Count Data Models

• Count Data models are used to characterize realizations of a non-negative integer-valued random variable
• Typically concerned with low probability events and many zeroes
• The dual to count data models are duration models, focusing on the length of time between the occurrence of an event (e.g., a recreation trip or the visit to a doctor)
Examples of Count Data Models

- Examples abound in the literature (145 hits in EconLit), including
  - Frequency of doctor visits
  - Patents
  - Recreation demand
  - Takeover bids
  - Bank failures
  - Accident frequency
  - Number of loan defaults
  - Presidential appointments to the Supreme Court
  - Number of criminal offenses
  - Manufacturing Defects

Example #1: Doctors Visits
(Cameron, Trivedi, Milne, and Piggott, 1988)
Example #2: Takeover Bids
(Jaggia and Thosar, 1993)

Example #3: Recreational Trips to Lake Somerville, Texas
(Ozuna and Gomez, 1995)
Outline

• Single Equation Models
  – Poisson
    • Distributional Assumptions
    • Estimation
    • Interpretation
    • Prediction
    • Limitations
      – Overdispersion
      – Excess zeros
  – Generalizations
    • Accounting for Overdispersion
    • The Negative Binomial Specification
    • Alternative Mixture Models

• Multivariate Count Data Models

• Comparisons of Continuous and Count Data Demand Systems

Sources – Univariate Setting

• *Cameron, A., and P. Travedi (1998), Regression Analysis of Count Data, New York: Cambridge University Press, Chapter 1, Sections 3.1 to 3.5.


The Poisson Distribution

- The most basic of the count data models is based on the Poisson distribution, developed in 1837 by Poisson.
- Classic early study by Bortkiewicz (1898) analyzed the number of annual deaths in Prussian army from mule kicks.
- If $Y$ is a discrete random variable that is distributed Poisson, then
  \[
  \Pr [Y = y] = \frac{e^{-\mu} (\mu t)^y}{y!}, \quad y = 0, 1, 2, \ldots
  \]
  where
  - $\mu$ denotes the intensity or rate parameter, $\mu > 0$
  - $t$ denotes the exposure period

The Poisson Distribution (cont’d)

- Usually normalize $t=1$, so that
  \[
  \Pr [Y = y] = \frac{e^{-\mu} (\mu)^y}{y!}, \quad y = 0, 1, 2, \ldots
  \]
- A key feature of the Poisson distribution is that
  \[
  E[Y] = Var[Y] = \mu
  \]
  referred to as the equidispersion property.
- Additivity property also holds; i.e., if $Y_i \sim i.i.d. P[\mu_i]$, then
  \[
  S = \sum_i Y_i \sim P \left[ \sum_i \mu_i \right]
  \]
“Rare Events” Interpretation

- Consider a sequence of \( n \) independent Bernoulli trials, each with success probability of \( \pi \).

- Let \( Y_{n,\pi} \) denote the total number of successes

\[ P_{n,\pi}(k) = P[Y_{n,\pi} = k] = \binom{n}{k} \pi^k (1 - \pi)^{n-k} \]

and \( \mu \equiv n\pi > 0 \)

then, for a fixed \( \mu \),

\[ \lim_{n \to \infty, \pi \to 0} P_{n,\pi}(k) = \lim_{n \to \infty, \pi \to 0} P_\mu(k) = \mu^k e^{-\mu} \]

The Poisson Regression Model

- The standard Poisson regression model results if the intensity parameter is assumed to be a function of observed characteristics.

- Let \( y_i \) denote the observed number of occurrences of the event of interest. The Poisson regression model assumes that

\[ f(y_i | x_i) = \frac{e^{-\mu(x_i; \beta)} \mu^{y_i}}{y_i!} \]

so that

\[ E[y_i | x_i] = Var[y_i | x_i] = \mu(x_i; \beta) \]
The Log-Linear Poisson Regression Model

- The most frequently used functional form assumes that
  \[ \mu(x_i; \beta) = \exp(\beta'x_i) \]
  insuring that \( \mu > 0 \), as required.
- The resulting nonlinear regression model is heteroskedastic, with
  \[ y_i = E[y_i | x_i] + (y_i - E[y_i | x_i]) \]
  \[ = \exp(\beta'x_i) + \varepsilon_i \]
  and
  \[ \text{Var}[\varepsilon_i | x_i] = \exp(\beta'x_i) \]

Estimation

- While either LS or WLS can be applied to the nonlinear regression model, ML is typically employed
- The log-likelihood function is given by
  \[ LL(y, x; \beta) = \sum_{i=1}^{N} \{ y_i \beta'x_i - \exp(\beta'x_i) - \ln(y_i!) \} \]
  with first order conditions
  \[ \sum_{i=1}^{N} [y_i - \exp(\beta'x_i)]x_i = 0 \]
  and
  \[ \frac{\partial^2 LL}{\partial \beta \partial \beta'} = -\sum_{i=1}^{N} \exp(\beta'x_i)x_ix_i' \]
  i.e., \( LL \) is globally concave
Properties of MLE

- Consistency of the $\hat{\beta}_{ML}$ requires only that conditional mean of $y_i$ is correctly specified; i.e., it need not be Poisson distributed, we need only
  \[ E[y_i | x_i] = \exp(\beta'x_i) \]

  however, the resulting ML standard errors will be incorrect unless $y_i$ is Poisson.

- The standard errors can be corrected

- More efficient estimates can be obtained if equidispersion does not hold

Properties of MLE (cont’d)

- If the counts do follow a Poisson process, then
  \[ \hat{\beta}_{ML} \sim N(\beta, \Omega_{ML}) \]
  \[ \Omega_{ML} = -E\left[ \frac{\partial^2 LL}{\partial \beta \partial \beta'} \right] = \left( \sum_{i=1}^{N} \exp(\beta'x_i)x_i' \right)^{-1} \]

  Alternatively, one can replace $\Omega_{ML}$ with
  \[ \Omega_{MLOP} = \left( \sum_{i=1}^{N} [y - \exp(\beta'x_i)]^2 x_i' \right)^{-1} \]
Overdispersion

- One of the chief criticisms of the Poisson specification is that it assumes equidispersion; i.e.,

\[ E[y_i \mid x_i] = Var[y_i \mid x_i] \]

- In practice, overdispersion tends to hold; i.e.,

\[ E[y_i \mid x_i] < Var[y_i \mid x_i] \]

- This manifests itself in terms of fatter tails and a greater number of zeros than would characterize a distribution with equidispersion

Example #3: Recreational Trips to Lake Somerville, Texas
(Ozuna and Gomez, 1995)
Solutions to Overdispersion

- Specify an alternative functional form for the conditional variance; e.g.,

\[
V[y_i | x_i] = \omega_i = \omega(\mu_i, \alpha)
\]

Given that this specification being correct, the appropriate asymptotic standard errors for \( \hat{\beta}_{ML} \) can be constructed.

