Correlation in The Bivariate Poisson Regression Model

William Greene[*]
Department of Economics, Stern School of Business,
New York University,
May, 2007

Abstract

We consider a bivariate Poisson model that is based on the lognormal heterogeneity model. Two recent applications have used this model. We suggest that the correlation estimated in their model frameworks is an ambiguous measure of the correlation of the variables of interest, and may substantially overstate it. We conclude with a detailed application of the proposed method using the data employed in one of the two aforementioned bivariate Poisson studies.

JEL classification: C14; C23; C25

Keywords: Poisson regression; Heterogeneity; Lognormal; Bivariate Poisson; Simulation

[*] 44 West 4th St., Rm. 7-78, New York, NY 10012, USA, Telephone: 001-212-998-0876; e-mail: wgreene@stern.nyu.edu, URL www.stern.nyu.edu/~wgreene.

1 Introduction

Models for count data have been prominent in many branches of the recent applied literature, for example, in health economics (e.g., in numbers of visits to health facilities[1]) management (e.g., numbers of patents[2]) and industrial organization (e.g., numbers of entrants to markets[3]). The foundational building block in this modeling framework is the Poisson regression model.[4] Two recent applications, Munkin and Trivedi (1999) and Riphahn et al. RWM (2003), have used a form of the bivariate Poisson/lognormal model in which the correlation is introduced through additive correlated latent variables in the conditional mean functions. This approach overstates the correlation coefficient estimated in their model frameworks. What they have specified is correlation between the logs of the conditional mean functions. This paper derives the implied correlation between the observed random variables and applies the results to the data set used in one of the studies mentioned.

The study is organized as follows: Section 2 will detail the basic modeling frameworks for count data, the Poisson and negative binomial (NB) models and will propose models for observed and unobserved heterogeneity in count data. We propose the lognormal model as an alternative to the log gamma model that produces the NB specification. Section 3 will detail the specification of the bivariate Poisson model. The results derived for the model are applied to the RWM panel data on health care utilization in Section 4. Some conclusions are drawn in Section 5.

2 Basic Functional Forms for Count Data Models

The literature abounds with alternative models for counts – see, e.g., Cameron and Trivedi (CT) (1998) and Winkelmann (2003). However, the Poisson and a few forms of the negative binomial model overwhelmingly dominate the received applications. [See, as well, Hilbe (2007).]

2.1 The Poisson Regression Model

The canonical regression specification for a variable *Y* that is a count of events is the Poisson regression,

$$(2\text{-}1) \qquad \text{Prob}[Y = y_i \mid \mathbf{x}_i] = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{\Gamma(1 + y_i)}, \lambda_i = \exp(\alpha + \mathbf{x}_i'\boldsymbol{\beta}), y_i = 0,1,...,i = 1,...,N \,,$$

where $\mathbf{x}_i$ is a vector of covariates and, $i = 1,…,N$, indexes the *N* observations in a random sample. For reasons that will emerge below, we explicitly assume that there is a constant term in the model.]The regression model is developed in detail in a vast number of standard references such as CT (1986, 1998, 2005), Winkelmann (2003) and Greene (2008a).] The Poisson model has the convenient feature that

$$(2\text{-}2) \qquad E[y_i|\mathbf{x}_i] = \lambda_i.$$

It has the undesirable characteristic that

---

[1] Contoyannis, Jones and Rice (2004), Munkin and Trivedi (1999), Riphahn, Wambach and Million (RWM) (2003). See, as well, Cameron and Trivedi (2005).
[2] Hausman, Hall and Griliches (HHG) (1984) and Wang, Cockburn and Puterman (1998).
[3] Asplund and Sandin (1999).
[4] HHG (1984), Cameron and Trivedi (1986, 1998), and Winkelmann (2003).

(2-3) $\text{Var}[y_i|\mathbf{x}_i] = \lambda_i.$

This is the 'equidispersion' aspect of the model. Since observed data will almost always display pronounced *over*dispersion, analysts typically seek alternatives to the Poisson model, such as the negative binomial model described below.

Estimates of the parameters of the model using a sample of N observations on $(y_i, \mathbf{x}_i)$, $i = 1,\ldots,N$, are obtained by maximizing the log likelihood function,

(2-4) $\ln L = \sum_{i=1}^{N} \left[y_i(\alpha + \mathbf{x}_i'\boldsymbol{\beta}) - \lambda_i - \ln\Gamma(1 + y_i) \right].$

2.2 The Negative Binomial and Poisson Lognormal Regression Models

The Poisson model imposes the (usually) transparently restrictive assumption that the conditional variance equals the conditional mean. The typical alternative is the negative binomial (NB) model. The model can be motivated as an attractive functional form simply in its own right that allows overdispersion. However, it is useful for present purposes to obtain the specification through the introduction of unobserved heterogeneity in the Poisson regression model. We consider two possible cases, the conventional approach based on the log gamma distribution and, we will argue, a more flexible approach based on the lognormal distribution.

2.2.1 The Negative Binomial Model

To introduce latent heterogeneity into the count data model, we write

(2-5) $E[y_i|\mathbf{x}_i, \varepsilon_i] = \exp(\alpha + \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i) = h_i\lambda_i,$

where $h_i = \exp(\varepsilon_i)$ is assumed to have a one parameter gamma distribution, $G(\theta,\theta)$ with mean 1 and variance $1/\theta = \kappa$. That is

(2-6) $f(h_i) = \dfrac{\theta^\theta \exp(-\theta h_i)h_i^{\theta-1}}{\Gamma(\theta)}, h_i \geq 0, \theta > 0.$ [5]

The nonzero mean of ε_i will be absorbed in the constant term of the index function. By integrating h_i out of the conditional Poisson mixture mdoel, we obtain the (NB) model,

(2-7) $\text{Prob}[Y = y_i|\mathbf{x}_i] = \dfrac{\Gamma(\theta + y_i)r_i^\theta(1 - r_i)^{y_i}}{\Gamma(1 + y_i)\Gamma(\theta)}, y_i = 0,1,\ldots, \theta > 0, r_i = \theta/(\theta + \lambda_i).$

[A full derivation of this result may be found in numerous references including Greene (2007b).] The conditional mean and variance of the NB random variable relate to the Poisson moments as follows:

(2-8) $E[y_i|\mathbf{x}_i] \qquad = \lambda_i,$

(2-9) $\text{Var}[y_i|\mathbf{x}_i] \qquad = \lambda_i[1 + (1/\theta)\lambda_i]$

---

[5] This general approach is discussed at length by Gourieroux, Monfort and Trognon (1984), CT (1986, 1998), Winkelmann (2003) and HHG (1984).

$$= \lambda_i \, [1 + \kappa \, \lambda_i]$$

where $\quad \kappa \quad = \mathrm{Var}[h_i]$.

