POLS571 - Longitudinal Data Analysis
December 11, 2001
Panel/TSCS Models for Event Count Data

1 Introduction

There are basically three kinds of panel/TSCS models for event count data: fixed-effects, random-effects and marginal (e.g., GEE) models.

2 Event Count Data

Data where the response variable is a nonnegative integer count – the number of occurrences of some event in a given period (or other defined unit).

The most basic model for counts of events is the Poisson:

$$Pr(Y_i = y) = \frac{exp(-\lambda_i)\lambda_i^y}{y!} \tag{1}$$

where μ_i is the expected value of Y_i and we typically introduce covariates as $\mu_i = exp(X_i\beta)$. I won't go into the Poisson distribution that much here; take the MLE class for the details. Importantly, though, the Poisson distribution has the property that $E(Y) = Var(Y)$; if this is not the case, then other models (e.g., the negative binomial model) are necessary.

3 Fixed Effects Models for Count Data

The fixed-effects Poisson model is:

$$Y_{it} \sim Poisson(\mu_{it} = \alpha_i\lambda_{it}) \tag{2}$$

where, as above, $\lambda_{it} = exp(X_{it}\beta)$. Note that this specification means that the conditional expectation of Y is:

$$
\begin{aligned}
E(Y_{it} \mid X_{it}, \alpha_i) &= \mu_{it} \\
&= \alpha_i exp(X_{it}\beta) \\
&= exp(\delta_i + X_{it}\beta) \qquad (3)
\end{aligned}
$$

where $\delta_i = ln(\alpha_i)$.

If N is small (and fixed), this model can be estimated in the standard fashion, simply by including dummy variables for each of the N units i. Thus, as with other kinds of models, fixed-effects are reasonable only when N is small relative to T, and when N is not increasing (that is, when one's asymptotics don't need to depend on N). However, Lancaster (in Cameron and Trivedi 1996) shows that there is no "incidental parameters problem" for Poisson regression, using a "concentrated likelihood (see Camon and Trivedi, pp. 281-2). One therefore needs to use a "conditional" approach, similar to that for fixed-effects logit, in which the fixed effects are conditioned on the sum of the event counts within the panel and then concentrated out of the likelihood. As with conditional fixed-effects logit, this relies on the fact that observations across units are independent (i.e., that the fixed effects completely capture the unobserved heterogeneity in the data).

In Stata, this is accomplished through the `-xtpois, fe-` command. Interpretation is standard for Poisson models, with the usual caveat that results are conditional on the values of the fixed effects. Note as well that, for overdispersed data, there is also a conditional fixed-effects negative binomial model (in Stata, `-xtnbreg, fe-`).

4 Random-Effects Models for Event Count Data

The random-effects Poisson model is similar to that for (e.g.) logit and probit, in that we assume that the α_is are distributed as some i.i.d. random variable, which we then integrate out of the likelihood:

$$
\begin{aligned}
Pr(Y_{i1} = y_{i1}, ...Y_{iT} = y_{iT}) &= \int_0^\infty Pr(Y_{i1} = y_{i1}, ...Y_{iT} = y_{iT}) f(\alpha_i) d\alpha_i \\
&= \int_0^\infty [\prod_{t=1}^{T} Pr(Y_{it} \mid \alpha_i)] f(\alpha_i) d\alpha_i \qquad (4)
\end{aligned}
$$

The simplest distribution to use for the α_is is the *gamma* with an expected value and variance equal to some parameter θ. (The gamma is the conjugate distribution for the Poisson.) The result is a likelihood (see Cameron and Trivedi p. 288) which is relatively straightforward to maximize. The random-effects Poisson model has $E(Y_{it}) = \lambda_{it}$ and $Var(Y_{it}) = \lambda_{it} + \frac{\lambda_{it}^2}{\theta}$. The Stata command is `-xtpois, re-`; and interpretation is also as with other conditional-effect models.

One can also estimate this model assuming that the α_is are distributed normally, in which case the quadrature procedure used for random-effects probit (or some other approach, e.g. MCMC, a la Chib et. al. 1998) is necessary. Note as well that there is also a random-effects negative binomial model due to Hausman, Hall and Griliches (1984), which is available only for gamma-distributed α_is and which is estimable in Stata using `-xtnbreg, re-`.

5 Marginal/GEE Models for Event Count Data

GEE models as available for both Poisson and negative binomial models, in all their forms...

6 An Example...