- Specify an alternative distributional assumption (e.g., negative binomial) and apply ML estimation procedures.

Poisson Pseudo-ML

- Given that the mean and variance have been correctly specified, and \( \hat{\beta}_{ML} \) derived from the Poisson model, it can be shown that

\[
\hat{\beta}_{ML} \sim N(\beta, \Omega_{PML})
\]

where

\[
\Omega_{PML} = \left( \sum_{i=1}^{N} \mu_i x_i x_i' \right)^{-1} \left( \sum_{i=1}^{N} \omega_i x_i x_i' \right) \left( \sum_{i=1}^{N} \mu_i x_i x_i' \right)^{-1}
\]
NB1 Model

- One common specification is:

\[ V[y_i | x_i] = \omega(\mu_i, \alpha) \]
\[ = (1 + \alpha) \mu_i \]
\[ = \phi_i E[y_i | x_i] \]

- \( \hat{\Omega}_{PML} \) is then constructed using

\[ \hat{\phi}_{NB1} = \frac{1}{n-k} \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2}{\mu_i} \]

NB2 Model

- A second specification assumes that:

\[ V[y_i | x_i] = \omega(\mu_i, \alpha) \]
\[ = \mu_i + \alpha \mu_i^2 \]

- \( \hat{\Omega}_{PML} \) is then constructed using

\[ \hat{\alpha}_{NB2} = \frac{1}{n-k} \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2 - \hat{\mu}_i}{\hat{\mu}_i^2} \]
RS (robust sandwich) Model

- Finally, we can proceed without specifying an exact functional form for the variance
- \( \hat{\Omega}_{PML} \) is then constructed using

\[
\hat{\Omega}_{PML} = \left( \sum_{i=1}^{N} \mu_i x_i \right)^{-1} \left( \sum_{i=1}^{N} y_i^2 \right) \left( \sum_{i=1}^{N} \mu_i x_i \right)^{-1}
\]

Example: Doctor Visits

- Cameron, Trivedi, Milne, and Piggott (1998) studied the number of visits to doctors during a 2 week period
- Explanatory variables include
  - Gender (Sex=1 for females)
  - Age (divided by 100)
  - Private health insurance dummy variable (LEVYPLUS)
  - Low income free government health insurance (FREEPOOR)
  - Old age/disability free government health insurance (FREEREPA)
  - Number of recent illnesses (ILLNESS)
  - Number of reduced activity days (ACTDAYS)
  - Health score (HSCORE)
  - Chronic conditions (CHCOND1 and CHCOND2)
- The mean number of visits is 0.302, whereas the variance is 0.637, suggesting overdispersion
## Parameter Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ML</td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>-2.224</strong></td>
<td>0.190</td>
</tr>
<tr>
<td>Sex</td>
<td><strong>0.157</strong></td>
<td>0.056</td>
</tr>
<tr>
<td>Age</td>
<td>1.056</td>
<td>1.001</td>
</tr>
<tr>
<td>Age²</td>
<td>-0.849</td>
<td>1.078</td>
</tr>
<tr>
<td>Income</td>
<td><strong>-0.205</strong></td>
<td>0.088</td>
</tr>
<tr>
<td>LevyPlus</td>
<td>0.123</td>
<td>0.072</td>
</tr>
<tr>
<td>FreePoor</td>
<td><strong>-0.440</strong></td>
<td>0.180</td>
</tr>
<tr>
<td>FreeREPA</td>
<td>0.080</td>
<td>0.092</td>
</tr>
<tr>
<td>Illness</td>
<td><strong>0.187</strong></td>
<td>0.018</td>
</tr>
<tr>
<td>Actdays</td>
<td><strong>0.127</strong></td>
<td>0.005</td>
</tr>
<tr>
<td>Hscore</td>
<td><strong>0.030</strong></td>
<td>0.010</td>
</tr>
<tr>
<td>ChCond1</td>
<td>0.114</td>
<td>0.066</td>
</tr>
<tr>
<td>ChCond2</td>
<td>0.141</td>
<td>0.083</td>
</tr>
</tbody>
</table>

## Negative Binomial Distribution

- The previous methods rely upon the Poisson-based ML estimates, correcting for overdispersion.
- However, they are inefficient in that they do not use the overdispersion characteristic in estimation.
- An alternative is to rely upon a parametric distribution without equidispersion; e.g., the Negative Binomial

\[
f(y_i \mid \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1) \Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i}\right)^{y_i}
\]

This reduces to the Poisson distribution if \(\alpha_i = 0\)
ML Estimation for NB Specification

- The corresponding log-likelihood in this case is given by

$$LL(y, x; \beta, \alpha) = \sum_{i=1}^{N} \left\{ \sum_{j=0}^{y_{i}-1} \ln\left(j + \alpha^{-1}\right) - \ln\left(y_{i}!\right) + \left(y_{i} + \alpha^{-1}\right) \ln\left(1 + \alpha \exp(\beta' x_{i})\right) + y_{i} \ln \alpha + y_{i} \beta' x_{i}\right\}$$

The corresponding first order conditions are straightforward and the resulting variance-covariance matrix is block diagonal.

Parameter Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Poisson</th>
<th></th>
<th>NB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.224</td>
<td>0.207</td>
<td>-2.190</td>
<td>0.222</td>
</tr>
<tr>
<td>Sex</td>
<td>0.157</td>
<td>0.062</td>
<td>0.217</td>
<td>0.066</td>
</tr>
<tr>
<td>Age</td>
<td>1.056</td>
<td>1.112</td>
<td>-0.216</td>
<td>1.333</td>
</tr>
<tr>
<td>Age²</td>
<td>-0.849</td>
<td>1.210</td>
<td>0.609</td>
<td>1.380</td>
</tr>
<tr>
<td>Income</td>
<td>-0.205</td>
<td>0.096</td>
<td>-0.142</td>
<td>0.098</td>
</tr>
<tr>
<td>LevyPlus</td>
<td>0.123</td>
<td>0.077</td>
<td>0.118</td>
<td>0.085</td>
</tr>
<tr>
<td>FreePoor</td>
<td>-0.440</td>
<td>0.188</td>
<td>-0.497</td>
<td>0.175</td>
</tr>
<tr>
<td>FreeREPA</td>
<td>0.080</td>
<td>0.102</td>
<td>0.145</td>
<td>0.117</td>
</tr>
<tr>
<td>Illness</td>
<td>0.187</td>
<td>0.021</td>
<td>0.214</td>
<td>0.026</td>
</tr>
<tr>
<td>Actdays</td>
<td>0.127</td>
<td>0.006</td>
<td>0.144</td>
<td>0.008</td>
</tr>
<tr>
<td>Hscore</td>
<td>0.030</td>
<td>0.012</td>
<td>0.038</td>
<td>0.014</td>
</tr>
<tr>
<td>ChCond1</td>
<td>0.114</td>
<td>0.071</td>
<td>0.099</td>
<td>0.077</td>
</tr>
<tr>
<td>ChCond2</td>
<td>0.141</td>
<td>0.092</td>
<td>0.190</td>
<td>0.095</td>
</tr>
<tr>
<td>LL</td>
<td>-3355.5</td>
<td></td>
<td>-3198.7</td>
<td></td>
</tr>
</tbody>
</table>
Indications of Overdispersion

