, which is a restrictive specification with respect to various time series applications.

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  $\varepsilon_t$  so that

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

Conventionally one assumes  $\{\varepsilon_t\}$  to be an iid sequence with  $E(\varepsilon_t) = 1$  and  $V(\varepsilon_t) = \sigma^2$ . The conditional and unconditional moments are then internally equal, i.e.

$$\begin{aligned} E(y_t) &= E(y_t|F_{t-1}) = \lambda_t \\ V(y_t) &= V(y_t|F_{t-1}) = \lambda_t + \sigma^2 \lambda_t^2, \end{aligned} \tag{1}$$

but the means and variances are no longer equal. When  $\varepsilon_t$  is assumed gamma distributed the unconditional  $y_t$  has a negative binomial, NB2, distribution. To estimate, either such a fully parametric model as NB2 may be estimated by ML or, e.g., a GMM estimator based on only the given moments may be applied.

Zeger (1988) suggested an extension of the overdispersed Poisson model for time series data. One sets  $E(y_t|\varepsilon_t) = V(y_t|\varepsilon_t) = \varepsilon_t\lambda_t$  and assumes the stationary  $\{\varepsilon_t\}$  sequence to again have  $E(\varepsilon_t) = 1$  and  $V(\varepsilon_t) = \sigma^2$ . Besides implying overdispersion this model gives serially correlated counts, though conditionally on  $\varepsilon_t$  and  $\varepsilon_s$ , respectively,  $y_t$  and  $y_s$  are uncorrelated. The unconditional mean and variance are those of the overdispersed model.

For this model one may show that

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

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  $\sigma_u^2$ . The parametrization is such that  $E(\varepsilon_t) = 1$ . Then  $E(\varepsilon_t|F_{t-1}) = \theta\varepsilon_{t-1} + (1 - \theta)$  and  $V(\varepsilon_t|F_{t-1}) = V(\varepsilon_t) = \sigma_u^2$ . Hence

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

In this case the conditional mean is affected by the correlation structure in  $\{\varepsilon_t\}$  in the same way as the conditional variance is.

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

To add flexibility in the conditional heteroskedasticity we may consider two approaches. First, we redefine  $\sigma^2$  to become time dependent and possibly dependent on previous observations. By this the size of the overdispersion becomes time dependent. Second, we alter the basic conditional moment expressions.

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

$$V(y_t|F_{t-1}) = \lambda_t + \sigma_t^2(F_{t-1})\lambda_t^2. \quad (3)$$

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

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  $\sigma_t^2$  remains positive an exponential form appears reasonable. Corresponding to exponential generalized ARCH or EGARCH specifications (Nelson, 1991) we could specify, say,

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

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  $\alpha_i$  parameters are feasible. Alternatively with  $\varepsilon_t$  gamma distributed,  $y_t$  follows a NB2 distribution and then ML estimation is feasible.

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

$$E(y_t|\varepsilon_t, \sigma_t) = V(y_t|\varepsilon_t, \sigma_t) = \lambda_t + (\varepsilon_t - 1)\sigma_t\lambda_t,$$

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

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

An obvious drawback with this specification arises from the requirement that  $\lambda_t + (\varepsilon_t - 1)\sigma_t\lambda_t \geq 0$ . This has a bearing on the functional form for  $\sigma_t$  for  $\varepsilon_t < 1$ . If, for example,  $\varepsilon_t = 0$  then  $\sigma_t < 1$  must hold, while no such restriction need to hold for  $\varepsilon_t > 1$ .

Approximately, the same moment properties can be obtained from the conditional representation  $\lambda_t \exp(\varepsilon_t \sigma_t)$ . If  $E(\varepsilon_t) = 0$ ,  $V(\varepsilon_t) = 1$  and  $\varepsilon_t \sigma_t$  is small, a first order Taylor expansion gives  $E(\exp(\varepsilon_t \sigma_t)) \approx 1$  and  $V(\exp(\varepsilon_t \sigma_t)) \approx 1 + \sigma_t^2$ . For this specification there are only size restrictions on  $\varepsilon_t \sigma_t$ . Note that a conditional specification  $\lambda_t \varepsilon_t \sigma_t$ , which appears closer to the continuous variable specification, would with  $E(\varepsilon_t) = 1$  result in a model where it would be more difficult to separate the mean and variance effects. We could obviously also express the model on a form closer to the mainstream conditional heteroskedasticity literature. By using  $y_t = E(y_t|F_{t-1}) + u_t$ , where  $E(u_t) = 0$  and  $V(u_t) = \lambda_t$ , we get results corresponding to the Poisson model. If we set  $u_t = \varepsilon_t \lambda_t$  with  $E(\varepsilon_t) = 0$  and  $V(\varepsilon_t) = \sigma_t^2(F_{t-1})$  we get  $V(y_t|F_{t-1}) = \lambda_t + \sigma_t^2(F_{t-1})\lambda_t^2$ . Distributionally this route is far from easy.

