AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

CANTER

DP 98-11

Department of Economics
UNIVERSITY OF CANTERBURY,
CHRISTCHURCH, NEW ZEALAND

ISSN 1171-0705

LIBRARY
WITHDRAWN
NOV 1 1 1998
GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS

RANDOM EFFECTS MODELS FOR PANEL COUNT DATA

Rainer Winkelmann

Discussion Paper

No. 9811

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9811

September 1998

RANDOM EFFECTS MODELS FOR PANEL COUNT DATA

Rainer Winkelmann

Random Effects Models for Panel Count Data

Rainer Winkelmann

University of Canterbury and Centre for Economic Policy Research, London.

August 1998

This paper discusses the specification and extimation of random effects count data models. A new multivariate count data model with negative binomial marginals is derived. In contrast to the existing multivariate Poisson model, this model allows for over-dispersion, a phenomenon that is frequently encountered in real data applications. In addition, a multi-factor Poisson model with general individual specific covariance structure is formulated, and an algorithm for estimating the parameters of the model by Simulated Maximum Likelihood is presented.

Keywords: Poisson convolution; multi-factor models; Simulated Maximum Likelihood.

JEL Classification: C23, C25.

---

[1] *Address for correspondence*: Department of Economics, PB 4800, Christchurch, New Zealand. Fax 643 364 2635. E-mail: r.winkelmann@econ.canterbury.ac.nz.

1 Introduction

This paper is concerned with the problems of specification and estimation of panel count data models with random effects. These problems were first addressed by Hausman, Hall and Griliches (1984) who suggested the use of gamma distributed individual random effects, and by Gourieroux, Monfort and Trognon (1984) who proposed a semi-parametric approach. Recent contributions to the literature include Brännäs and Johannson (1996), Cincera (1997), and Chib, Greenberg, and Winkelmann (1997, 1998).

This paper summarizes the existing methods and offers some extensions. Firstly, I discuss a model based on additive Poisson distributed random effects. The resulting model is easy to implement and has several desirable properties. In particular, the marginal distribution of the counts has a covariance structure very similar to the one of the linear random effects model. Furthermore, the model is readily extended to allow for extra-Poisson variation, or *over-dispersion*. Secondly, I show how the parameters of a multi-factor model with correlated errors can be estimated without much hassle by Simulated Maximum Likelihood.

Throughout this paper, the conditional expectation function is kept the same for all models, so that estimated regression parameters are directly comparable across the different specifications. An application section illustrates the practicalities of implementing the different models, and studies the sensitivity of the results with regard to the choice of model.

2 One-factor models

I consider the situation where data are observed in clusters such that $y_i = (y_{i1}, \ldots, y_{iT})$ and $x_i = (x_{i1}, \ldots, x_{iT})$ denote observations on the ith cluster, $i = 1, \ldots, n$. Typically, t represents time and i represents individuals.[2] The most well known example of a one-factor

---

[2] All the methods are easily extended to the case of a variable cluster size T_i.

model for continuous data is the normal linear model with cluster specific random effects (see Hsiao, 1986)

$$y_{it} = x'_{it}\beta + u_i + \varepsilon_{it}\,, \tag{1}$$

where u_i and ε_{it} are independently normal distributed with means 0 and variances σ_u^2 and σ_ε^2, respectively. Then $y_i \sim MVN(x_i'\beta, \Omega)$ where

$$\Omega = \sigma_\varepsilon^2 I_T + \sigma_u^2 \iota_T \iota_T'\,. \tag{2}$$

MVN is the multivariate normal distribution, and ι denotes a $(T \times 1)$ vector of ones.

Now consider the situation where y_i represents counts rather than continuous responses. The benchmark model is the Poisson regression model

$$f(y_i|x_i) = \prod_{t=1}^{T} \frac{\exp(-\lambda_{it})\lambda_{it}^{y_{it}}}{\Gamma(y_{it}+1)} \tag{3}$$

where $\lambda_{it} = \exp(x'_{it}\beta)$.

Hausman, Hall and Griliches (1984) have proposed to introduce cluster specific heterogeneity in a multiplicative way. In particular, the mixture multivariate density of y_i takes the form

$$f(y_i|x_i) = \int \left[\prod_{t=1}^{T} \frac{\exp(-\lambda_{it}u_i)(\lambda_{it}u_i)^{y_{it}}}{\Gamma(y_{it}+1)}\right] g(u_i)\mathrm{d}u_i \tag{4}$$

If u_i is gamma distributed with $\mathrm{E}(u_i) = 1$ and $\mathrm{Var}(u_i) = \gamma$ it can be shown that the joint PDF of y_i is of a negative binomial variety. This model will be referred to as the **Poisson-Gamma** model.[3] Define $y_{i.} = \sum_t y_{it}$ and $\lambda_{i.} = \sum_t \lambda_{it}$. Then

$$f(y_i|x_i) = \left[\prod_{t=1}^{T} \frac{(\lambda_{it})^{y_{it}}}{\Gamma(y_{it}+1)}\right] \frac{\Gamma(y_{i.}+\gamma^{-1})}{\Gamma(\gamma^{-1})}(1+\sigma_u^2\lambda_{i.})^{\gamma} \left(\frac{\gamma}{1+\gamma\lambda_{i.}}\right)^{y_{i.}} \tag{5}$$

---

[3] If, instead, u_i was log-normal the integral would not have a closed form solution but rather require numerical methods. An algorithm based on Gauss-Hermite quadrature is discussed in Winkelmann (1998).