- Comparing the unconditional mean and variance of the count data will give an indication of overdispersion, but will tend to overstate its degree, since
  \[
  \frac{\text{Var}[y_i]}{E[y_i]} > \frac{\text{Var}[y_i | x_i]}{E[y_i | x_i]}
  \]

- Cameron and Trivedi (1998) suggest a good rule of thumb is that overdispersion is an issue if
  \[
  \frac{\text{Var}[y_i]}{E[y_i]} > 2
  \]

Tests of Overdispersion

- One can estimate both the Negative Binomial and Poisson regression models and test for overdispersion
- The distribution of the test statistics are not standard due to the restriction that \( \alpha > 0 \)
  - LR test statistic for a level of \( \delta \) is \( \chi^2_{1-2\delta} (1) \)
  - Wald test is implemented as a one-sided \( t \) test
- One can also test for overdispersion by estimating the Poisson model and running the auxiliary regression
  \[
  \frac{(y_i - \hat{\mu}_i)^2 - y_i}{\mu_i} = \alpha \hat{\mu}_i + u_i,
  \hat{\mu}_i = \exp \left( \hat{\beta} x_i \right)
  \]
Model Interpretation

- As with all nonlinear regression models, the parameters do not indicate the marginal impact of an explanatory variable.
- For the case of an exponential conditional mean; i.e.,

\[ E[y_i | x_{ik}] = \exp(\beta'x_i) \]

We have

\[ \frac{\partial E[y_i | x_{ik}]}{\partial x_{ik}} = \beta_k \exp(\beta'x_i) \]

or in elasticity form

\[ \frac{\partial \ln E[y_i | x_{ik}]}{\partial \ln x_{ik}} = \beta_k x_{ik} \]

Overall Response

- One is often interested in the overall response in the population.
- One overall measure of response for the same level change is

\[ \frac{1}{N} \sum_{i=1}^{N} \beta_k \exp(\beta'x_i) = \beta_k \left[ \frac{1}{N} \sum_{i=1}^{N} \exp(\beta'x_i) \right] \]

For the Poisson ML estimates, when there is a constant in the model, this reduces considerable, since

\[ \frac{1}{N} \sum_{i=1}^{N} \exp(\beta'x_i) = \frac{1}{N} \sum_{i=1}^{N} y_i = \bar{y} \Rightarrow \frac{1}{N} \sum_{i=1}^{N} \frac{\partial E[y_i | x_i]}{\partial x_{ik}} = \beta_k \bar{y} \]
Overall Response (cont’d)

- Alternatively, if you are looking at the same percentage change in the explanatory variable, then

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\partial \ln E[y_i | x_i]}{\partial \ln x_{ik}} = \frac{1}{N} \sum_{i=1}^{N} \beta_k x_{ik} = \beta_k \bar{x}_{ik}$$

- As we have seen in other nonlinear regression models, one does not want to rely on the response as the average characteristic

$$\frac{\partial E[y_i | x_i]}{\partial x_{ik}} \bigg|_{\tau} = \beta_k \exp(\beta^T \bar{x}) < \frac{1}{N} \sum_{i=1}^{N} \frac{\partial E[y_i | x_i]}{\partial x_{ik}}$$

Miscellaneous Notes on Response Measures

- While the parameters do not indicate the marginal impact, their relative sizes do indicate the relative strength of each variable’s effect; i.e.,

$$\begin{pmatrix}
\frac{\partial E[y_i | x_i]}{\partial x_{ik}} \\
\frac{\partial E[y_i | x_i]}{\partial x_{ij}}
\end{pmatrix} = \begin{pmatrix}
\beta_k \\
\beta_j
\end{pmatrix}$$

- Finally, while the marginal effects are nonlinear functions of the estimated parameters, their standard deviation can be constructed by simulation or bootstrapping
### Variable Effects: Doctor Visits

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>Average</th>
<th>At Avg.</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.224</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.157</td>
<td>0.047</td>
<td>0.035</td>
<td>0.082</td>
</tr>
<tr>
<td>Age</td>
<td>1.056</td>
<td>0.319</td>
<td>0.241</td>
<td>0.430</td>
</tr>
<tr>
<td>Age²</td>
<td>-0.849</td>
<td>-0.256</td>
<td>-0.193</td>
<td>-0.176</td>
</tr>
<tr>
<td>Income</td>
<td>-0.205</td>
<td>-0.062</td>
<td>-0.047</td>
<td>-0.120</td>
</tr>
<tr>
<td>LevyPlus</td>
<td>0.123</td>
<td>0.037</td>
<td>0.028</td>
<td>0.055</td>
</tr>
<tr>
<td>FreePoor</td>
<td>-0.440</td>
<td>-0.133</td>
<td>-0.100</td>
<td>-0.019</td>
</tr>
<tr>
<td>FreeREPA</td>
<td>0.080</td>
<td>0.024</td>
<td>-0.018</td>
<td>0.017</td>
</tr>
<tr>
<td>Illness</td>
<td>0.187</td>
<td>0.056</td>
<td>0.043</td>
<td>0.268</td>
</tr>
<tr>
<td>Actdays</td>
<td>0.127</td>
<td>0.038</td>
<td>0.029</td>
<td>0.109</td>
</tr>
<tr>
<td>Hscore</td>
<td>0.030</td>
<td>0.009</td>
<td>0.007</td>
<td>0.037</td>
</tr>
<tr>
<td>ChCond1</td>
<td>0.114</td>
<td>0.034</td>
<td>0.026</td>
<td>0.046</td>
</tr>
<tr>
<td>ChCond2</td>
<td>0.141</td>
<td>0.043</td>
<td>0.032</td>
<td>0.016</td>
</tr>
</tbody>
</table>