In the generalized model of Winkelmann & Zimmermann (1991) the conditional variance is of the form  $\lambda_t + (\sigma^2 - 1)\lambda_t^{k+1}$ , where both  $\sigma^2$  and  $k$  are treated as unknown parameters. Therefore, in this model a quadratic effect is not assumed. Our specification is more general in the sense that a quadratic effect can be very strongly present when  $\sigma_t^2$  is large, and the conditional variance can be close to the variances of the negative binomial (NB1 with variance  $\lambda_t(1 + \sigma_t^2)$ ) or the Poisson for smaller  $\sigma_t^2$ .

### 3. ESTIMATION

We discuss some likelihood based approaches to estimation and GMM estimation. The unknown parameter vectors are  $\beta$  in the  $\lambda_t = \exp(\mathbf{x}_t\beta)$  function and  $\alpha$  in the  $\sigma_t^2$  specification.

For a general case, the log-likelihood function conditional on the first observation is

$$\ln L = \sum_{t=r}^T \ln L_t = \sum_{t=r}^T \ln f(y_t|F_{t-1}),$$

where  $r$  is one plus the largest lag found in the conditional variance expression or in  $\lambda_t$ . The conditional density

$$f(y_t|F_{t-1}) = \int_0^\infty f(y_t|\varepsilon_t, F_{t-1})f(\varepsilon_t) d\varepsilon_t = E_\varepsilon [f(y_t|\varepsilon_t, F_{t-1})] \quad (4)$$

obviously requires distributional assumptions on the conditional variable  $y_t$  and on  $\varepsilon_t$ . For the former we assume throughout a Poisson distribution. Unfortunately, there are in general no strong justifications for an assumption about  $\varepsilon_t$ .

If the basic Poisson model is used for estimation, i.e. no account is given to  $\sigma_t^2$ , it is straightforward to show the consistency and asymptotic normality of the ML estimator of  $\beta$ . The covariance matrix estimator is a simple modification (from a constant  $\sigma^2$  to a time dependent  $\sigma_t^2$ ) of the sandwich estimator for overdispersed data (e.g., Gouriéroux, Monfort & Trognon, 1984; Cameron & Trivedi, 1998, p. 65):

$$V(\hat{\beta}) = \left( \sum_{t=r}^T \lambda_t \mathbf{x}_t' \mathbf{x}_t \right)^{-1} \left( \sum_{t=r}^T (\lambda_t + \sigma_t^2 \lambda_t^2) \mathbf{x}_t' \mathbf{x}_t \right) \left( \sum_{t=r}^T \lambda_t \mathbf{x}_t' \mathbf{x}_t \right)^{-1}.$$

The estimator is evaluated at the Poisson ML estimates, but also requires estimation of a  $\sigma_t^2$  model. An estimator that avoids this problem is of the form:

$$V(\hat{\beta}) = \left( \sum_{t=r}^T \lambda_t \mathbf{x}_t' \mathbf{x}_t \right)^{-1} \left( \sum_{t=r}^T (y_t - \lambda_t)^2 \mathbf{x}_t' \mathbf{x}_t \right) \left( \sum_{t=r}^T \lambda_t \mathbf{x}_t' \mathbf{x}_t \right)^{-1}.$$

Hence, if  $\sigma_t^2$  is considered a function of no interest the Poisson ML estimator can be applied and using a sandwich covariance matrix estimator, testing about functions of  $\beta$  is straightforward.

A technically convenient assumption about  $\varepsilon_t$  is the gamma distribution with expected value one and variance  $\sigma_t^2$ , i.e.

$$f(\varepsilon_t) = \frac{1}{\Gamma(\sigma_t^{-2})(\sigma_t^2)^{\sigma_t^{-2}}} \varepsilon_t^{\sigma_t^{-2}-1} \exp(-\varepsilon_t/\sigma_t^2).$$