The marginals of the Poisson-Gamma model are univariate negative binomial,

$$E(y_i) = \lambda_i$$

and

$$\mathrm{Var}(y_i) = \Lambda_i + \Lambda_i \gamma \iota \iota' \Lambda_i$$

where $\lambda_i = (\lambda_{i1}, \ldots, \lambda_{iT})$ and $\Lambda_i = \mathrm{diag}(\lambda_i)$.[4]

Hence, the Poisson-Gamma model allows for *over-dispersion*, a phenomenon frequently encountered in applications. The within cluster covariances are an increasing function of both γ, the variance of the multiplicative heterogeneity term, and the product of the expected values. The correlation matrix is given by

$$\mathrm{Corr}(y_i) = C_i \Lambda_i C_i' + C_i \Lambda_i \gamma \iota \iota' \Lambda_i C_i'$$

where $C_i = \mathrm{diag}\{(\lambda_{i1} + \gamma \lambda_{i1}^2)^{-1/2}, \ldots, (\lambda_{iT} + \gamma \lambda_{iT}^2)^{-1/2}\}$

Unlike the normal linear model with random effects, the Poisson-Gamma model does not lead to equi-covariances or equi-correlation within clusters. However, like for the normal model, within cluster correlations are bound to be positive.

3 An alternative model

While the previous model was based on a multiplicative error, it turns out to be possible to construct a simple parametric model with additive error as well. This approach exploits that the Poisson distribution (like the normal distribution) is closed under convolution and

---

[4] These results follow from the law of iterative expectations by which $E(y_i) = E_u[E(y_i|u_i)]$ and $\mathrm{Var}(y_i) = E_u[\mathrm{Var}(y_i|u_i)] + \mathrm{Var}_u[E(y_i|u_i)]$.

models the data as the sum of a count that is both individual and time specific plus a count that is specific to the individual.

$$y_{it} = z_{it} + u_i \tag{6}$$

The **Poisson-Poisson** model has been used before in the context of multivariate and "seemingly unrelated" count data (Gourieroux, Montfort and Trognon, 1984, King 1989, Jung and Winkelmann, 1993, Gurmu and Elder, 1998). However, I adopt it here to the situation of panel data with individual specific effect. Arguably, the approach is much better suited for longitudinal data than for multivariate data since the simple "one-factor" structure does only allow for non-negative within-cluster correlations. This restriction is likely to be undesirable in the context of systems of equations or multivariate counts, but it fits quite plausibly to the structure of panel data as was already noted in the linear random effects model. I also show that one further criticism of the model, its apparent inability to account for over-dispersion, can be overcome in an extended model.

Assume that z_{it} and u_i are independently Poisson distributed with parameters $\lambda_{it} - \gamma$ and γ, respectively. Then $y_i = (y_{i1} + u_i, \ldots, y_{iT} + u_i)$ has a multivariate Poisson distribution with the following properties: Each element y_{it} is Poisson distributed with

$$\mathrm{E}(y_{it}) = \mathrm{Var}(y_{it}) = \lambda_{it}$$

where $\lambda_{it} = \exp(x'_{it}\beta)$. Moreover,

$$\begin{aligned} \mathrm{Var}(y_{it}, y_{js}) &= \mathrm{Var}(z_{it} + u_i, z_{js} + u_j) \\ &= \begin{cases} \lambda_{it} & \text{for } i = j \text{ and } t = s \\ \gamma & \text{for } i = j \text{ and } t \neq s \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

In cluster-specific notation

$$\mathrm{Var}(y_i) = \Lambda_i + \gamma(\iota_T \iota'_T - I_T)\,,$$

Hence, as in the linear random effects model, all covariances are the same. However, since the Poisson model is inherently heteroscedastic, the Poisson-Poisson model is not equicorrelated. Correlations are given by $\gamma/[(\lambda_{it})^{-1/2}(\lambda_{is})^{-1/2}]$, or, in matrix-form,

$$\mathrm{Corr}(y_i) = C_i\Lambda_i C_i' + C_i\gamma\iota\iota' C_i'$$

where $C_i = \mathrm{diag}\{(\lambda_{i1})^{-1/2}, \ldots, (\lambda_{iT})^{-1/2}\}$