### Problems with the Poisson Regression Model

- **Overdispersion**
  - Typically attributed to omitted and/or unobserved sources of heterogeneity
  - Can use Poisson ML estimates with corrected standard errors
  - Alternatively, can use models without equidispersion
    - Negative Binomial
    - mixed Poisson
- **Truncation (especially zero truncation)**
- **Excess zeros**
- **Serial correlation**
Mixed Poisson

- Mixture Models in the context of count data models are much like in the mixed logit context – capturing unobserved heterogeneity
- Suppose
  \[ E[y_i | x_i, v_i] = \mu_i v_i \]
  where
  \[ \mu_i = \exp(\beta' x_i) \]
  and
  \[ v_i = \exp(\varepsilon_i) \]
  with \( \varepsilon_i \) capturing unobserved omitted explanatory variables

Mixed Poisson (cont’d)

- If, for example, one assumes that \( \varepsilon_i \) is iid with
  \[ E[v_i | x_i] = 1 \]
  and
  \[ \text{Var}[\varepsilon_i | x_i] = \sigma^2 \]
  with \( y_i \) distributed according to a Poisson distribution, conditional on \( x_i \) and \( v_i \); i.e.,
  \[ E[y_i | x_i, v_i] = \text{Var}[y_i | x_i, v_i] = \mu_i \]
  then
  \[ E[y_i | x_i] = E_{v_i}[\mu_i v_i] = \mu_i \]
  As in NB2
  and
  \[ \text{Var}[y_i | x_i] = E_{v_i} \left( \text{Var}[y_i | x_i, v_i] \right) + \text{Var}_{v_i} \left( E[y_i | x_i, v_i] \right) = \mu_i \left[ 1 + \sigma^2 \mu_i \right] \]
Mixture Models Properties

- In general, mixture models are characterized by
  - Overdispersion
  - A greater number of zeros relative to equidispersion model
  - Thicker tails than the parent distribution (in this case Poisson)
- If $v_j$ is drawn from a gamma distribution, then the resulting mixed Poisson model is in fact the Negative Binomial model
- Other mixing distributions are, of course, feasible

Truncation and Censored in Count Data Models

- It is not uncommon to find truncation and censoring problems in count data settings.
- Truncation frequently occurs with “zeros” being unobserved, resulting the “Positive Poisson” model
- Censoring can arise when “large” realizations are pooled into a single category (e.g., over 10).
- The fundamental steps required to correct of these data problems are similar to those seen for continuous variables, resulting in a corrected pdf or cdf
Truncation

- In general, suppose that $y_i$ had a discrete pdf of $h(y_i, \Lambda)$ where $\Lambda$ denotes the parameters of the distribution.

- Let

$$H(y_i, \Lambda) = \sum_{i=0}^{y_i} h(y_i, \Lambda), \quad y_i = 0, \ldots$$

denote the corresponding cdf

- For a sample truncated from below, excluding values of $y_i$ less than $r$, then

$$f(y_i, \Lambda | y_i \geq r) = \begin{cases} \frac{h(y_i, \Lambda)}{1 - H(r-1, \Lambda)} & y_i = r, r+1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Truncated Poisson

- Consider a situation in which only positive counts are observed; e.g.,
  - On site surveys (recreational, doctor visits, etc.)
  - Accident frequencies from police reports, etc.

- If the counts follow a Poisson process, then

$$H(0, \Lambda) = \Pr[y_i = 0] = \exp(-\mu_i)$$

and

$$f(y_i, \Lambda | y_i \geq 1) = \begin{cases} \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i! (1 - e^{-\mu_i})} & y_i = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$
Truncated Poisson (cont’d)

- The central moments of the truncated Poisson are given by

\[
E[y_i | y_i > 0] = \frac{\mu_i}{1-e^{-\mu_i}}
\]

\[
\text{Var}[y_i | y_i > 0] = \frac{\mu_i}{1-e^{-\mu_i}} \left[ 1 - \frac{\mu_i e^{-\mu_i}}{1-e^{-\mu_i}} \right] < \mu_i < E[y_i | y_i > 0]
\]

Truncated Poisson (cont’d)

For the Truncated Poisson regression model, the corresponding log-likelihood function is then given by

\[
LL(y,x; \beta) = \sum_{i=1}^{N} \left[ y_i \beta' x_i - \exp(\beta' x_i) - \ln(y_i!) - \ln \left( 1 - \exp[-\exp(\beta' x_i)] \right) \right]
\]

- Without the truncation correction, the parameter estimates will be biased
- Unlike its untruncated counterpart, the ML parameters estimates will be biased, even after correcting for truncation if overdispersion exists
- This latter results suggests using a model which allows for overdispersion (e.g., negative binomial)
Truncated Negative Binomial

• If we have a NB distribution truncated at zero, then

\[
h(y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1) \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}} \left( \frac{\mu_i}{\alpha^{-1} + \mu_i} \right)^{y_i}
\]

\[
H(0, \Lambda) = Pr[y_i = 0] = \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}}
\]

Which can in turn be used to derive truncated distribution and log-likelihood for estimation

Example:
Creel and Loomis (1990), *AJAE*

• Analyzed deer hunter trips in California in 1987
• Total number of trips were analyzed using
  – OLS
  – NLS
  – Truncated NLS (truncated at 0.5)
  – Poisson
  – Truncated Poisson
  – Negative Binomial
  – Truncated Negative Binomial
Example:
Creel and Loomis (1990) (cont’d)

- 2223 observations with a mean number of trips of 2.76
- Three subsets of the data were created
  - Specification (707)
  - Estimation (764)
  - Prediction (752)
- Explanatory Variables
  - Travel Cost (TC)
  - Travel Time (TIME)
  - Average Trip Length (DAYS)
  - Number of prior trips deer hunting (YEARS)
  - Success last year (BAG, dummy)
  - Number of prior passed up opportunities to bag a deer (PASNO)
  - Number of deer seen on past trip (DEERSN)
  - Household Income (INCOME)
  - Zonal hunting season length (SEASON)

Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS</th>
<th>NLS</th>
<th>TNLS</th>
<th>POIS</th>
<th>TPOIS</th>
<th>NB</th>
<th>TNB</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>-4.674</td>
<td>1.827</td>
<td>2.190</td>
<td>1.560</td>
<td>1.603</td>
<td>1.514</td>
<td>1.332</td>
</tr>
<tr>
<td>TC</td>
<td>-0.012</td>
<td>-0.006</td>
<td>-0.027</td>
<td>-0.006</td>
<td>-0.013</td>
<td>-0.006</td>
<td>-0.014</td>
</tr>
<tr>
<td>TIME</td>
<td>-0.052</td>
<td>-0.098</td>
<td>-0.175</td>
<td>-0.024</td>
<td>-0.033</td>
<td>-0.017</td>
<td>-0.017</td>
</tr>
<tr>
<td>DAYS</td>
<td>-0.154</td>
<td>-0.055</td>
<td>-0.057</td>
<td>-0.044</td>
<td>-0.048</td>
<td>-0.038</td>
<td>-0.050</td>
</tr>
<tr>
<td>YEARS</td>
<td>0.037</td>
<td>0.012</td>
<td>0.019</td>
<td>0.010</td>
<td>0.011</td>
<td>0.009</td>
<td>0.011</td>
</tr>
<tr>
<td>BAG</td>
<td>-0.252</td>
<td>-0.104</td>
<td>-0.189</td>
<td>-0.078</td>
<td>-0.104</td>
<td>-0.0747</td>
<td>-0.147</td>
</tr>
<tr>
<td>PASNO</td>
<td>0.226</td>
<td>0.027</td>
<td>0.033</td>
<td>0.034</td>
<td>0.030</td>
<td>0.040</td>
<td>0.062</td>
</tr>
<tr>
<td>DEERSN</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>INCOME</td>
<td>-0.019</td>
<td>-0.008</td>
<td>-0.014</td>
<td>-0.006</td>
<td>-0.008</td>
<td>-0.006</td>
<td>-0.011</td>
</tr>
<tr>
<td>SEASON</td>
<td>0.030</td>
<td>0.024</td>
<td>0.026</td>
<td>0.021</td>
<td>0.033</td>
<td>0.018</td>
<td>0.036</td>
</tr>
<tr>
<td>log-L</td>
<td>-1892.9</td>
<td>-1810.4</td>
<td>-1523.5</td>
<td>-1539.1</td>
<td>-1295.8</td>
<td>-1443.2</td>
<td>-1145.6</td>
</tr>
</tbody>
</table>
Out-of-Sample Predictions

<table>
<thead>
<tr>
<th>Statistical Model</th>
<th>OLS</th>
<th>NLS</th>
<th>TNLS</th>
<th>POIS</th>
<th>TPOIS</th>
<th>NB</th>
<th>TNB*</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.233</td>
<td>0.297</td>
<td>0.027</td>
<td>0.346</td>
<td>0.334</td>
<td>0.328</td>
<td>0.301</td>
</tr>
<tr>
<td>Act-Pred</td>
<td>-121.9</td>
<td>50.5</td>
<td>-775.0</td>
<td>16.0</td>
<td>-132.8</td>
<td>-99.1</td>
<td>-74.1</td>
</tr>
<tr>
<td>%Err</td>
<td>-6.6</td>
<td>2.7</td>
<td>-40.9</td>
<td>0.9</td>
<td>-7.2</td>
<td>-5.4</td>
<td>-4.0</td>
</tr>
<tr>
<td>CS/pred trips</td>
<td>117.25</td>
<td>172.82</td>
<td>36.72</td>
<td>153.62</td>
<td>74.71</td>
<td>163.05</td>
<td>70.07</td>
</tr>
</tbody>
</table>

Systems of Count Models

- Count data models are often applied to the estimation of consumer demand
- In these settings, a system of counts is need to allow for substitution possibilities among commodities, each good taking the form of counts
- Approaches
  - Independent Poisson
  - Seemingly Unrelated Poisson Regression Model (SUPREME)
  - Multivariate Poisson Log-normal
Sources


The SUPREME Specification

• The Supreme specification is fundamentally a mixed Poisson specification, where the mixing distribution is another independent Poisson

• Consider a two good example, with

\[
\begin{align*}
y_i^* &\sim Poisson[\mu_i] \quad i = 1, 2 \\
\omega &\sim Poisson[\xi] \\
\end{align*}
\]

assumed independent

Then

\[
y_i = y_i^* + \omega \sim Poisson[\mu_i + \xi] \quad i = 1, 2
\]

\[
E[y_i] = Var[y_i] = \mu_i + \xi = \theta_i
\]

\[
Cov(y_i, y_j) = \xi
\]
The SUPREME Specification
(cont’d)

The resulting bivariate Poisson distribution for \((y_1, y_2)\)
becomes

\[
P(y_1, y_2 | \theta_1, \theta_2, \xi) = \exp(\xi \theta_1 - \theta_2)
\]

\[
\times \sum_{j=0}^{\min(\xi y_2, y_2)} \frac{\xi^j}{j!} \left[ \frac{(\theta_1 - \xi)^{(y_1 - j)}}{(y_1 - j)!} \right] \left[ \frac{(\theta_2 - \xi)^{(y_2 - j)}}{(y_2 - j)!} \right]
\]

The SUPREME Specification
(cont’d)

• The “regression” version of this model comes about by
specifying that

\[
\theta_i = E[y_i | x_i, \beta_i] \quad i = 1, 2
\]

Typically, one uses a linear exponential form, with

\[
\theta_i = \exp(\beta_i^T x_i) \quad i = 1, 2
\]

where \(x_i (i=1,2)\) are \(k \times 1\) vectors of exogenous
explanatory variables and the \(\beta_i\)'s are unknown
parameter vectors.
Estimation the Supreme Model

- The SUPREME specification is estimated using maximum likelihood estimation, with the log-likelihood function given by:

\[
LL\left(y_1, y_2; x_1, x_2, \beta_1, \beta_2, \xi\right) = \sum_{i=1}^{N} \left[ \xi - \exp\left( \beta_1' x_{1i} \right) - \exp\left( \beta_2' x_{2i} \right) + \ln\left( \sum_{j=0}^{\min\left(y_{1i}, y_{2i}\right)} A_{ij} \right) \right]
\]

where

\[
A_{ij} = \frac{\xi^j}{j!} \left[ \frac{\left( e^{\beta_1 x_{1i}} - \xi \right)^{y_{1i} - j}}{(y_{1i} - j)!} \right] \left[ \frac{\left( e^{\beta_2 x_{1i}} - \xi \right)^{y_{2i} - j}}{(y_{2i} - j)!} \right]
\]

Merits of the SUPREME Model

- Advantages
  - Relatively easy to implement
  - Efficiency gain over independent Poisson, since it allows for covariance (even when explanatory variables are the same in the two equations)
  - With simultaneous estimation, one can test cross equation restrictions