We get the negative binomial distribution (NB2)

$$f(y_t | F_{t-1}) = \frac{\Gamma(y_t + \sigma_t^{-2})}{y_t! \Gamma(\sigma_t^{-2})} \left( \frac{\lambda_t}{\lambda_t + \sigma_t^{-2}} \right)^{y_t} \left( \frac{\sigma_t^{-2}}{\lambda_t + \sigma_t^{-2}} \right)^{\sigma_t^{-2}} \quad (5)$$

and using  $\ln [\Gamma(y + \sigma^{-2})/\Gamma(\sigma^{-2})] = \sum_{j=0}^{y-1} \ln(j + \sigma^{-2})$  and  $\lambda_t = \exp(\mathbf{x}_t \beta)$  the log-likelihood function is

$$\begin{aligned} \ln L = \sum_{t=2}^T \ln L_t &= \sum_{t=r}^T \left[ \sum_{j=0}^{y_t-1} \ln(j + \sigma_t^{-2}) - \ln y_t! - (y_t + \sigma_t^{-2}) \ln[1 + \sigma_t^2 \lambda_t] \right. \\ &\quad \left. + y_t \ln \sigma_t^2 + y_t \ln \lambda_t \right]. \end{aligned}$$

Cameron & Trivedi (1998, p. 71) and others give the likelihood equations  $\partial \ln L / \partial \beta$  and  $\partial \ln L / \partial \sigma^2$  for the the constant  $\sigma^2$  parameter NB2 model. For the time dependent model the likelihood equation for  $\beta$  is more complicated if  $\sigma_t^2$  is a function of  $u_t = y_t - \lambda_t$ . The gradient for a particular  $\sigma_t^2$  model is given in the Appendix. With other distributional assumptions about  $\varepsilon_t$  there is, in general, no explicit solution arising from (4).

An obvious approach is to use numerical integration for the integral in (4). For the log-normally distributed  $\varepsilon_t$  Gaussian-Hermite quadrature has been used by, e.g., Hinde (1982).

In this case the log-likelihood function takes the form

$$\ln L = \sum_{t=r}^T \ln \sum_{j=1}^M w_j \frac{\exp(-\exp(\psi_{j,t}) + y_t \psi_{j,t})}{\sqrt{\pi} y_t!},$$

where  $\psi_{j,t} = \mathbf{x}_t \boldsymbol{\beta} + v_j \sqrt{2 \ln(1 + \sigma_t^2)} - \ln(1 + \sigma_t^2)/2$ . Here,  $w_j$  and  $v_j, j = 1, \dots, M$ , are the weights and abscissas for the quadrature. Obviously, other distributional assumptions may lead to integrals that are better approached by other devices.

An alternative likelihood based approach is the simulated maximum likelihood (SML) estimator of Gouriéroux & Monfort (1991). By this, one can avoid the complications due to nonexplicit solutions in (4). Numerical integration is then replaced by simulation. A natural estimator of  $f(y_t|F_{t-1}) = E_\varepsilon[f(y_t|\varepsilon_t, F_{t-1})]$  is the sample moment, i.e.

$$\hat{f}(y_t|F_{t-1}) = \frac{1}{S} \sum_{s=1}^S f(y_t|\varepsilon_t^{(s)}, F_{t-1}),$$

where  $\varepsilon_t^{(s)}$  is the  $s$ th randomly generated replication from some assumed marginal distribution for  $\varepsilon_t$  having mean one and variance  $\sigma_t^2$ . The estimator is used to obtain numerical values on the log-likelihood function, i.e.

$$\ln L_S = \sum_{t=r}^T \ln \frac{1}{S} \sum_{s=1}^S f(y_t|\varepsilon_t^{(s)}, F_{t-1}),$$

for given values on  $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}')' \in \boldsymbol{\Theta}$  and where the conditional density is Poisson. The resulting SML estimator is obtained as  $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \ln L_S$ . The asymptotic properties of the estimator are summarized by Gouriéroux & Monfort (1996, ch. 3).

Estimation is also feasible under weaker assumptions using the generalized method of moment (GMM) estimator. The GMM estimator is consistent and asymptotically normal under weak assumptions (e.g., Hansen, 1982; Davidson & MacKinnon, 1993). As the estimator is based on the conditional moment structure, no full distributional assumptions are required for neither the conditional distribution of  $y_t$  nor for  $\varepsilon_t$ . We may therefore expect the estimator to be more robust with respect to distributional misspecification, but to be less efficient than an ML estimator should the distributional assumptions be correct. As a moment condition for the conditional mean function the likelihood equation  $(T-r)^{-1} \sum_{t=r}^T \mathbf{x}_t'(y_t - \lambda_t) = (T-r)^{-1} \sum_{t=r}^T \mathbf{m}_{1t}$  is an obvious candidate. To generate moment conditions for the conditional heteroskedasticity parameters the variance expression in (3) is basic. To catch the conditional variance we may depart from  $V(y_t|F_{t-1})$ , and use  $(y_t - \lambda_t)^2 - \lambda_t - \sigma_t^2 \lambda_t^2$  as a residual in forming conditions. In this case too,  $\mathbf{x}_t$  can be used as instruments. The resulting moment condition is  $(T-r)^{-1} \sum_{t=r}^T \mathbf{m}_{2t}$ . We set  $\mathbf{m}' = (T-r)^{-1} [\sum_{t=r}^T \mathbf{m}_{1t}' : \sum_{t=r}^T \mathbf{m}_{2t}']$ .