Maximum likelihood estimation and inference about the parameters of the NB model $(\alpha, \boldsymbol{\beta}, \theta)$ is straightforward, as documented in, e.g., Greene (2007a), CT (1990, 1998, 2005) and Hilbe (2007).

2.2.2 Poisson Lognormal Mixture Model

Consider, instead, introducing the heterogeneity in (2-5) as a normally distributed variable with mean zero and standard deviation σ, which we introduce into the model explicitly by standardizing ε_i. Then, the conditional Poisson model is

$$(2\text{-}10) \qquad P(y_i|\mathbf{x}_i,\varepsilon_i) = \frac{\exp(-h_i\lambda_i)(h_i\lambda_i)^{y_i}}{\Gamma(1+y_i)}, \; h_i\lambda_i = \exp(\alpha + \mathbf{x}_i'\boldsymbol{\beta} + \sigma\varepsilon_i) , \; \varepsilon_i \sim \mathrm{N}[0,1].$$

The unconditional density would be

$$(2\text{-}11) \qquad P(y_i|\mathbf{x}_i) = \int_{-\infty}^{\infty} \frac{\exp[-\exp(\sigma\varepsilon_i)\lambda_i][\exp(\sigma\varepsilon_i)\lambda_i]^{y_i}}{\Gamma(1+y_i)}\phi(\varepsilon_i)\mathrm{d}\varepsilon_i ,$$

where here and in what follows, $\phi(\varepsilon_i)$ denotes the standard normal density. The unconditional log likelihood function is

$$(2\text{-}12) \qquad \ln L \quad = \sum_{i=1}^{N} \ln P(y_i \mid \mathbf{x}_i)$$

$$= \sum_{i=1}^{N} \ln\left\{ \int_{-\infty}^{\infty} \frac{\exp[-\exp(\sigma\varepsilon_i)\lambda_i][\exp(\sigma\varepsilon_i)\lambda_i]^{y_i}}{\Gamma(1+y_i)}\phi(\varepsilon_i)\mathrm{d}\varepsilon_i \right\} .$$

Maximum likelihood estimates of the model parameters are obtained by maximizing the unconditional log likelihood function with respect to the model parameters $(\alpha, \boldsymbol{\beta}, \sigma)$.

The integrals in the log likelihood function do not exist in closed form. The quadrature based approach suggested by Butler and Moffitt(1982) is a convenient method of approximating them. [See, e.g., Greene (2007a).] Simulation is another effective approach to maximizing the log likelihood function. [See Train (2003) and Greene (2008).] The log likelihood function is

$$(2\text{-}13) \qquad \ln L = \sum_{i=1}^{N} \ln \int_{-\infty}^{\infty} P(y_i|\mathbf{x}_i,\varepsilon_i)] \; \phi(\varepsilon_i)d\varepsilon_i .$$

The simulated log likelihood would be

$$(2\text{-}14) \qquad \ln L_S = \sum_{i=1}^{N} \ln \frac{1}{M}\sum_{m=1}^{M} P(y_i|\mathbf{x}_i,\sigma\varepsilon_{im})]$$

where ε_{im} is a set of M random draws from the standard normal population. [We would propose to improve this part of the estimation by using Halton sequences, instead. See Train (2003, pp. 224-238) and Greene (2008a).] Extensive discussion of maximum simulated likelihood estimation

appears in Gourieroux and Monfort (1996), Munkin and Trivedi (1999), Train (2003) and Greene (2008a).][6]

The mean and variance of the lognormal variable are

$$(2\text{-}15) \qquad E[\exp(\sigma\varepsilon_i)] \quad = \exp(\sigma^2/2),$$

$$\mathrm{Var}[\exp(\sigma\varepsilon_i)] \quad = E[\exp(\sigma\varepsilon_i)^2] - \{E[\exp(\sigma\varepsilon_i)]\}^2$$

$$= \exp(\sigma^2)[\exp(\sigma^2) - 1].$$

The conditional mean in the Poisson lognormal model is

$$(2\text{-}16) \qquad E[y_i|\, \mathbf{x}_i,\varepsilon_i] \; = \; \lambda_i \exp(\sigma\varepsilon_i).$$

It follows that

$$E[y_i|\mathbf{x}_i] \; = \; E_\varepsilon[E[y_i|\mathbf{x}_i,\varepsilon_i]]$$

$$(2\text{-}17) \qquad\qquad = \; \lambda_i \exp(\sigma^2/2)$$

$$= \; \exp[(\alpha + \sigma^2/2) + \mathbf{x}_i'\boldsymbol{\beta}].$$

To obtain the unconditional variance, we use

$$(2\text{-}18) \qquad \mathrm{Var}[y_i|\mathbf{x}_i] \; = \; E_{\varepsilon_i}\left[\mathrm{Var}[y_i|\, \mathbf{x}_i,\varepsilon_i] + Var_{\varepsilon_i}[E[y_i|\, \mathbf{x}_i,\varepsilon_i]]\right].$$

Combining the results above, we find

$$(2\text{-}19) \qquad \mathrm{Var}[y_i|\mathbf{x}_i] \quad = \; \lambda_i\exp(\sigma^2/2)\{1 + \lambda_i\exp(\sigma^2/2)[\exp(\sigma^2)-1]\}$$

$$= \; E[y_i|\, \mathbf{x}_i,\varepsilon_i]\{1 + \tau\, E[y_i|\, \mathbf{x}_i,\varepsilon_i]\},\ \tau = [\exp(\sigma^2)-1].$$

Thus, the variance in the lognormal model has the same quadratic form as that in the negative binomial model in (2-9).