In order to estimate the parameters of the Poisson-Poisson model by maximum-likelihood, I next derive the likelihood function. Let $s_i = \min(y_{i1}, \ldots, y_{iT})$. The generic joint PDF for cluster i under independence of z_{it} and u_i can be written as

$$f(y_{i1}, \ldots, y_{iT}) = \sum_{k=0}^{s_i} f(k) \prod_{t=1}^{T} f(y_{it} - k) \tag{7}$$

where f is the Poisson PDF. The intuition behind this joint probability function is as follows. First, we know that the cluster specific count u_i cannot exceed any of the observed counts $(y_{i1}, \ldots, y_{iT})$, because each count is the sum of u_i and a non-negative count z_{it}. Secondly, the probability of an observed count y_{it} is the sum over the joint probabilities $f(u_i)$ and $f(z_{it}) = f(y_{it} - u_i)$, where $u_i = 0, \ldots, s_i$. Multiplication follows from the independence assumption.

In explicit notation

$$\begin{aligned} f(y_{i1}, \ldots, y_{iT}) &= \sum_{k=0}^{s_i} \frac{e^{-\gamma}\gamma^k}{k!} \prod_{t=1}^{T} \frac{e^{-(\lambda_{it}-\gamma)}(\lambda_{it}-\gamma)^{y_{it}-k}}{(y_{it}-k)!} \\ &= \exp\left[(T-1)\gamma - \sum_{t=1}^{T}\lambda_{it}\right] \sum_{k=0}^{s_i} \frac{\gamma^k}{k!} \prod_{t=1}^{T} \frac{(\lambda_{it}-\gamma)^{y_{it}-k}}{(y_{it}-k)!} \end{aligned} \tag{8}$$

and for clusters with at least one zero observation, the joint probability function simplifies to

$$f(y_{i1}, \ldots, y_{iT})|_{\min(y_{i1},\ldots,y_{iT})=0} = \exp\left[(T-1)\gamma - \sum_{t=1}^{T}\lambda_{it}\right] \prod_{t=1}^{T} \frac{(\lambda_{it}-\gamma)^{y_{it}}}{(y_{it})!}$$

Also, for $\gamma = 0$ the joint PDF collapses to the standard Poisson model. As for the normal distribution, zero correlation implies independence in this particular model.

The Poisson-Poisson model can be readily extended to allow for over-dispersion. Perhaps the simplest way of doing so is to assume from the outset that z_{it} and u_i are independently negative binomial distributed. One possible reason could be that

$$z_{it}|\nu_{it} \sim \text{Poisson}(\lambda_{it}\nu_{it})$$

and

$$u_i|\kappa_i \sim \text{Poisson}(\gamma\kappa_i)$$

where ν_{it} and κ_i are independently gamma distributed. As before

$$y_{it} = z_{it} + u_i$$

Furthermore, assume that the parametrization is such that[5]

$$z_{it} \sim \text{NB with } \mathrm{E}(z_{it}) = \lambda_{it} - \gamma,\ \mathrm{Var}(z_{it}) = (\lambda_{it} - \gamma)(1 + \sigma)$$

and

$$u_i \sim \text{NB with } \mathrm{E}(u_i) = \gamma,\ \mathrm{Var}(u_i) = \gamma(1 + \sigma)$$

It can then be shown that the marginals of y_{it} are negative binomial with expectation $\mathrm{E}(y_{it}) = \lambda_{it}$ and variance[6]

$$\mathrm{Var}(y_{it}, y_{js}) = \begin{cases} \lambda_{it}(1+\sigma) & \text{for } i = j \text{ and } t = s \\ \gamma(1+\sigma) & \text{for } i = j \text{ and } t \neq s \\ 0 & \text{otherwise} \end{cases}$$

---

[5] At present, the variance term is forced to be the same for the z and u parts. Relaxing this restriction in the future is desirable. Also, this model implies a linear variance function in contrast to the quadratic variance function of the previous section.

[6] A derivation is given in the Appendix.

or, in the usual notation,

$$\operatorname{Var}(y_i) = [\Lambda_i + \gamma(\iota_T\iota_T' - I_T)](1+\sigma)$$

Hence, the covariance matrix is proportional to the covariance matrix of the Poisson-Poisson model, and over-dispersion exists whenever $\sigma > 0$. The joint PDF for cluster i is obtained along the lines of (7) where f now stands for the univariate negative binomial PDF. This model will be referred to as the **Negbin-Negbin** model.