The GMM estimator minimizes the criterion function

$$q = \mathbf{m}' \mathbf{W}^{-1} \mathbf{m},$$

where  $\mathbf{W}$  in a first step is set equal to the identity matrix. For the second step  $\hat{\mathbf{W}} = (T-r)^{-1} \sum_{t=r}^T \hat{\mathbf{m}}_t \hat{\mathbf{m}}_t'$  is based on the first step estimates. The estimator of the asymptotic covariance matrix is

$$V(\hat{\boldsymbol{\theta}}) = (T-r)^{-1} \left( \hat{\mathbf{G}}' \hat{\mathbf{W}}^{-1} \hat{\mathbf{G}} \right)^{-1},$$

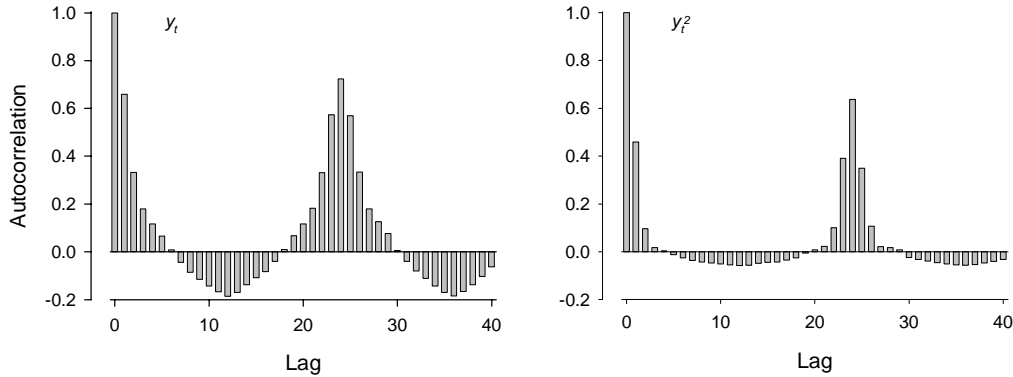


Figure 1: Autocorrelation functions for  $y_t$  and  $y_t^2$ .

where  $\mathbf{G} = \partial \mathbf{m} / \partial \theta'$  and  $\mathbf{W}$  are evaluated at final estimates.

This basic GMM framework can be extended by incorporating additional instrumental variables beyond  $\mathbf{x}_t$ . Examples of such variables could be lagged variables, though lagging, e.g., dummy and trend variables is not a good idea, and other explanatory variables. Another extension is to include more moment conditions arising, e.g., from a restriction of no serial correlation. The GMM estimator is more efficient the more moment conditions that are incorporated.

For testing purposes Wald, likelihood ratio or Lagrange multiplier types of tests against conditional heteroskedasticity and/or overdispersion can be obtained for testing  $\alpha = \mathbf{0}$ , while testing only the 'slope' parameters of this vector in the  $\sigma_t^2$  model is a test against conditional heteroskedasticity. Such tests can be based on either the ML or GMM approaches. An overall test of the GMM tested model is  $(T - r)\hat{q} \stackrel{a}{\sim} \chi^2(l)$ , where  $l$  is the number of overidentifying restrictions.

#### 4. EMPIRICAL RESULTS

The illustration is based on data from a related study of Brännäs et al. (2001) on time learning and anticipatory activity in fish. The studied time series records the numbers of trigger actuations of self-feeders, hourly over 137 days, for 84 experimental Arctic charr fish. Observations started January 29, 1999, and the time series length is  $T = 3312$ . In the study of Brännäs et al. (2001) the specification was static using hourly dummy variables and Poisson ML estimation was employed with a robust covariance matrix estimator. In this study we employ a dynamic model and explicitly model  $\sigma_t^2$ .

The autocorrelation functions for  $y_t$  and  $y_t^2$  indicate that the series contain a 24 hour cycle and a trend component, cf. Figure 1.