One could argue that the lognormal model is a more natural specification. If the heterogeneity captures the aggregate of individually small influences, then an appeal to the central limit theorem would motivate the normal distribution more than the log gamma. [See Winkelmann (2003).] The attraction in this development is the ease with which the normal mixture model can be extended and adapted to new models and formulations, such as the two part models below. The log gamma model that underlies the familiar negative binomial specification provides no means doing so. [See, as well, RWM (1003, p. 395) and Million (1998).]

2.3. Models for Panel Data

A random effects Poisson model can be formed by writing

$$(2\text{-}20) \qquad \lambda_{it} = \exp(\mathbf{x}_{it}'\boldsymbol{\beta} + u_i)$$

---

[6] One could preserve the log gamma specification by drawing h_{im} from a gamma(1,1) population and using the logs in the simulation, rather than using draws from N[0,1] for w_{im}. This approach, which obviates deriving the unconditional distribution analytically, was used in Munkin and Trivedi (1999).

where u_i is independent of $\mathbf{x}_{it}$. The Poisson RE model could also be specified with lognormal heterogeneity. Analysis would follow precisely along the lines of Section 2.2.2. The joint probability would be computed from

$$P(y_{i1},...,y_{iT} \mid \mathbf{X}_i) = \int_{u_i} \prod_{t=1}^{T} \frac{\exp(-\exp(u_i)\lambda_{it})(\exp(u_i)\lambda_{it})^{y_{it}}}{\Gamma(1+y_{it})} f(u_i)du_i$$

(2-21)

$$= \prod_{t=1}^{T} \left[\frac{\lambda_{it}^{y_{it}}}{\Gamma(1+y_{it})} \right] \int_{u_i} \exp\left[-\exp(u_i)\Sigma_{t=1}^{T}\lambda_{it} \right] \left[\exp(u_i) \right]^{\Sigma_{t=1}^{T} y_{it}} f(u_i)du_i.$$

This function and its derivatives can be approximated using either quadrature or simulation.

3. The Bivariate Poisson Model

There have been a variety of proposals for a bivariate (or multivariate) count data model. The earliest form is that of Kocherlakota and Kocherlakota (1992) which is based on the trivariate reduction method. Let z_1, z_2 and u denote three Poisson distributed random variables. Then, the observed random variables,

$$y_1 = z_1 + u$$

(3-1)   and

$$y_2 = z_2 + u$$

have a bivariate Poisson distribution with correlation $\lambda_u / \left[\sqrt{(\lambda_1 + \lambda_u)(\lambda_2 + \lambda_u)} \right]$. This model does produce a pair of correlated Poisson variables, however the correlation must be positive, which severely limits the generality of this specification. (For the outcomes examined in Section 4, doctor visits and hospital visits, a negative correlation would not be surprising.) [7] Munkin and Trivedi (1999) present a survey of other approaches.

Two recently developed approaches considered here [Munkin and Trivedi (1999) and RWM (2003)] build the bivariate model into latent heterogeneity structures, as employed in the various models proposed above. These allow the sign of the correlation to vary. However, they shift the impact of the bivariate distribution from the variables of interest, as in the trivariate model above, to the unobservables in the conditional mean function. The bivariate count outcomes model is still preserved. However, the estimated correlations in these models do not provide a clear picture of the implied correlations between the outcome variables that was the objective to begin with. As a general proposition, the correlation between the observed counts will be less, potentially far less, than the estimated correlation between the underlying unobserved heterogeneity.

The bivariate probit model specified in Munkin and Trivedi (1999) and Riphahn, Rambach and Million (2003) is

$$\exp(\mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \sigma_1\varepsilon_{1i}) = \lambda_{1i}\exp(\sigma_1\varepsilon_{1i})$$

$$\exp(\mathbf{x}_{2i}'\boldsymbol{\beta}_2 + \sigma_2\varepsilon_{2i}) = \lambda_{2i}\exp(\sigma_2\varepsilon_{2i})$$

(3-2) $(\varepsilon_{1i},\varepsilon_{2i}) \sim N_2[(0,0),(1,1),\rho]$

---

[7] The trivariate reduction method was employed e.g., by Jung and Winkelmann (1993), Karlis and Ntzoufras (2003) and King (1989).

$$P(y_{ji}|\mathbf{x}_{ji},\varepsilon_{ji}) = \frac{\exp(-\exp(\sigma\varepsilon_{ji})\lambda_{ji})(\exp(\sigma\varepsilon_{ji})\lambda_{ji})^{y_{ji}}}{\Gamma(1+y_{ji})}, j=1,2$$

Both studies build the empirical measurement of correlation of the two outcomes around the estimation of ρ. However, as we now demonstrate, the correlation coefficient, ρ, provides a misleading description of this correlation. Superficially, this is obvious from the construction. The coefficient ρ is not the correlation between y_{1i} and y_{2i}; it is the correlation between $\ln E[y_{1i}|\mathbf{x}_{1i},\varepsilon_{1i}]$ and $\ln E[y_{2i}|\mathbf{x}_{2i},\varepsilon_{2i}]$. How this relates to $Corr[y_{1i},y_{2i}|\mathbf{x}_{1i},\mathbf{x}_{2i}]$ is less than clear. To deduce this from the model specification, we proceed as follows:

(3-3)
$$Corr[y_{1i},y_{2i}|\mathbf{x}_{1i},\mathbf{x}_{2i}] = \frac{Cov[y_{1i},y_{2i}|\mathbf{x}_{1i},\mathbf{x}_{2i}]}{\sqrt{Var[y_{1i}|\mathbf{x}_{1i}]}\sqrt{Var[y_{2i}|\mathbf{x}_{2i}]}}$$
$$= \frac{E[Cov[y_{1i},y_{2i}|\mathbf{x}_{1i},\mathbf{x}_{2i},\varepsilon_{1i},\varepsilon_{2i}]]+Cov[E[y_{1i}|\mathbf{x}_{1i},\varepsilon_{1i}],E[y_{2i}|\mathbf{x}_{2i},\varepsilon_{2i}]]}{\sqrt{E[Var[y_{1i}|\mathbf{x}_{1i},\varepsilon_{1i}]]+Var[E[y_{1i}|\mathbf{x}_{1i},\varepsilon_{1i}]]}\sqrt{E[Var[y_{2i}|\mathbf{x}_{2i},\varepsilon_{2i}]]+Var[E[y_{2i}|\mathbf{x}_{2i},\varepsilon_{2i}]]}}.$$

For convenience, let

(3-4) $\mu_{ji} = \lambda_{ji}\exp(\sigma_j^2/2)$.