- Disadvantages
  - Cannot accommodate overdispersion
  - Cannot capture negative correlations between equations
  - Typically restricted to bivariate setting, though can be theoretically extended to more than two
Example: Ozuna and Gomez (1990)

- Examined recreation trips to two Texas Lakes
  - Conroe
  - Somerville
- 659 observations from a survey of boat owners
- Explanatory Variables
  - Income
  - Travel costs to each site (and two substitute sites)
  - Quality ratings for four sites
  - Dummy variables for
    - water skiing
    - annual pass to Lake Somerville
    - overnight stay

Results

<table>
<thead>
<tr>
<th></th>
<th>SUR</th>
<th>SUPREME</th>
<th>Independent Poisson</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conroe</td>
<td>Somerville</td>
<td>Conroe</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.492</td>
<td>0.725</td>
<td>0.257</td>
</tr>
<tr>
<td>C_conroe</td>
<td>-0.111</td>
<td>0.033</td>
<td>-0.103</td>
</tr>
<tr>
<td>C_livingston</td>
<td>0.010</td>
<td>-0.003</td>
<td>0.064</td>
</tr>
<tr>
<td>C_somerville</td>
<td>0.033</td>
<td>-0.104</td>
<td>0.028</td>
</tr>
<tr>
<td>C_chuston</td>
<td>-0.021</td>
<td>0.070</td>
<td>0.005</td>
</tr>
<tr>
<td>Income</td>
<td>-0.065</td>
<td>-0.120</td>
<td>-0.044</td>
</tr>
<tr>
<td>Q_conroe</td>
<td>1.390</td>
<td>0.004</td>
<td>0.013</td>
</tr>
<tr>
<td>Q_somerville</td>
<td>0.833</td>
<td>0.507</td>
<td>0.043</td>
</tr>
<tr>
<td>Waterski</td>
<td>1.019</td>
<td>0.507</td>
<td>0.486</td>
</tr>
<tr>
<td>Q_livingston</td>
<td>-0.151</td>
<td>-0.065</td>
<td>-0.064</td>
</tr>
<tr>
<td>Overnight</td>
<td>-1.453</td>
<td>-0.571</td>
<td>-0.064</td>
</tr>
<tr>
<td>Fee</td>
<td>5.904</td>
<td>0.647</td>
<td>-0.647</td>
</tr>
<tr>
<td>( \xi )</td>
<td></td>
<td></td>
<td>0.023</td>
</tr>
</tbody>
</table>
Mixed Multivariate Poisson Models

- The concept underlying the SUPREME specification can be extended to a multivariate setting much like Mixed Logit extends the MNL model.

- Suppose that
  \[y_{ij} \mid x_{ij}, \beta_j \sim \text{Poisson}[\mu_{ij}] \quad j = 1, \ldots, J\]

  where
  \[
  \mu_{ij} = \exp(\beta'_j x_{ij})
  \]
  \[\beta_j \sim f(\beta \mid \theta)\]

  then
  \[
  \Pr[y_{ij} = k \mid x_{ij}, \beta_j] = \frac{e^{-\mu_{ij}} \mu_{ij}^k}{k!}
  \]

Mixed Multivariate Poisson Models (cont’d)

- The unconditional choice probabilities are then obtained by integrating over the chosen parametric distributions, with
  \[
  \Pr[y_{ir} = k \mid x_{ir}] = \left[\prod_{j=1}^J \frac{e^{-\mu_{ij}} \mu_{ij}^k}{k!}\right] f(\beta \mid \theta) d\beta
  \]

- Simulation methods can be used in ML estimation
- The distributional assumptions associated with the parameters can be used to induce
  - correlation between equations
  - overdispersion
Comparing Continuous and Count Data Demand Systems

- von Haefen and Phaneuf (2003) provide a comparison of demand systems based on
  - Kuhn-Tucker framework
  - Count data models
- They focus on both
  - conceptual differences in the modeling approaches and
  - the empirical performance of each model in analyzing wetland recreation usage in Iowa

Review of Kuhn-Tucker Model

- Recall that the KT model starts with a direct utility function with an integrated error term capturing unobserved factors causing preference heterogeneity; i.e.,

\[ U(x, q, z, \varepsilon) \]

where
- \( x \) denotes the \( M \times 1 \) vector of commodities,
- \( q \) denotes an \( M \times K \) vector of commodity characteristics,
- \( z \) is the numeraire good and
- \( \varepsilon \) denotes the vector capturing unobserved preference factors
Review of Kuhn-Tucker Model (cont’d)

• The individual is assumed to solve

\[
\max_x U(x, q, y - p'x, \varepsilon) \quad \text{s.t. } x \geq 0
\]

yielding first order conditions

\[
p_i \geq \frac{U(x^*, q, y - p'x^*, \varepsilon)}{U_x(x, q, y - p'x^*, \varepsilon)}
\]

\[\equiv \xi_{xi}(p, q, y, \varepsilon) \quad \forall i = 1, \ldots, M\]

where \(\xi_{xi}\) represents a “virtual price” for commodity \(i\)

• Thus,

\[
x_i \begin{cases}
> 0 & \text{if } p_i = \xi_{xi}(p, q, y, \varepsilon) \\
= 0 & \text{if } p_i > \xi_{xi}(p, q, y, \varepsilon)
\end{cases}
\]

• The individual’s derived demands can be written as

\[
x^* = f[\xi_{xi}(p, q, y, \varepsilon), q, y, \varepsilon]
\]

• The KT conditions are used to derive log-likelihood functions used in estimation
Count Data Demand Systems

- Ideally a count data model proceeds from the maximization problem

\[
\max_x U \left( x, q, y - p'x, \varepsilon \right) \quad \text{s.t. } x_i \in \{0,1,\ldots\} \forall i
\]

- However, the count nature of the choice set precludes the use of differential calculus

- The solution relies on comparison across a large choice set

- Count data models instead rely on utility maximization in the specification of the expected number of trips and the count distributional assumptions to generate trips

Count Data Demand Systems (cont’d)

- Formally, expected trips are specified, with

\[
E[x] = f(p, q, y)
\]

- Utility theory is employed at this stage by requiring that these expected trip demand equations satisfy standard integrability restrictions; i.e.,
  - adding up
  - homogeneity of degree zero in prices and income
  - symmetry and negative semi-definite Slutsky matrix

- Note: These assumptions are not imposed on the actual trip demands, just their expectations
Count Data Demand Systems
(cont’d)

- Welfare analysis proceeds by using the expected demands to integrate back to an indirect utility function
  \[ v(p,q,y) \]