4 Multi-factor models

An alternative to the one-factor approach, the multi-factor approach models dependence within clusters by way of correlated errors.[7] Let $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$. $\mathrm{E}(\varepsilon_i\varepsilon_i') = D$ is left unrestricted. In the normal linear model it is relatively simple to directly estimate the $t(t+1)/2$ elements of D, for instance by maximum likelihood, although this option is rarely pursued in practice (most likely, because the one factor model has an intutitive interpretation as a model for individual specific unobserved heterogeneity that the multi-factor model does not have).

Estimation of multi-factor models is more difficult in the context of non-linear count data models. The basic Poisson model is augmented by correlated individual and time specific heterogeneity components ε_{it}. The marginals are obtained by integrating the joint PDF over the multivariate (T-dimensional) distribution of ε_i. The generic model with multiplicative multivariate normal heterogeneity structure can be written as follows.

Counts are Poisson distributed conditionally on the random effects $u_{it} = e^{\varepsilon_{it}}$, i.e.,

$$y_{it}|\beta, \varepsilon_{it} \sim \operatorname{Poisson}(\lambda_{it}e^{\varepsilon_{it}}),$$

---

[7] "Correlated" errors refers to correlation of errors *within clusters*, not between errors and regressors x

where

$$\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})' \sim MVN_T(0, D),$$

and MVN_T denotes the T-variate normal distribution. Hence, $\exp(\varepsilon_i)$ has a multivariate log-normal distribution with mean $\mathrm{E}[\exp(e_{ij})] = \exp(0.5\sigma_j^2)$ and covariances

$$v_{jk} = \mathrm{Cov}[\exp(e_{ij})\exp(e_{ik})] = [\exp(\rho_{jk}\sigma_j\sigma_k) - 1]\exp[0.5(\sigma_j^2 + \sigma_k^2)] \tag{9}$$

(See Johnson and Kotz, 1972, p. 20). Let V have typical element $[v_{jk}]$. Then

$$\mathrm{Var}(y_i) = \Lambda_i + \Lambda_i V \Lambda_i'$$

This model allows for both positive and negative covariances. In particular, (9) is negative, positive, or zero if and only if the bivariate coefficient of correlation ρ_{jk} is negative, positive, or zero, respectively.

Under conditional independence, we obtain the likelihood function for observations on the ith cluster

$$f(y_i|x_i, \beta, \varepsilon_i) = \prod_{t=1}^{T} p(y_{it}|x_i, \beta, \varepsilon_i)$$

and

$$f(y_i, \varepsilon_i|x_i, \beta, D) = f(y_i|x_i, \beta, \varepsilon_i)\,\phi(\varepsilon_i|0, D)$$

is the joint density of (y_i, ε_i), where p is the Poisson mass function with conditional mean $\lambda_{it}e^{\varepsilon_{it}}$ and $\phi(\varepsilon_i|0, D)$ is the density of the normal distribution with variance covariance matrix D. Each likelihood contribution may be written in explicit form as

$$\begin{aligned} L(y_i|x_i, \beta, D) &= \int \cdots \int \left[\prod_{t=1}^{T} \frac{\exp(-\exp(x_{it}'\beta + \varepsilon_{it}))(\exp(x_{it}'\beta + \varepsilon_{it}))^{y_{it}}}{\Gamma(y_{it} + 1)}\right] \\ &\times \phi(\varepsilon_{i1}, \ldots, \varepsilon_{iT}; D)\mathrm{d}\varepsilon_{i1} \cdots \mathrm{d}\varepsilon_{iT} \end{aligned} \tag{10}$$

The likelihood function of the parameters given $y = (y_1, \ldots, y_n)$ is then the product of the n likelihood contributions $L_i(y_i|\beta, D)$:

$$\begin{aligned} L(y|\beta, D) &= \prod_{i=1}^{n} \int f(y_i, \varepsilon_i|\beta, D)\, d\varepsilon_i \\ &\equiv \prod_{i=1}^{n} L_i(y_i|\beta, D), \end{aligned} \tag{11}$$

The intractability of the likelihood function arises from the difficulty of evaluating the multidimensional integral.

5 Simulated Maximum Likelihood

One possible approach to estimate the model parameters is to replace $L(y_i|x_i, \beta, D)$ by a consistent estimator $\hat{L}(y_i|x_i, \beta, D)$ and maximize the resulting simulated likelihood function (See Gourieroux and Monfort, 1993). This approach is particularly easy to implement whenever a relevant conditioning exists, i.e., when $f(y_i|x_i, \beta, D, u_i)$ has closed form and the distribution of u_i is known and, in particular, does not depend on further parameters. In such a situation, the actual density

$$f(y_i|x_i, \beta, D) = \int f(y_i|x_i, \beta, D, u_i) g(u_i) \mathrm{d}u_i$$

can be replaced by a simulated density

$$\hat{f}(y_i|x_i, \beta, D) = \frac{1}{H} \sum_{h=1}^{H} f(y_i|x_i, \beta, D, u_{ih})$$

where u_{ih} are random number draws from $g(u)$ and H is the number of simulations per observation. The draws $[u_{ih}]$ are kept constant over the iterations of the ML routine in order to avoid discontinuities in the log-likelihood function. The objective function is then

$$\hat{L}(\beta, D) = \sum_{i=1}^{N} \log \hat{f}(y_i|x_i, \beta, D)$$

and the Simulated Maximum Likelihood estimator

$$\text{argmax}_{\beta,D}\hat{L}(\beta, D)$$

is consistent and asymptotically efficient whenever H and N tend to infinity in such a way that $\sqrt{N}/H$ tends to zero (Gourieroux and Monfort, 1993).