Table 1 gives the parameter estimates for a model without added conditional heteroskedasticity (estimated by Poisson ML) and three specifications that encompass added conditional heteroskedasticity (estimated by negative binomial, Poisson-lognormal ML with  $M = 100$  and GMM). The period variables are throughout dummy variables. Periods 1, 3

Table 1: Poisson, negative binomial (NB2) and Poisson - lognormal ML and GMM estimates.

Variable	Poisson ML		Negative Binomial ML		Poisson - Lognormal ML		GMM	
	Est	s.e.	Est	s.e.	Est	s.e.	Est	s.e.
$\lambda_t$ -part								
$y_{t-1}^*$	0.357	0.029	0.475	0.027	0.395	0.027	0.369	0.030
$y_{t-2}^*$	-0.097	0.022	-0.136	0.029	-0.077	0.026	-0.063	0.022
$y_{t-24}^*$	0.166	0.028	0.307	0.024	0.118	0.024	0.202	0.028
$d_{t-1}$	-0.324	0.121	-0.081	0.066	-0.468	0.081	-0.060	0.101
$d_{t-2}$	-0.202	0.077	-0.445	0.068	-0.292	0.074	-0.042	0.073
$d_{t-24}$	-0.512	0.115	-0.303	0.071	-0.597	0.082	-0.705	0.106
Trend	1.726	0.373	0.838	0.300	0.721	0.318	0.751	0.339
Seasonal	0.542	0.065	0.635	0.050	0.453	0.066	0.575	0.063
Period 1	0.388	0.279	-0.056	0.212	-0.632	0.221	-0.050	0.251
Period 2	-0.064	0.185	0.116	0.168	-0.623	0.167	-0.304	0.172
Period 3	0.403	0.124	-0.053	0.127	-0.266	0.120	0.117	0.115
Period 4	-0.291	0.092	-0.098	0.110	-0.905	0.100	-0.273	0.091
Food	1.101	0.101	0.375	0.111	1.158	0.091	0.779	0.099
Pre-food	0.975	0.132	0.645	0.147	0.606	0.146	0.817	0.124
Post-food	0.075	0.108	-0.662	0.185	0.187	0.124	-0.127	0.105
Constant	0.019	0.355	0.287	0.294	1.386	0.311	0.636	0.326
EGARCH-part								
Constant			0.406	0.036	1.016	0.116	0.403	0.082
$u_{t-1}$			-0.015	0.001	-0.034	0.004	-0.036	0.005
$u_{t-2}$			0.004	0.002	0.010	0.004	-0.035	0.015
$u_{t-24}$			-0.005	0.001	-0.014	0.004	-0.037	0.007
Seasonal			-0.797	0.054	-1.336	0.135	-1.844	0.494
$s_{t-1}$			0.448	0.041	0.486	0.047	0.193	0.207
$\ln L$	-12974		-7124		-7047			
$q, p$ -value							21.77	0.02

Estimation by repeated use of a simplex algorithm. For the negative binomial and Poisson-log-normal ML estimators a BHHH covariance matrix estimator based on numerical derivatives is employed.

Table 2: Constants  $c_i$  for Poisson and negative binomial ML and GMM estimates.

	Poisson ML		Negative Binomial ML		Poisson - Lognormal ML		GMM	
	Est	s.e.	Est	s.e.	Est	s.e.	Est	s.e.
$c_1$	0.404	0.05	0.396	0.03	0.306	0.08	0.850	0.03
$c_2$	8.024	1.19	9.281	14.4	44.35	153	1.948	0.08
$c_{24}$	0.046	0.02	0.373	0.10	0.006	0.02	0.030	0.01



and 5 (base period) provided free access to food 24 hours a day. In Period 2 food is available 1200-1400, and in Period 4 0900-1100. Period lengths varied between 12 and 55 days. Pre-food is a dummy variable for the hour before restricted feeding became available in Periods 2 and 4, while Post-food is a corresponding dummy variable for the hour after food availability. The trend variable is defined as  $t/T$  and the seasonal variable as  $\sin(2\pi(h-6)/24)$  with  $h$  the hour within the day. The light and dark periods were both uninterrupted 12 hours long throughout. The model is of the Zeger & Qaqish (1988) type that contains  $\ln \max(y_{t-i}, c_i)$  as explanatory variables. The respecification of Cameron & Trivedi (1998, pp. 239-240) uses  $y_{t-i}^{**} = y_{t-i}$  and  $d_{t-i} = 0$  for  $y_{t-i} > 0$  and  $y_{t-i}^{**} = 1$  and  $d_{t-i} = 1$  for  $y_{t-i} = 0$ . Finally,  $y_{t-i}^* = \ln(y_{t-i}^{**})$  and  $d_{t-i}, i = 1, 2$  and 24 are used as explanatory variables. If the parameter in front of  $y_{t-i}^{**}$  is  $\rho_i$  the parameter in front of  $d_{t-i}$  is  $\rho_i \ln c_i$ . Then  $\hat{c}_i = \exp(\widehat{\rho_i \ln c_i / \hat{\rho}_i})$ .