- It is not clear whose preferences this indirect utility represents
  - It is derived from average or expected trip demands, not any one individual’s demands
  - All prices are likely to enter these average demands, so they will enter the counterpart representative indirect utility function,
  - This representative utility function is in turn used for welfare analysis for all agents so all price changes matter for all agents
  - This contrast with the KT approach, for which only consumed commodity price enter an individual indirect utility function

Model Specification - KT

- The authors use a Stone-Geary utility function
  \[ \sum_{i=1}^{M} \Psi_i \ln (\phi_i x_i + \theta_i) + \ln (y - p'x) \]

where
- \( \Psi_i \) is a quality index for good \( i \)
- \( \phi_i \) is a repackaging parameter for good \( i \)
- \( \theta_i \) is a translation parameter for good \( i \)
Model Specification – KT (cont’d)

- Two specifications are considered

Specification #1

\[ \Psi_i = \exp(\delta' s + \gamma' q_i + \varepsilon_i) \]
\[ \phi_i = 1 \]

where \( s \) denotes individual socio-demographic characteristics and \( q_i \) denotes attributes of good \( i \)

Specification #2

\[ \Psi_i = \exp(\delta' s + \varepsilon_i) \]
\[ \phi_i = \exp(\gamma' q_i) \]

authors refer to this as weak complementary model

Model Specification – KT (cont’d)

- The resulting KT conditions are given by

\[ \varepsilon_i \leq \ln p_i + \ln \left( x_i + \frac{\theta_i}{\phi_i} \right) - \ln (y - p'x) - \ln \Psi_i = g_i \]

where

\[ \Psi_i = \exp(\ln \Psi_i - \varepsilon_i) \]

Assuming iid extreme value errors, the likelihood function becomes

\[ L = \text{abs} \left| \prod_{i=1}^{M} 1_{x_i>0} \frac{\exp(-g_i / \nu_i)}{\nu_i} \right| \exp \left( - \sum_{i=1}^{M} \exp(-g_i / \nu_i) \right) \]
Model Specification – Counts

• The Count Data model starts with a specification of expected counts for each commodity. Shonkwiler (1999), started with

\[
E[x_i] = \exp \left( \alpha_i + \sum_{j=1}^{M} \beta_{ij} p_j + \lambda_i \ln y \right) \quad \forall i = 1, \ldots, M
\]

Integrability restrictions can be imposed on this system of incomplete demand equations by assuming that:
\[
\alpha_i > 0 \quad \lambda_i = \lambda_i \quad \forall i
\]
\[
\beta_{ii} < 0 \quad \beta_{ij} = 0 \quad \forall j \neq i
\]

Model Specification – Counts
(cont’d)

• The resulting system of expected demands becomes

\[
E[x_i] = \exp \left( \alpha_i + \beta_{ii} p_i + \lambda \ln y \right) \quad \forall i = 1, \ldots, M
\]

• Notice that we now have
  – no cross price effects
  – a common income effect
• vonHaefen and Phaneuf use a slight modification of this functional form

\[
E[x_i] = \frac{1}{\phi_i} \exp \left( \alpha_i + \beta_{ii} \frac{p_i - \alpha_i}{\phi_i} + \lambda \ln y \right) \quad \forall i = 1, \ldots, M
\]
Model Specification – Counts (cont’d)

• Two specifications are considered
  Specification #1: “Simple repackaging”
  \[ \omega_i = 0 \]
  \[ \phi_i = \gamma' q_i \]

  where \( \omega_i \) is a “cross-product” parameter and \( \phi_i \) is the “repackaging” parameter

  Specification #2
  \[ \omega_i = \gamma' q_i \]
  \[ \phi_i = 1 \]

Accounting for Excess Zeros

• To complete the model specification, the authors note that the data are characterized by excess zeros.

• One way to handle this problem is to use a “zero-inflated” count model, which assumes that

  \[ \Pr[x_i = 0] = \pi_i + (1 - \pi_i) e^{-\mu_i} \]
  \[ \Pr[x_i = r] = (1 - \pi_i) \frac{e^{-\mu_i} \mu_i^r}{r!} \]

  This is a finite mixture model in which one of the distributions is degenerate at zero and the other is Poisson
Accounting for Excess Zeros (cont’d)

• Notice that for such models

\[ \text{Var}[x_i] = (1 - \pi_i)(\mu_i + \pi_i\mu_i^2) > (1 - \pi_i)\mu_i = E[x_i] \]

thus excess zeros imply overdispersion

• One can then specify a functional form for spike at zero; e.g., logistic with

\[ \pi_i = \frac{\exp(\tau'z)}{1 + \exp(\tau'z)} \]

Error Specification – Count Model

• von Haefen and Phaneuf take this model one step further, assuming that the counts are independent zero inflated negative binomial counts with

\[ \mu_i = \frac{1}{\phi_i} \exp\left(\alpha_i + \beta_i \frac{P_i - \omega_i}{\phi_i} + \lambda \ln y\right) \quad \forall i = 1, \ldots, M \]

\[ \pi_i = \frac{\exp(\rho_i + \delta's)}{1 + \exp(\rho_i + \delta's)} \]

Note: This allows the spike probability, \( \pi_i \), to vary by commodity, but requires the marginal impact of socio-demographic characteristics to be the same
Count Model Log-Likelihood

The resulting log-likelihood function is given by

\[
L = \prod_{i=1}^{M} \left\{ \pi_i \left(1 - \pi_i \right) \left( \frac{1}{1 + v_i \eta_i} \right)^{\frac{1}{v_i}} \right\} \\
+ \prod_{x_i > 0} \left\{ \frac{\Gamma \left( \frac{1}{v_i} + x_i \right)}{\Gamma \left( \frac{1}{v_i} \right)} \left( \frac{1}{1 + v_i \eta_i} \right)^{\frac{1}{v_i}} \left( \frac{v_i \eta_i}{1 + v_i \eta_i} \right)^{x_i} \right\}
\]

Application – Iowa Wetlands

- Analysis based on 1997 recreational trips to wetlands in Iowa
- 2891 observations
- vonHaefen and Phaneuf aggregated destinations to 5 “mega-zones”
- explanatory variables include
  - travel costs
  - pheasant counts in each zone
  - socio-demographic characteristics
    - gender
    - hunting license
    - age
    - college education dummy variable
### Descriptive Statistics
#### Average Trip Data

<table>
<thead>
<tr>
<th>Site</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trips</td>
<td>0.81</td>
<td>1.18</td>
<td>3.60</td>
<td>1.44</td>
<td>1.47</td>
<td>8.51</td>
</tr>
</tbody>
</table>
| Range| [0,40]| [0,40]| [0,49]| [0,48]| [0,45]| [0,49]|}
| %0’s | 88    | 84    | 60    | 79    | 78    | 31?   |
| %≥40 | 0.03  | 0.07  | 0.45  | 0.28  | 0.21  | 2.56  |
| Cost | 156   | 119   | 76    | 118   | 106   | 115   |
| Phsnt.| 18    | 51    | 56    | 28    | 40    | 38    |