In the present context, let $u_{ih} \sim MVN(0, I_T)$ and $v_{ih}(D) = \text{chol}(D)'u_{ih}$. Then

$$\hat{f}(y_i|x_i, \beta, D) = \frac{1}{H}\sum_{h=1}^{H}\left[\prod_{t=1}^{T}\frac{\exp(-\exp(x'_{it}\beta + v_{ith}))(\exp(x'_{it}\beta + v_{ith}))^{y_{it}}}{\Gamma(y_{it}+1)}\right]$$

6 Application

Diggle, Liang, and Zeger (1994) provide data on seizure counts (y_{it}) for each of 59 epileptics, some of whom are treated with a drug progabide (two observation are eliminated from the dataset because of the "unusually high pre- and post-randomization seizure counts"). In this illustration, I use data for the base period and two post-treatment periods ($T = 3$). The covariates are

$$x_{i2} = \begin{cases} 1 & \text{if} \quad \text{visit 1 or 2} \\ 0 & \text{if} \quad \text{baseline} \end{cases}$$

$$x_{i3} = \begin{cases} 1 & \text{if} \quad \text{treatment group} \\ 0 & \text{if} \quad \text{control} \end{cases}$$

and

x_{i4}, an interaction between x_{i2} and x_{i3}

Following Diggle, Liang, and Zeger, the counts are modelled by a Poisson link with mean

$$\log E(y_{it}|\beta) = \beta_1 + \beta_2 x_{it2} + \beta_3 x_{it3} + \beta_4 x_{it2} x_{it3}$$

As the periods are of different lenght (8 weeks for the base period and 2 weeks each for the post-treatment visits), β_2 accounts both for any genuine difference in the seizure rates before and after treatment for the control group, in addition to the effect of the longer base period. For instance, if seizure counts were strictly proportional to the length of the period, we would expect an estimated coefficient of $\ln(1/4) = -1.39$. As the assignment to the drug was randomized, we would expect β_3 to be close to zero. Finally, the coefficient β_4 measures the treatment effect. A negative value corresponds to a greater reduction (or smaller increase) in the seizure counts for the treatment group, and thereby indicates that the treatment was effective.

Table 1 gives the maximum likelihood estimates for the standard Poisson model without random effects, and for the two additive random effects Poisson models that were introduced above (Standard errors in parentheses).[8]

The Poisson parameters in the first column of Table 1 show that, as expected, β_2 is negative and close to the value that would give strict proportionality to the period at risk. The treatment effect is negative (i.e., a reduction in the expected number of epileptic seizures by 31 percent) and significant. The two random effects models produce quite similar parameter estimates. These models illustrate the importance of allowing for individual specific heterogeneity. However, the estimated treatment effect decreases in the Poisson-Poisson and Negbin-Negbin models (in absolute value), and is insignificant in the latter.

Both t- and likelihood ratio tests clearly reject the restriction $\gamma = 0$ of the Poisson model against the Poisson-Poisson model. In the Poisson-Poisson model, the value of 3.019 gives the within cluster covariance. In this model, the conditional variance and expectation are set equal, an assumption that seems not to fit well with the data, as the Negbin-Negbin model that separately estimates the variance and covariance terms, has a much higher log-likelihood value. Note that the Negbin-Negbin model implies a linear variance function

---

[8] The data and the programs (written in Gauss) are available from the author upon request.

whereas the models that are shown in Table 2 all imply a quadratice variance function.

The Table displays the results for the three models with multiplicative error. In the Poisson-Gamma model, the multiplicative individual specific heterogeneity term is Gamma distributed. The Poisson normal model has the same structure, only that the heterogeneity term is now log-normally distributed. Finally, the last column gives the Poisson model with correlated errors (Poisson-MVN). While the Poisson-Gamma model has closed form expressions for the marginal distribution $y|x$ the other two models don't, and they were both estimated by Simulated Maximum Likelihood with 300 draws for each observation.