The terms in the denominator were derived earlier. The unconditional variance is

(3-5)
$$Var[y_{ji}|\mathbf{x}_{ji}] \quad = \lambda_{ji}\exp(\sigma_j^2/2)[1+\lambda_{ji}\exp(\sigma_j^2/2)[\exp(\sigma_j^2)-1]]$$
$$= \mu_{ji}\{1+\mu_{ji}[\exp(\sigma_j^2)-1]\}, j=1,2.$$

For the terms in the numerator, the first is zero, since conditioned on ε_{1i} and ε_{2i}, y_{1i} and y_{2i} (given $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$) are independent. Thus, what remains to derive is

(3-6)
$$Cov\{E[y_{1i}|\mathbf{x}_{1i},\varepsilon_{1i}],E[y_{2i}|\mathbf{x}_{2i},\varepsilon_{2i}]\} = Cov\{[\lambda_{1i}\exp(\sigma_1\varepsilon_{1i})],[\lambda_{2i}\exp(\sigma_2\varepsilon_{2i})]\}$$
$$= (\lambda_{1i}\lambda_{2i})Cov\{[\exp(\sigma_1\varepsilon_{1i})],[(\exp(\sigma_2\varepsilon_{2i})]\}$$
$$= (\lambda_{1i}\lambda_{2i})\{E[\exp(\sigma_1\varepsilon_{1i})\exp(\sigma_2\varepsilon_{2i})]-E[\exp(\sigma_1\varepsilon_{1i})]E[\exp(\sigma_2\varepsilon_{2i})]\}.$$

The two conditional means are the means for the univariate lognormals,

(3-7) $E[\exp(\sigma_j\varepsilon_{ji})] = \exp(\sigma_j^2/2)$.

The remaining term is straightforward;

(3-8)
$$E[\exp(\sigma_1\varepsilon_{1i})\exp(\sigma_2\varepsilon_{2i})] = E[\exp(\sigma_1\varepsilon_{1i}+\sigma_2\varepsilon_{2i})]$$
$$= \exp[(\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)/2].$$

Combining terms and manipulating the expression produces

(5-9)
$$Cov[y_{1i},y_{2i}|x_{1i},x_{2i}] = \left[\lambda_{1i}\exp(\sigma_1^2/2)\right]\left[\lambda_{2i}\exp(\sigma_2^2/2)\right]\left[\exp(\rho\sigma_1\sigma_2)-1\right]$$
$$= \mu_{1i}\mu_{2i}\left[\exp(\rho\sigma_1\sigma_2)-1\right].$$

Combining all terms and simplifying (slightly), we obtain the final result

$$(5\text{-}10) \qquad Corr(y_{1i}, y_{2i} \mid \mathbf{x}_{1i}, \mathbf{x}_{2i}) = \frac{\sqrt{\mu_{1i}\mu_{2i}}\left[\exp(\rho\sigma_1\sigma_2) - 1\right]}{\sqrt{1 + \mu_{1i}[\exp(\sigma_1^2) - 1]}\sqrt{1 + \mu_{2i}[\exp(\sigma_2^2) - 1]}} .$$

How this relates to ρ is unclear. It has the same sign, but the magnitudes are likely to be essentially unrelated. We will examine it in the application below.

Finally, RWM (2003) extended this development to a panel data setting. Their random effects model is

$$\ln\lambda_{it,1} = \alpha_1 + \mathbf{x}_{it,1}'\boldsymbol{\beta}_1 + u_{i,1} + \varepsilon_{it,1} = \alpha_1 + \mathbf{x}_{it,1}'\boldsymbol{\beta}_1 + v_{it,1},$$

$$\ln\lambda_{it,2} = \alpha_2 + \mathbf{x}_{it,2}'\boldsymbol{\beta}_2 + u_{i,2} + \varepsilon_{it,2} = \alpha_2 + \mathbf{x}_{it,2}'\boldsymbol{\beta}_2 + v_{it,2},$$

$$(3\text{-}11) \qquad (\varepsilon_{it,1}, \varepsilon_{it,2}) \sim N_2[(0,0),(\sigma_1,\sigma_2),\rho],$$

$$(u_{i,1}, u_{i,2}) \sim N_2[(0,0),(\omega_1,\omega_2),0].$$

The correlation between $\varepsilon_{it,1}$ and $\varepsilon_{it,2}$ creates the bivariate model. In the notation of our earlier formulation, the correlation of interest, between $v_{it,1}$ and $v_{it,2}$, is

$$(3\text{-}12) \qquad \rho_{12} = \frac{\rho\sigma_1\sigma_2}{\sqrt{\omega_1^2 + \sigma_1^2}\sqrt{\omega_2^2 + \sigma_2^2}} .$$

And the counterparts to σ_1 and σ_2 are the two terms in the denominator. The model also implies a T-variate Poisson-lognormal mixture model for each group for each of the two variables. The implied correlation is $\rho_{ts,j} = \omega_j^2/(\omega_j^2 + \sigma_j^2)$, $j = 1,2$. As they note and discuss, ρ is the correlation between the unique unobservable factors in the two equations. One could, however, misinterpret the magnitude of the value as representative of the correlation between the composed heterogeneity or, worse yet, between the outcome variables, themselves. For example, for their equation system applied to the males in their sample, they report $\rho = 0.599$, $\sigma_1 = 0.996$, $\sigma_2 = 1.244$, $\omega_1 = 0.795$ and $\omega_2 = 1.195$. The computation above produces $\rho_{12} = 0.276$. The calculation is relevant because the unobservable propensities are difficult to partition neatly into time varying and time invariant parts. It is speculative to assume that ρ in isolation captures the full correlation of the unobservables apart from persistent, time invariant components (and leaves $u_{i,j}$ truly unexplained). We will revisit the computation of the implied correlation between the two outcomes below.