### Other Summary Statistics

<table>
<thead>
<tr>
<th>%Visiting</th>
<th>Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Sites</td>
<td>Mean Income</td>
</tr>
<tr>
<td></td>
<td>%Male</td>
</tr>
<tr>
<td>1 Site</td>
<td>Mean Age</td>
</tr>
<tr>
<td>2 Sites</td>
<td>% License</td>
</tr>
<tr>
<td>3 Sites</td>
<td>% 4yr college</td>
</tr>
<tr>
<td>4 Sites</td>
<td></td>
</tr>
<tr>
<td>5 Sites</td>
<td></td>
</tr>
</tbody>
</table>
### KT Results

<table>
<thead>
<tr>
<th></th>
<th>Specification 1</th>
<th>Specification 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Likelihood</td>
<td>-14,919</td>
<td>-14,940</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>5.522</td>
<td>6.963</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>6.915</td>
<td></td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>6.656</td>
<td></td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>5.475</td>
<td></td>
</tr>
<tr>
<td>$\theta_5$</td>
<td>5.843</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\text{constant}}$</td>
<td>-5.508</td>
<td>-5.065</td>
</tr>
<tr>
<td>$\delta_{\text{permit}}$</td>
<td>0.344</td>
<td>0.341</td>
</tr>
<tr>
<td>$\delta_{\text{male}}$</td>
<td>-0.035</td>
<td>-0.034</td>
</tr>
<tr>
<td>$\delta_{\text{age}}$</td>
<td>-0.005</td>
<td>-0.004</td>
</tr>
<tr>
<td>$\delta_{\text{college}}$</td>
<td>-0.095</td>
<td>-0.092</td>
</tr>
<tr>
<td>$\gamma_{\text{pheasants}}$</td>
<td>0.011</td>
<td>0.003</td>
</tr>
</tbody>
</table>

### Count Data Model Results

**Spike Parameters**

<table>
<thead>
<tr>
<th></th>
<th>Specification 1</th>
<th>Specification 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Likelihood</td>
<td>-15,370</td>
<td>-15,400</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-2.34</td>
<td>-1.96</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.97</td>
<td>-0.44</td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>-1.74</td>
<td>-1.67</td>
</tr>
<tr>
<td>$\rho_4$</td>
<td>-1.63</td>
<td>-1.60</td>
</tr>
<tr>
<td>$\rho_5$</td>
<td>-1.12</td>
<td>-1.42</td>
</tr>
<tr>
<td>$\delta_{\text{permit}}$</td>
<td>-2.34</td>
<td>-2.18</td>
</tr>
<tr>
<td>$\delta_{\text{male}}$</td>
<td>-0.63</td>
<td>-0.60</td>
</tr>
<tr>
<td>$\delta_{\text{age}}$</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>$\delta_{\text{college}}$</td>
<td>-1.04</td>
<td>-1.07</td>
</tr>
</tbody>
</table>
### Count Data Model Results

#### Count Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Specification 1</th>
<th>Specification 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>-3.19</td>
<td>-5.55</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-1.26</td>
<td></td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>-1.80</td>
<td></td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>-0.39</td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>-0.22</td>
<td>-0.015</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>-1.18</td>
<td>-0.026</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>-2.14</td>
<td>-0.038</td>
</tr>
<tr>
<td>$\beta_{44}$</td>
<td>-0.65</td>
<td>-0.023</td>
</tr>
<tr>
<td>$\beta_{55}$</td>
<td>-1.48</td>
<td>-0.032</td>
</tr>
<tr>
<td>$\gamma_{\text{pheasants}}$</td>
<td>1 (not estimated)</td>
<td>0.808</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.71</td>
<td>0.699</td>
</tr>
</tbody>
</table>

### Elasticity Implications

<table>
<thead>
<tr>
<th>Own Price</th>
<th>KT 1</th>
<th>KT 2</th>
<th>Count 1</th>
<th>Count 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Price</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.60</td>
<td>-3.00</td>
<td>-1.91</td>
<td>-2.31</td>
</tr>
<tr>
<td>2</td>
<td>-2.69</td>
<td>-2.94</td>
<td>-2.79</td>
<td>-3.08</td>
</tr>
<tr>
<td>3</td>
<td>-2.11</td>
<td>-1.94</td>
<td>-2.90</td>
<td>-2.89</td>
</tr>
<tr>
<td>4</td>
<td>-2.66</td>
<td>-2.66</td>
<td>-2.78</td>
<td>-2.70</td>
</tr>
<tr>
<td>5</td>
<td>-2.36</td>
<td>-2.57</td>
<td>-3.86</td>
<td>-3.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
<th>KT 1</th>
<th>KT 2</th>
<th>Count 1</th>
<th>Count 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.61</td>
<td>2.92</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>2</td>
<td>2.85</td>
<td>2.75</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>3</td>
<td>2.01</td>
<td>1.89</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>4</td>
<td>2.75</td>
<td>2.70</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>5</td>
<td>2.39</td>
<td>2.53</td>
<td>0.71</td>
<td>0.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Own Quality</th>
<th>KT 1</th>
<th>KT 2</th>
<th>Count 1</th>
<th>Count 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.49</td>
<td>0.12</td>
<td>0.91</td>
<td>0.21</td>
</tr>
<tr>
<td>2</td>
<td>1.63</td>
<td>0.31</td>
<td>1.79</td>
<td>1.06</td>
</tr>
<tr>
<td>3</td>
<td>1.14</td>
<td>0.15</td>
<td>1.90</td>
<td>1.72</td>
</tr>
<tr>
<td>4</td>
<td>0.79</td>
<td>0.16</td>
<td>1.78</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>1.04</td>
<td>0.23</td>
<td>2.86</td>
<td>1.04</td>
</tr>
</tbody>
</table>
### Welfare Estimates ($’s)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>KT Models</th>
<th>Count Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spec. 1</td>
<td>Spec. 2</td>
</tr>
<tr>
<td>$50 fee at Sites 1 and 2</td>
<td>-76.67 (2.89)</td>
<td>-84.28 (3.11)</td>
</tr>
<tr>
<td>20% increase in Pheasants at Site 2</td>
<td>16.47 (2.38)</td>
<td>1.98 (0.53)</td>
</tr>
</tbody>
</table>