If ε_{ii} is normally distributed with mean 0 and variance d_{ii}, then $\exp(\varepsilon_{ii})$ has a log-normal distribution with variance $\epsilon^{d_{ii}}(\epsilon^{d_{ii}} - 1)$. For instance, the estimated variances of the heterogeneity term in the Poisson-MVN model range from 0.65 to 2.50. The negative treatment effect in all cases indicates a significant reduction in the number of seizures for the treated group. The Poisson-Gamma and Poisson-normal models yield almost identical log-likelihood values whereas the Poisson-MVN model provides a substantially better "fit". A formal comparison of the Poisson-MVN model against the Poisson-normal model based on a log-likelihood ratio test leads to a rejection of the one-factor restriction at any conventinal significance level.

7 Discussion

This paper has explored various issues that arise in the specification and estimation of count data models with random effects. While the discussion was centered around panel data, i.e. repeated observations on the same cross-sectional unit over time, all of the presented methods equally well can be used to model multivariate counts of systems of seemingly unrelated count data regressions.

It is important to stress that the presented methods relied on the assumption of independence between the cluster specific random effect and the regressors x. In situations where the validity of this assumption is questionable one would need to use alternative methods that allow for conditional inference as discussed in Hausman, Hall and Griliches (1984) and more recently in Blundell, Griffith and Van Reenan (1995). It is possible to test the validity of the random effects assumption through a Hausman type test.

Appendix

The negative binomial distribution has probability generating function

$$\mathcal{P}(s) = \mathrm{E}(s^X) = [1 + \theta(1 - s)]^{-\alpha}$$

and PDF

$$f(y) = \frac{\Gamma(\alpha + y)}{\Gamma(\alpha)\Gamma(y + 1)} \left(\frac{1}{1 + \theta}\right)^{\alpha} \left(\frac{\theta}{1 + \theta}\right)^{y},$$

The mean and variance are given by

$$\mathrm{E}(X) = \mathcal{P}'(1) = \alpha\theta \tag{12}$$

and

$$\mathrm{Var}(X) = \mathcal{P}''(1) + \mathcal{P}'(1) - [\mathcal{P}'(1)]^2 = \alpha\theta(1 + \theta) \,. \tag{13}$$

Assume that the two random variables Y_1 and Y_2 are distributed with $NB(\alpha_1, \theta_1)$ and $NB(\alpha_2, \theta_2)$ respectively. Under independence,

$$\begin{aligned} \mathcal{P}_{Y_1+Y_2}(s) &= \mathcal{P}_{Y_1}(s)\mathcal{P}_{Y_2}(s) \\ &= [1 + \theta_1(1 - s)]^{-\alpha_1}[1 + \theta_2(1 - s)]^{-\alpha_2} \end{aligned}$$

(See Feller, 1968). Next consider the following two common parametrizations:

Negbin I $\alpha = \lambda/\theta$. Then $\mathrm{E}(Y) = \lambda$ and $\mathrm{Var}(Y) = \lambda(1+\theta)$.

Negbin II $\theta = \lambda/\alpha$. Then $\mathrm{E}(Y) = \lambda$ and $\mathrm{Var}(Y) = \lambda(1+\lambda/\alpha)$.

We see that the Negbin 1 model is closed under convolution if $\theta_1 = \theta_2$ (whereas the Negbin II model isn't), since

$$\begin{aligned} \mathcal{P}_{Y_1+Y_2}(s) &= [1+\theta(1-s)]^{-\lambda_1/\theta}[1+\theta(1-s)]^{-\lambda_2/\theta} \\ &= [1+\theta(1-s)]^{-(\lambda_1+\lambda_2)/\theta} \end{aligned}$$

and $Y_1 + Y_2$ is negative binomial distributed with mean $\lambda_1 + \lambda_2$ and variance $(\lambda_1+\lambda_2)(1+\theta)$.

To obtain the random effects negative binomial model, let $z_{it} \sim \mathrm{NB}(\lambda_{it} - \gamma, \sigma)$ and $u_i \sim \mathrm{NB}(\gamma, \sigma)$. Then $y_{it} \sim \mathrm{NB}(\lambda_{it}, \sigma)$ with PDF

$$f(y_{it}) = \frac{\Gamma(\lambda_{it}/\sigma + y_{it})}{\Gamma(\lambda_{it}/\sigma)\Gamma(y_{it}+1)} \left(\frac{1}{1+\sigma}\right)^{\lambda_{it}/\sigma} \left(\frac{\sigma}{1+\sigma}\right)^{y_{it}} .$$

References

Blundell, R. R. Griffith and J. Van Reenan, 1995, Dynamic Count Data Models of Technological Innovation, *Economic Journal* 105, 333-344.

Brännäs, K. and Johannson, 1996, Panel Data Regression for Counts, *Statistical Papers*, 37, 191-213.

Chib, S. and E. Greenberg 1995, Understanding the Metropolis-Hastings Algorithm, *American Statistician*, 49, 327-335.