The conditional variance specification is  $\sigma_t^2 = \exp(s_t)$ , where  $s_t = \alpha_0 + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \alpha_{24} u_{t-24} + \gamma_1 \text{Seasonal}_t + \gamma_2 s_{t-1}$ . To arrive at this specification a number of model versions were estimated by GMM and compared with respect to GMM criterion values and whether variables had significant effects. There is significant overdispersion ( $p = 0$ ) after Poisson ML estimation and employing a likelihood ratio test against either the negative binomial or Poisson-lognormal model the Poisson model can be rejected. The Poisson model is also rejected against a constant  $\sigma^2$ , negative binomial model. The overall test of the GMM estimated model ( $p = 0.02$ ) raises some questions about the validity of the model specification or of the used instruments in  $\mathbf{X}$ .

In most cases the signs and sizes of estimates as well as standard errors remain the same across the estimated models. The biologically most interesting effects are those of the food availability variables. During food availability periods the number of feeder actuations are significantly larger than when no food is available. There is a significant anticipatory effect with an increased number of actuations before food becomes available, while the effect after such periods is generally not significant. Qualitatively identical findings were reported by Brännäs et al. (2001). Figure 2 suggests that the effect of time restricted feeding gives larger actuation frequencies during the observed high activity hours than if feeding is available throughout the day. The base case is related to that of Figure 3, while the alternative of free access is obtained by setting the Period 4 and food dummy variables equal to zero throughout the period.<sup>1</sup>

The estimates of the constants  $c_i$  are given in Table 2. With the exception of the  $\hat{c}_2$ s the sizes are reasonable as we expected estimates smaller than one. When it comes to the estimates in the  $\sigma_t^2$  specification the estimates of lagged  $u_{t-i}$  are quite small and in the GMM case of quite similar size. Note that the effects of lagged  $u_{t-i}$  are asymmetric with large effects of negative residuals and much less effect of positive ones (cf. Figure 3). The effect of the seasonal variable is negative which dampens the conditional variance at peak hours in activity (see also Figure 4). The persistence effect of  $s_{t-1}$  is positive but small for GMM and not too large for the other estimators.

Figure 3 shows fitted values after GMM estimation. There is a reasonable fit for lower counts, but for this selection of hours we evidently miss the high frequency noted at 10 AM for the first day and the one at 9 AM for the second day. The conditional standard deviation has a pattern not unlike the observed outcome though of smaller amplitude. Note that  $\sigma_t^2$

---

<sup>1</sup>The dynamic simulation starts from the identity  $y_t = \lambda_t(y_{t-1}, y_{t-2}, y_{t-24}, \mathbf{x}_t) + \xi_t$ , and sets  $\xi_t = 0$ , since  $E(\xi_t) = E(y_t - \lambda_t) = 0$ .

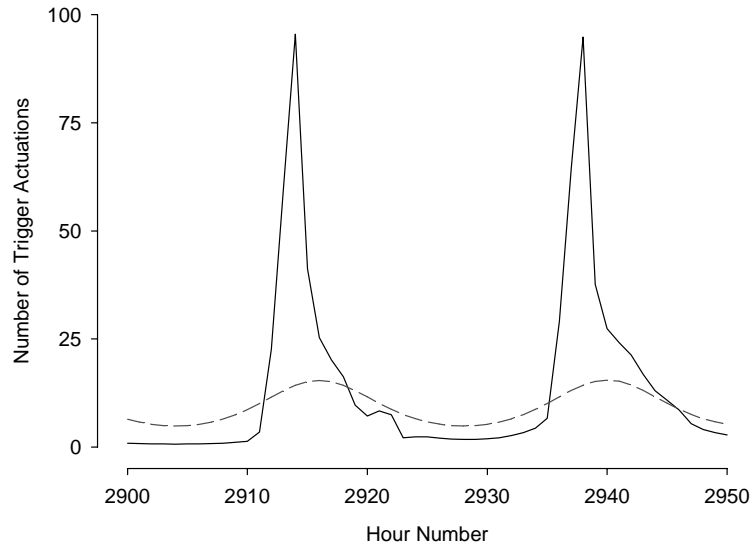


Figure 2: Simulated dynamic effects of a change in Period 4 to free access (dashed line) from time restricted feeding (solid line) starting at hour number 2900.