4. Applications

In "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation," Riphahn, Wambach and Million (2003) employed a part of the German Socioeconomic Panel (GSOEP) data set to analyze two count variables, *DocVis*, the number of doctor visits in the last three months and *HospVis*, the number of hospital visits in the last year. The authors employed a bivariate panel data (random effects) Poisson model to study these two outcome variables. A central focus of the investigation was the role of the choice of private health insurance in the intensity of use of the health care system, i.e., whether the data contain evidence of moral hazard. We will use these data to illustrate the model extensions described above.[8] The authors of this study presented estimates for the Poisson-lognormal model in Section 2.2.2 and the bivariate Poisson model in (3-11). We will analyze the single equation and two part

---

[8] The raw data are published and available for download on the *Journal of Applied Econometrics* data archive website, The URL is given below Table 1.

models in some detail, but only analyze the correlation structure developed for the bivariate Poisson model in Section 5. (We have not proposed any extensions for this model; our analysis has only provided a more detailed interpretation of the existing model results.) In order to keep the amount of reported results to a manageable size, we will also restrict attention to *DocVis*, the count of doctor visits. Analysis of the count of hospital visits is left for further research.

4.1 The Data

The RWM data set is an unbalanced panel of 7,293 individual families observed from one to seven times. The number of observations varies from one to seven (1,525, 1,079, 825, 926, 1,051, 1000, 887) with a total number of observations of 27,326.

The variables in the data file are listed in Table 1 with descriptive statistics for the full sample. They estimated separate equations for males and females and did not report any estimates based on the pooled data. Table 2 reports descriptive statistics for the two subsamples. The figures given all match those reported by RWM. (See their Table II, page 393.) The outcome variables of interest in the study were doctor visits in the last three months and number of hospital visits last year. Histograms for these variables for the full data set are shown in Figures 2 and 3. (Figure 1 was truncated at 20 visits. Figure 2 was truncated at 10. These remove about 200 observations from the sample used to form the figures.)

The base case count model used by the authors included the following variables in addition to the constant term:

$$\mathbf{x}_{it} = \quad (Age,\ Agesq,\ HSat,\ Handdum,\ Handper,\ Married,\ Educ,\ Hhninc,$$
$$Hhkids,\ Self,\ Civil,\ Bluec,\ Working,\ Public,\ AddOn)$$

and a set of year effects,

$$\mathbf{t} = (YEAR1985,\ YEAR1986,\ YEAR1987,\ YEAR1988,\ YEAR1991,\ YEAR1994).$$

The same specification was used for both *DocVis* and *HospVis*. We will use their specification in our count models. The estimated year effects are omitted from the reported results in the paper.

4.2 Functional Forms and Heterogeneity

Table 3 presents estimates of the Poisson regression models for males and females. The pooled (across genders and across time) results appear in the first column. We tested for homogeneity of the coefficient vectors for males and females using a likelihood ratio test; the chi squared statistic is

$$\lambda_{LR} = 2[90097.4 - (42927.6+46275.1)] = 1789.4.$$

This is substantially larger than the critical chi squared with 16 degrees of freedom (26.30), so the hypothesis that the same model applies to males and females is rejected for the Poisson model. The Poisson specification is, itself, rejected in favor of a model with heterogeneity, so we repeated the homogeneity test with the log gamma (negative binomial) results. The log likelihood for the pooled data is -58082.0 – the pooled NB results are not shown – so the LR statistic for the NB model is 678.60, with 17 degrees of freedom.

The immediate impression is that the presence of public insurance and private addon insurance in the pooled model both have a significant influence on usage of physician visits. However, when the models are fit separately for males and femalse, the latter effect is dissipated.

9

It appears that generally, the former effects disappears from the models that account for latent heterogeneity – of the four sets of results in Table 3, the effect of Addon remains significant only in the NB (log gamma) model for females.

The third column of the two groups of estimates present the lognormal model as an alternative specification to the log gamma (negative binomial). These are the counterparts to RWM's results in their Table IV. Our estimates differ slightly; the difference appears small enough to be attributable to difference in the approximation methods. We used a 48 point Hermite approximation. RWM do not note what method they used for the heterogeneous Poisson model. They used a modification of the Hermite quadrature for the bivariate Poisson model. For example, for the log likelihood function, their reported value is -27411.4 vs. our -27408.6. The counterparts for females are -30213.4 for RWM and -30214.7 for ours. Based on the likelihoods, the lognormal model appears to be superior to the negative binomial model. Since the models are not nested, a direct test based on these values is inappropriate. The Vuong statistic suggested in Greene (2007b) equals 2.329 in favor of the lognormal model.

RWM note based on comparing the Poisson-lognormal to the bivariate model that the significance and, in some cases, the signs of the coefficients change with the specification. We find generally, that this applies to the marginal variables, but that the pattern of significance of most of the variables in the equation is extremely stable. The very important exception is the variables that were the focus of the study, the insurance variables. What we find is that as the model is extended to account for latent heterogeneity, the importance of the private insurance variable diminishes consistently.

4.3 The Bivariate Poisson Model

Result (5-10) provides the implied correlation between y_{i1} and y_{i2} in the bivariate Poisson model in which

$$\lambda_{i1} = \exp(\mathbf{x}_{i1}'\boldsymbol{\beta}_1 + \sigma_1\varepsilon_{i1})$$

and

$$\lambda_{i1} = \exp(\mathbf{x}_{i2}'\boldsymbol{\beta}_2 + \sigma_2\varepsilon_{i2})$$

where $\quad (\varepsilon_{i1},\varepsilon_{i2}) \sim N_2[(0,1),(1,1),\rho].$

Munkin and Trivedi (1999) used this specification in a model for the joint determination of the counts of emergency room visits and hospital visits for a sample of 4,406 elderly Americans drawn from the National Medical Expenditure Survey from 1987 and 1988. The authors report the estimates of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ and, in addition, $\hat{\sigma}_1 = 1.39$, $\hat{\sigma}_2 = 1.36$ and $\hat{\rho} = 0.92$. The last of these might lead one to suspect that emergency room visits and hospital visits were extremely highly correlated. However, as derived in Section 5, the 0.92 reflects only the correlation between the latent effects in the conditional means. In order to evaluate the correlation between the two outcomes, we propose to evaluate (3-10) at the sample observations, and then average the outcomes. However, without the Munkin and Trivedi data in hand, we resort to an approximation. To a reasonable approximation, the sample average of λ_i evaluated at the individual data will equal the mean of the outcome variable. (The result is exact in the base case Poisson model – this is the likelihood equation for the constant term.) The authors report sample means of $\bar{y}_1 = 0.26$ and $\bar{y}_2 = 0.92$. Thus, in (3-10), we use

$$\mu_{i1} \approx \bar{y}_1 \exp(.5 \times 1.39^2) = 0.683162$$
$$\mu_{i2} \approx \bar{y}_2 \exp(.5 \times 1.36^2) = 0.756409,$$

and complete the computation with a hand calculator. The result is an estimated correlation of 0.668929, which is substantially less than 0.92. Munkin and Trivedi do not report the sample correlation of their two outcome variables, so we cannot measure the implied estimate against the sample statistic.