Chib, S., E. Greenberg and R. Winkelmann 1998, Posterior Simulation and Bayes Factors in Panel Count Data Models, *Journal of Econometrics*, 86, 33-54.

Cincera, M. 1997, Patents, R&D, and Technological Spillovers at the Firm Level: Some Evidence from Econometric Count Models for Panel Data, *Journal of Applied Econometrics*, 12, 265-80.

Diggle, P.J., K.-Y. Liang and S.L. Zeger 1994, *Analysis of Longitudinal Data*, Oxford University Press.

Feller, W. 1968, *An Introduction to Probability Theory and Its Applications*, Vol. 1, Wiley.

Gourieroux, C. , A. Monfort and A. Trognon, 1984, Pseudo maximum likelihood methods: Applications to Poisson models, *Econometrica*, 52, 701-720.

Gourieroux, C. and A. Monfort, 1993, Simulation based inference, *Journal of Econometrics*, 59, 5-33.

Gurmu, S. and J. Elder 1998, Estimation of multivariate count regression models with applications to health care utilization, mimeo, University of Virginia.

Hausman, J., B.H. Hall, and Z. Griliches 1984, Econometric Models for Count Data with an Application to the Patents-R&D Relationship, *Econometrica*, 52, 909-938.

Hsiao, C. 1986, *Analysis of Panel Data*, Cambridge University Press.

Johnson, N.L. and S. Kotz 1972, *Distribution in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons.

Jung, R.C. and R. Winkelmann 1993, Two aspects of labor mobility: a bivariate Poisson regression approach, *Empirical Economics* 18, 543-556.

King, G. 1989, A seemingly unrelated Poisson regression model, *Sociological Methods and Research* 17, 235-255.

Winkelmann, R. 1997, *Econometric Analysis of Count Data*, Heidelberg, Springer.

Winkelmann, R. 1998, Count data models with selectivity, *Econometric Reviews* (in press).

Table 1. Poisson random effects models

	Poisson	Poisson-Poisson	Negbin-Negbin
β_1	3.326	3.309	3.242
	(0.036)	(0.035)	(0.095)
β_2	-1.255	-1.295	-1.161
	(0.061)	(0.058)	(0.134)
β_3	-0.005	0.024	0.091
	(0.050)	(0.047)	(0.123)
β_4	-0.273	-0.183	-0.156
	(0.087)	(0.078)	(0.169)
γ		3.019	4.151
		(0.302)	(0.789)
σ			6.335
			(0.925)
Log-Likelihood	-987.79	-900.02	-562.04

Table 2. Poisson Mixture Models (Simulated Maximum Likelihood)

	Poisson-Gamma	Poisson-normal	Poisson-MVN
β_1	3.325	2.975	3.114
	(0.093)	(0.125)	(0.083)
β_2	-1.255	-1.255	-1.385
	(0.060)	(0.060)	(0.101)
β_3	-0.005	0.216	-0.115
	(0.060)	(0.162)	(0.089)
β_4	-0.273	-0.273	-0.333
	(0.086)	(0.087)	(0.126)
γ	2.468	0.408	
	(0.464)	(0.070)	
d_{11}			0.385
			(0.045)
d_{12}			0.286
			(0.059)
d_{22}			0.584
			(0.105)
d_{13}			0.362
			(0.053)
d_{23}			0.573
			(0.075)
d_{33}			0.767
			(0.066)
Log-Likelihood	-579.87	-579.95	-541.58