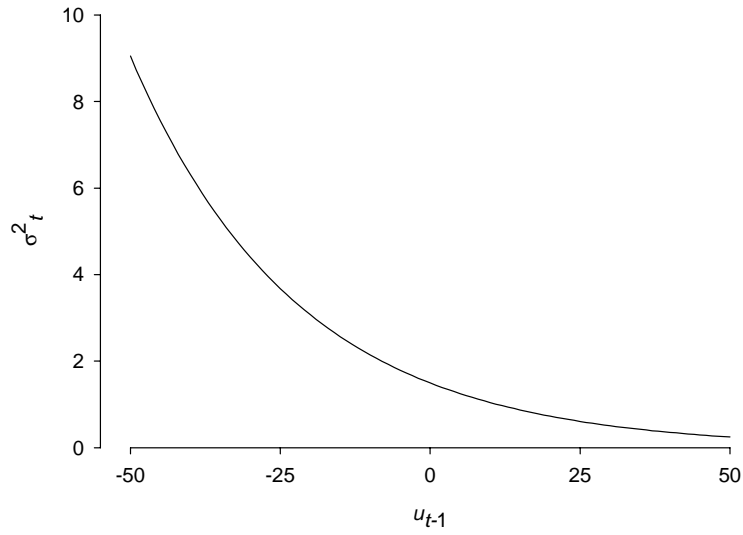


Figure 3: Effect of  $u_{t-1}$  on  $\sigma_t^2$  in a situation with a constant term and all other variables kept at zero.

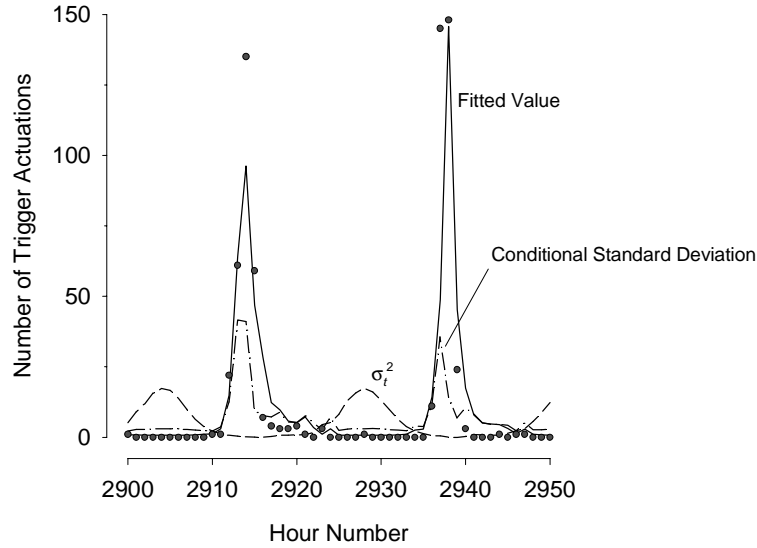


Figure 4: Observed number of trigger actuations (circle) with fitted values ( $\hat{\lambda}_t$ , solid line), conditional standard deviation  $((\hat{\lambda}_t + \hat{\sigma}_t^2 \hat{\lambda}_t^2)^{1/2})$ , dot-dashed line) and conditional variance function ( $\hat{\sigma}_t^2$ , dashed line) for a 50 hour period. Based on GMM estimation of full model.

varies contra-cyclically with large values when  $\hat{\lambda}_t$  is small while dampening the conditional variance (or standard deviation) when  $\hat{\lambda}_t$  is large.

## 5. CONCLUSIONS

The paper has suggested an extension of the overdispersed Poisson regression model or some of its conditional relatives to include time dependent overdispersion. This renders a more flexible conditional heteroskedasticity or variance function. ML estimation using the negative binomial and Poisson–lognormal models is developed. The Poisson ML estimator is consistent and asymptotically normal but empirically a robust covariance matrix estimator should be used. We also consider a basic framework for GMM estimation. More instruments and possibly employing additional moment conditions would probably reduce the estimated standard errors of this estimator.

Empirically we found that Arctic charr exhibits a feeding activity that is increased during the hours of restricted feeding. There is also a significant and positive anticipatory effect.

## ACKNOWLEDGMENTS

The partial financial support to the first author from the Wallander Foundation and The Research Institute of the Umeå School of Business and Economics and to the second author from The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning is gratefully acknowledged.