The RWM model is

(4-1)
$$\ln\lambda_{it,1} = \alpha_1 + \mathbf{x}_{it,1}'\boldsymbol{\beta}_1 + u_{i,1} + \varepsilon_{it,1} = \alpha_1 + \mathbf{x}_{it,1}'\boldsymbol{\beta}_1 + v_{it,1},$$

$$\ln\lambda_{it,2} = \alpha_2 + \mathbf{x}_{it,2}'\boldsymbol{\beta}_2 + u_{i,2} + \varepsilon_{it,2} = \alpha_2 + \mathbf{x}_{it,2}'\boldsymbol{\beta}_2 + v_{it,2},$$

$$(\varepsilon_{it,1},\varepsilon_{it,2}) \sim N_2[(0,0),(\sigma_1,\sigma_2),\rho],$$

$$(u_{i,1},u_{i,2}) \sim N_2[(0,0),(\omega_1,\omega_2),0].$$

for which we derived

(4-2)
$$\rho_{12} = \frac{\rho\sigma_1\sigma_2}{\sqrt{\omega_1^2 + \sigma_1^2}\sqrt{\omega_2^2 + \sigma_2^2}} .$$

The result in (3-10) can be used by using this expression for ρ and the two standard deviations, $\tau_1 = \sqrt{\omega_1^2 + \sigma_1^2}$ and $\tau_2 = \sqrt{\omega_2^2 + \sigma_2^2}$ for σ_1 and σ_2 in (3-10). The authors did not report the full set of estimated parameters (they omitted the coefficients on the year dummy variables). Rather than reestimate the full bivariate Poisson model, we proceeded as follows. Each of the equations in (4-1) can be consistently estimated in isolation. Moreover, we note that the marginal distribution of each of the observations, (i,t), in the sample, has a marginal Poisson distribution with normally distributed heterogeneity with mean zero and standard deviation τ_j. Thus, we estimated the four equations singly using the lognormal heterogeneity model discussed in Section 2.2.2. This provides consistent, albeit inefficient estimators of the parameters of the four equations. These are shown in Table 4. (RWM's counterparts are shown in Table 5 for comparison.). The estimated variance, σ^2, in each of these equations is an estimate of $\tau^2 = \sigma_\varepsilon^2 + \omega_u^2$ in the RWM model. This is also shown in Table 4. Only an estimate of ρ is needed to complete the calculations in (3-4), (3-10) and (3-12). We used the estimate of ρ reported by RWM for males and females, which appears in the last row of Table 4. For comparison purposes, we have decomposed the estimated variance from our estimates using the implied analysis of variance in RWM. The computations appear at the bottom of Table 4. The proportion denoted "p" in the table inferred from the RWM results is used to decompose the estimated variance from our model. With these statistics in hand, and with the estimated coefficient vectors, we are able to compute the implied correlations for the two models (males and females). Using (3-12), we obtain means of individual specific estimates of the correlations of the outcome variables of 0.06938 for males and 0.05795 for females. These are an order of magnitude less than the estimate of ρ reported in the paper, and moreover, only about half of the actual correlation between the outcomes in the data.

5. Conclusions

One of the recent applications of the methods developed in this paper is in a type of bivariate count model. We found that in these models, the introduction of a "correlation coefficient" into the model within the conditional means provides only a partial indication of the degree of correlation between the outcome variables. We derived the relationship between the structural parameters and the reduced form correlation between the outcome variables in the bivariate Poisson model. In the application carried out in this paper, we find that the estimated correlation coefficient is far higher than the actual correlation of the variables in the model. Moreover, the implied correlation

coefficient based on the model estimates, which is a function of the data and thus varies by observation, does a strikingly poor job of reproducing the actual, simple correlation of the outcome variables and, moreover, appears, on average in these data, to be a full order of magnitude less than the simple reported correlation coefficient. This calls into question the precise interpretation of this part of the model and whether this form of correlation is an effective approach to modeling the correlation across related count data outcome variables.

9. References

Asplund, M. and R. Sandin, "The Number of Firms and Production Capacity in Relation to Market Size," *The Journal of Industrial Economics*, Vol. 47, No. 1 (Mar., 1999), pp. 69-85

Butler, J. and Moffitt, R., "A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model," *Econometrica*, 50, 1982, pp. 761-764.

Cameron, A., and P. Trivedi. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests." *Journal of Applied Econometrics,* 1, 1986, pp. 29–54.

Cameron, A., and P. Trivedi. "Regression Based Tests for Overdispersion in the Poisson Model." *Journal of Econometrics,* 46, 1990, pp. 347–364.

Cameron, C., and P. Trivedi. *Regression Analysis of Count Data.* New York: Cambridge University Press, 1998.

Cameron, C. and P. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge, 2005.

Contoyannis, C., A. Jones and N. Rice, "The Dynamics of Health in the British Household Panel Survey," *Journal of Applied Econometrics*, 19, 4, 2004, pp. 473-503.

Gourieroux, C., and A. Monfort. *Simulation-Based Methods Econometric Methods*. Oxford: Oxford University Press, 1996.

Gourieroux, C., A. Monfort, and A. Trognon. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica,* 52, 1984, pp. 701–720.

Greene, W., "Functional Form and Heterogeneity in Models for Count Data," Working Paper 07-10, Department of Economics, Stern School of Business, New York University, 2007.

Greene, W., *Econometric Analysis, 6th ed.*, Prentice Hall, Englewood Cliffs, 2008a.