LIST OF DISCUSSION PAPERS

No. 9311 Optimal Long-Run Business Immigration Under Differential Savings Functions, by A. E. Woodfield and K-L. Shea.
No. 9312 The Welfare Cost of Taxation in New Zealand Following Major Tax Reforms, by P. McKeown and A. Woodfield.
No. 9313 The Power of the Goldfeld-Quandt Test when the errors are autocorrelated, by J.P. Small and R.J. Dennis.
No. 9314 The Nucleolus Strikes Back, by M. Carter and P. Walker.
No. 9315 The Output-Inflation Tradeoff in the United States: New evidence on the New Classical vs. New Keynesian Debate, by Alfred V. Guender.
No. 9401 Insurance Market Equilibrium and the Welfare Costs of Gender-Neutral Insurance Pricing under Alternative Regulatory Regimes by Alan E. Woodfield.
No. 9402 Labour Market Signalling and the Welfare Costs of Regulated Insurance Market Equilibria under Gender-neutral Pricing, by Alan E. Woodfield.
No. 9403 The New Classical Vs The New Keynesian debate On The Output – Inflation tradeoff: Evidence From Four industrialized Countries, by Alfred V. Guender.
No. 9404 Yield Spreads & Real Economic Activity: The Case of New Zealand & Australia, by Alfred V. Guender and Mathias Moersch.
No. 9405 Periodic Integration & cointegration with applications to the New Zealand Aggregate Consumption Function, by Robin Harrison and Aaron Smith.
No. 9406 Linear Programming with Mathematica, by Michael Carter.
No. 9407 Are People Really Risk Seeking When Facing Losses? by John Fountain, Michael McCosker & Dean Morris.
No. 9501 Pricing Internet: The New Zealand Experience by Michael Carter and Graeme Guthrie.
No. 9502 Long term and Short Term Aggregate Uncertainty and the Effect on Real Output, by Alfred V. Guender and Robin Young.
No. 9503 Framing and Incentive Effects on Risk Attitudes When Facing Uncertain Losses, by John Fountain, Michael McCosker and Bruce Macfarlane.
No. 9504 An Outline of the History of Game Theory, by Paul Walker.
No. 9505 A Drunk, Her Dog and a Boyfriend: An Illustration of Multiple Cointegration and Error Correction, by Aaron Smith and Robin Harrison.
No. 9506 Optimal Markup Responses to Trade Policy Reform in a Small, Open Economy: Evidence from New Zealand, by Liliana Winkelmann and Rainer Winkelmann.
No. 9507 Unemployment: Where Does It Hurt? by Liliana Winkelmann and Rainer Winkelmann.
No. 9508 Apprenticeship And After: Does It Really Matter? by Rainer Winkelmann.
No. 9601 Another Look at Work Contracts and Absenteeism, by Rainer Winkelmann.
No. 9602 Markov Chain Monte Carlo Analysis of Underreported Count Data With an Application to Worker Absenteeism, by Rainer Winkelmann.
No. 9603 Count Data Models With Selectivity, by Rainer Winkelmann.
No. 9604 Optimal Revenue Smoothing: The Case of New Zealand, by Alfred V. Guender and Kirdan Lees.
No. 9605 Posterior Simulation and Model Choice in Longitudinal Generalized Linear Models, by Siddhartha Chib, Edward Greenberg, Rainer Winkelmann.
No. 9606 Getting It Right: Superannuation and Savings in the U.S.A., by Philip Meguire.
No. 9607 Interconnection in Network Industries, by Michael Carter and Julian Wright.
No. 9608 Tariffs, Quotas and Terms-of-Trade: The Case of New Zealand, by Liliana Winkelmann and Rainer Winkelmann.

No. 9609 Credit Frictions, Debt Choice and the Transmission of Monetary Policy, by Julian Wright.

No. 9610 The Forward Premium Bias Under Different Monetary Policy Environments, by David Steele and Julian Wright.

No. 9611 Optimal Currency Denomination of Public Debt in New Zealand, by Kerryn Fowlie and Julian Wright.

No. 9701 Is There a Bank Lending Channel of Monetary Policy in New Zealand? by Alfred V. Guender.

No. 9702 Payback and the Value of Waiting to Invest, by Glenn W. Boyle and Graeme A. Guthrie.

No. 9703 Zermelo and the Early History of Game Theory, by Ulrich Schwalbe and Paul Walker.

No. 9801 Demon or Saviour? An Assessment of the ISO 9000 Quality Assurance Standards, by Philip Gunby.

No. 9802 Market-Implemented Monetary Policy with Open Mouth Operations, by Graeme Guthrie and Julian Wright.

No. 9803 Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration, by James G. MacKinnon, Alfred A. Haug and Leo Michelis.

No. 9804 Conflicts Among Tests for Cointegration, by Allan W. Gregory and Alfred A. Haug.

No. 9805 Alternative Monetary Policy Rules for Small Open Economies, by Richard T. Froyen and Alfred V. Guender

No. 9806 The Labour Market Outcomes of New Zealand's Old and New Immigrants, by Liliana Winkelmann and Rainer Winkelmann

No. 9807 Contracted Workdays and Absence, by Tim Barmby, Michael Nolan, Rainer Winkelmann

No. 9808 Is Job Stability Declining in Germany? Evidence from Count Data Models, by Rainer Winkelmann and Klaus F. Zimmermann

No. 9809 The Labor Market Performance of European Immigrants in New Zealand in the 1980s and 1990s, by Rainer Winkelmann

No. 9810 The Economic Benefits of Schooling in New Zealand: Comment and Update, by Rainer Winkelmann

No. 9811 Random Effects Models for Panel Count Data, by Rainer Winkelmann

* Copies of these Discussion Papers may be obtained for $4 (including postage, price changes occasionally) each by writing to the Secretary, Department of Economics, University of Canterbury, Christchurch, New Zealand. A list of the Discussion Papers prior to 1993 is available on request.

This paper is circulated for discussion and comments. It should not be quoted without the prior approval of the author. It reflects the views of the author who is responsible for the facts and accuracy of the data presented. Responsibility for the application of material to specific cases, however, lies with any user of the paper and no responsibility in such cases will be attributed to the author or to the University of Canterbury.