# APPENDIX

The gradient vector for the negative binomial model in (5)-(6) with  $\sigma_t^2 = \exp(s_t)$ , where  $s_t = \alpha_0 + \sum_{i=1}^q (\alpha_{1i} u_{t-i} + \alpha_{2i} u_{t-i}^2) + \sum_{i=1}^p \alpha_{3i} s_{t-i}$ , with  $u_t = y_t - \lambda_t$ , is:

$$\begin{aligned} \frac{\partial \ln L_t}{\partial \beta} &= - \sum_{j=0}^{y_t-1} \frac{\dot{\mathbf{s}}'_t}{1 + j \sigma_t^2} + (\mathbf{x}'_t + \dot{\mathbf{s}}'_t) \frac{y_t - \lambda_t}{1 + \sigma_t^2 \lambda_t} + \dot{\mathbf{s}}'_t \sigma_t^{-2} \ln(1 + \sigma_t^2 \lambda_t) \\ \frac{\partial \ln L_t}{\partial \sigma_t^2} &= \frac{1}{\sigma_t^4} \left[ \ln(1 + \sigma_t^2 \lambda_t) - \sum_{j=0}^{y_t-1} \frac{1}{j + \sigma_t^2} \right] + \mathbf{x}'_t \frac{y_t - \lambda_t}{1 + \sigma_t^2 \lambda_t} \\ \frac{\partial \sigma_t^2}{\partial \alpha} &= \frac{\partial \ln L_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \alpha}, \end{aligned}$$

where

$$\begin{aligned} \partial \sigma_t^2 / \partial \beta &= (\partial s_t / \partial \beta) \sigma_t^2 = \dot{\mathbf{s}}'_t \sigma_t^2 \\ \partial \sigma_t^{-2} / \partial \beta &= -\dot{\mathbf{s}}'_t \sigma_t^2 \\ \dot{\mathbf{s}}'_t &= \sum_{i=1}^q \mathbf{x}'_{t-i} \lambda_{t-i} (\alpha_{1i} - 2\alpha_{2i} u_{t-i}) + \sum_{i=1}^p \alpha_{3i} \dot{\mathbf{s}}'_{t-i} \\ \partial s_t / \partial \alpha_0 &= 1 + \sum_{i=1}^p \alpha_{3i} \dot{s}_{t-i}^{(0)} \\ \partial s_t / \partial \alpha_{11} &= u_{t-1} + \sum_{i=1}^p \alpha_{3i} \dot{s}_{t-i}^{(11)} \\ &\dots \\ \partial s_t / \partial \alpha_{3q} &= s_{t-q} + \sum_{i=1}^p \alpha_{3i} \dot{s}_{t-i}^{(3q)} \end{aligned}$$

and all initial conditions can be set at, e.g.,  $\dot{s}_{r-1} = \dots = \dot{s}_1 = 0$ .

## REFERENCES

- BRÄNNÄS, E., BERGLUND, U. & ERIKSSON, L-O. (2001). Time Learning and Anticipatory Activity in Groups of Arctic Charr. Submitted, Swedish University of Agricultural Science, Umeå.
- BRÄNNÄS, K. & JOHANSSON, P. (1994). Time Series Count Data Regression. *Communications in Statistics: Theory and Methods* **23**, 2907-2925.
- CAMERON, A.C. & TRIVEDI, P. (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.
- DAVIDSON, R. & MACKINNON, J.G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press, Oxford.
- ENGLE, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **50**, 987-1008.
- GOURIÉROUX, C. & MONFORT, A. (1991). Simulation Based Econometrics in Models with Heterogeneity. *Annales d'Economie et de Statistique* **20**, 69-107.
- GOURIÉROUX, C. & MONFORT, A. (1996). *Simulation-Based Econometric Methods*. Oxford University Press, Oxford.
- HANSEN, L.P. (1982). Large Sample Properties of Generalized Method of Moment Estimators. *Econometrica* **50**, 1029-1054.
- HINDE, J. (1982). Compound Poisson Regression Models. In Gilchrist, R. (ed.) *GLIM 1982: Proceedings of the International Conference on Generalized Linear Models*. Springer-Verlag, Berlin.
- JOHNSON, N.L. & KOTZ, S. (1969). *Discrete Distributions*. Houghton Mifflin, Boston.
- NELSON, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **59**, 347-370.
- WINKELMANN, R. & ZIMMERMANN, K.F. (1991). A New Approach for Modelling Economic Count Data. *Economics Letters* **37**, 139-143.
- ZEGER, S.L. (1988). A Regression Model for Time Series of Counts. *Biometrika* **75**, 621-629.
- ZEGER, S.L. & QAQISH, B. (1988). Markov Regression Models for Time Series: A Quasi-Likelihood Approach. *Biometrics* **44**, 1019-1031.