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

Hilbe, Joseph M. (2007). *Negative Binomial Regression*, Cambridge, UK: Cambridge University Press.

Jung, R. and R. Winkelmann, "Two Aspects of Labor Mobility: A Bivariate Poisson Regression Approach," *Empirical Economics*, 18, 1993, pp. 543-556.

Karlis, D. and I. Ntzoufras, "Analysis of Sports Data By Using Bivariate Poisson Models," *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52, 3, 2003, pp. 381–393.

Kocherlakota, S. and K. Kocherlakota, *Bivariate Discrete Distributions*, New York, Marcel Dekker, 1992.

Million, A., "Models for Correlated Count Data," Unpublished dissertation, University of Munich, 1998.

Munkin, M. and P. Trivedi, "Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application," *Econometrics Journal*, 2, 1999, pp. 29-49.

Riphahn, R., A. Wambach, and A. Million, A., "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation" *Journal of Applied Econometrics*, Vol. 18, No. 4, 2003, pp. 387-405

Train, K., *Discrete Choice Models with Simulation*, Cambridge University Press, Cambridge, 2003.

Vuong, Q., "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses," *Econometrica*, 57, 1989, pp. 307-344.

Wang, P., I. Cockburn, and L. Puterman, "Analysis of Patent Data − A Mixed-Poisson-Regression-Model Approach," *Journal of Business and Economic Statistics*, 16, 1, 1998, pp. 27-41

Winkelmann, R., *Econometric Analysis of Count Data*, Springer Verlag, Heidelberg, 4th ed. 2003.

Table 1. Variables in German Health Care Data File

Variable	Measurement	Mean	Standard Deviation
ID	household identification, 1,...,7293		
YEAR	calendar year of the observation	1987.82	3.17087
YEAR1984	dummy variable for 1984 observation	.141770	.348820
YEAR1985	dummy variable for 1984 observation	.138842	.345788
YEAR1986	dummy variable for 1984 observation	.138769	.345712
YEAR1987	dummy variable for 1984 observation	.134158	.340828
YEAR1988	dummy variable for 1984 observation	.164056	.370333
YEAR1991	dummy variable for 1984 observation	.158823	.365518
YEAR1994	dummy variable for 1984 observation	.123582	.329110
AGE	age in years	43.5257	11.3302
AGESQ**	age saquared/1000	2.02286	1.00408
FEMALE	female = 1; male = 0	.478775	.499558
MARRIED	married = 1; else = 0	.758618	.427929
HHKIDS	children under age 16 in the household = 1; else = 0	.402730	.490456
HHNINC***	household nominal monthly net income, German marks / 10000	.352084	.176908
EDUC	years of schooling	11.3206	2.32489
WORKING	employed = 1; else = 0	.677048	.467613
BLUEC	blue collar employee = 1; else = 0	.243761	.429358
WHITEC	white collar employee = 1; else = 0	.299605	.458093
SELF	self employed = 1; else = 0	.0621752	.241478
CIVIL	civil servant = 1; else = 0	.0746908	.262897
HAUPTS	highest schooling degree is Hauptschul = 1; else = 0	.624277	.484318
REALS	highest schooling degree is Realschul = 1; else = 0	.196809	.397594
FACHHS	highest schooling degree is Polytechnical= 1; else = 0	.0408402	.197924
ABITUR	highest schooling degree is Abitur = 1; else = 0	.117031	.321464
UNIV	highest schooling degree is university = 1; else = 0	.0719461	.258403
HSAT	health satisfaction, 0 - 10	6.78543	2.29372
NEWHSAT*'**	health satisfaction, 0 - 10	6.78566	2.29373
HANDDUM	handicapped = 1; else = 0	.214015	.410028
HANDPER	degree of handicap in pct, 0 - 100	7.01229	19.2646
DOCVIS	number of doctor visits in last three months	3.18352	5.68969
DOCTOR**	1 if DOCVIS > 0, 0 else	629108	.483052
HOSPVIS	number of hospital visits in last calendar year	.138257	.884339
HOSPITAL**	1 of HOSPVIS > 0, 0 else	.0876455	.282784
PUBLIC	insured in public health insurance = 1; else = 0	.885713	.318165
ADDON	insured by add-on insurance = 1; else = 0	.0188099	.135856

Data source: http://qed.econ.queensu.ca/jae/2003-v18.4/riphahn-wambach-million/.

From Riphahn, R., A. Wambach and A. Million "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation," *Journal of Applied Econometrics*, 18, 4, 2003, pp. 387-405.

Notes: * *NEWHSAT = HSAT*; 40 observations on HSAT recorded between 6 and 7 were changed to 7.
　　　** Transformed variable not in raw data file.
　　　*** Divided by 1,000 rather than 10,000 by RWM. We used this scale to ease comparison of coefficients.

Table 2. Descriptive Statistics by Gender

Variable	Males		Females	
	Mean	Standard Dev.	Mean	Standard Dev.
YEAR	1987.84	3.19003	1987.80	3.14985
YEAR1984	.141613	.348665	.141940	.349002
YEAR1985	.138875	.345828	.138806	.345757
YEAR1986	.138173	.345094	.139418	.346395
YEAR1987	.134171	.340848	.134144	.340820
YEAR1988	.162396	.368826	.165864	.371973
YEAR1991	.157551	.364332	.160208	.366813
YEAR1994	.127220	.333231	.119621	.324530
AGE	42.6526	11.2704	44.4760	11.3192
AGESQ	1.94628	.987385	2.10623	1.01543
FEMALE	.000000	.000000	1.00000	.000000
MARRIED	.765148	.423921	.751510	.432154
HHKIDS	.412975	.492386	.391577	.488122
HHNINC	.359054	.173564	.344495	.180179
EDUC	11.7287	2.43649	10.8764	2.10911
WORKING	.850312	.356777	.488420	.499885
BLUEC	.340237	.473805	.138730	.345677
WHITEC	.299937	.458246	.299243	.457944
SELF	.0856561	.279865	.0366124	.187815
CIVIL	.117812	.322397	.0277459	.164250
HAUPTS	.601137	.489682	.649469	.477155
REALS	.176086	.380907	.219369	.413835
FACHHS	.0536404	.225315	.0269051	.161812
ABITUR	.146949	.354068	.0844608	.278088
UNIV	.0961876	.294859	.0455553	.208527
HSAT	6.92436	2.25148	6.63417	2.32951
NEWHSAT	6.92459	2.25148	6.63441	2.32953
HANDDUM	.227295	.419007	.199559	.399538
HANDPER	8.13371	20.3288	5.79143	17.9562
DOCVIS	2.62571	5.21121	3.79080	6.11113
DOCTOR	.559503	.496464	.704884	.456112
HOSPVIS	.127782	.930209	.149660	.831416
HOSPITAL	.0779330	.268076	.0982191	.297622
PUBLIC	.861055	.345902	.912558	.282492
ADDON	.9175525	.131323	.0201789	.140617
Sample Size	14,243		13,083	

Table 3. Poisson Models and Heterogeneity in Poisson (t ratios in parentheses)

Variable	Pooled Poisson	Males			Females		
		Poisson	Log gamma	Lognormal	Poisson	Log gamma	Lognormal
Constant	2.639 (39.46)	2.771 (28.85)	3.1488 (13.74)	2.8079 (11.26)	2.546 (28.54)	3.0245 (15.03)	2.7556 (12.47)
AGE	−0.00732 (−2.64)	−0.02387 (−5.44)	−0.03983 (−4.07)	−0.05858 (−5.51)	−0.01320 (−3.64)	−0.03119 (−3.78)	−0.04485 (−4.90)
AGESQ	0.1407 (4.54)	0.3693 (7.45)	0.5467 (4.77)	0.7853 (6.45)	0.1794 (4.46)	0.3727 (4.02)	0.5421 (5.27)
HSAT	−0.2149 (−151.9)	−0.2253 (−104.1)	−0.2392 (−42.44)	−0.2650 (−50.93)	−0.2034 (−108.3)	−0.2080 (−47.30)	−0.2225 (−46.54)
HANDDUM	0.1011 (8.71)	0.06899 (4.09)	−0.02090 (−0.46)	−0.01093 (−0.23)	0.1379 (8.55)	0.1133 (2.79)	0.1011 (2.48)
HANDPER	0.001992 (10.73)	0.002858 (10.04)	0.006614 (8.05)	0.007398 (9.08)	0.002414 (9.48)	0.004359 (5.92)	0.004432 (6.14)
MARRIED	0.02058 (2.32)	0.05831 (3.89)	0.06582 (2.18)	0.1276 (3.67)	0.02718 (2.39)	0.02816 (1.13)	0.04590 (1.63)
EDUC	−0.01483 (−7.96)	−0.02348 (−8.43)	−0.02623 (−4.59)	−0.02297 (−3.43)	0.01473 (5.65)	0.007725 (1.36)	0.01318 (2.09)
HHNINC	−0.1729 (−7.27)	−0.2220 (−5.93)	−0.1917 (−2.48)	−0.1257 (−1.44)	−0.2063 (−6.53)	−0.1624 (−2.57)	−0.1417 (−1.92)
HHKIDS	−0.1108 (−12.86)	−0.07598 (−5.75)	−0.08440 (−3.32)	−0.09013 (−2.94)	−0.1338 (−11.63)	−0.1243 (−4.91)	−0.1360 (−4.81)
SELF	−0.2914 (−16.18)	−0.2110 (−8.98)	−0.2179 (−5.02)	−0.3590 (−6.81)	−0.2175 (−7.47)	−0.2424 (−4.51)	−0.2885 (−4.55)
CIVIL	−0.05026 (−2.64)	0.09144 (3.78)	0.08411 (1.56)	0.01916 (0.32)	−0.07113 (−1.91)	−0.01982 (−0.34)	−0.03188 (−0.39)
BLUEC	−0.08920 (−9.01)	0.01779 (1.24)	.03706 (1.20)	−0.03137 (−.93)	−0.03543 (−2.38)	−0.04010 (−1.31)	−0.09991 (−2.81)
WORKING	−0.07478 (−7.62)	−0.05539 (−3.17)	−0.01545 (−0.38)	0.03119 (0.78)	0.01490 (1.29)	0.03046 (1.23)	0.03851 (1.38)
PUBLIC	0.1145 (7.32)	0.1001 (4.27)	.09340 (1.83)	0.05150 (0.91)	0.1312 (6.22)	0.09530 (2.44)	0.08076 (1.72)
ADDON	0.06084 (2.39)	0.06655 (1.63)	0.05506 (0.50)	0.1954 (1.81)	0.02071 (0.63)	0.03088 (0.32)	0.1175 (1.25)
θ			**0.5707 (59.96)**			**0.8289 (64.44)**	
κ			1.7522 (59.96)			1.2064 (64.44)	
$\sigma(\varepsilon)$			1.9874 (72.19)	**1.2520 (104.61)**		1.4757 (84.33)	**1.0608 (114.80)**
$\sigma(h)$			1.3237 (119.92)	4.2651 (29.49)		1.1.0984 (128.89)	2.5325 (41.13)
ln *L*	−89641.2	−42774.7	−27480.4	−27408.6	−45900.2	−30262.3	−30214.7
n	27326	14243			13083		

Notes: Estimated coefficients for year dummy variables, excluding year 1984, are not reported.

θ = the estimated parameter for the log gamma (NB) model

$\kappa = 1/\theta = \text{Var}[h]$ for log gamma model.

$\sigma(\varepsilon) = \sqrt{\psi'(\theta)} = \text{Var}(\ln h_i)$ for the log gamma model. Estimated directly for the lognormal model.

$\sigma(h) = \sqrt{\kappa}$ for the log gamma model, $\sqrt{\exp(\sigma^2)[\exp(\sigma^2)-1]}$ for the lognormal model.

Table 4. Single Equations Estimates of Bivariate Poisson Models

Variable	Males		Females	
	DocVis	**HospVis**	**DocVis**	**HospVis**
Constant	2.95829046	-.71769354	2.65970461	-.92395935
AGE	-.06039694	-.01495334	-.03467601	-.04182897
AGESQ	.81046943	.17526254	.40417087	.32800832
HSAT	-.26979714	-.28467610	-.22492822	-.21283211
HANDDUM	-.10360835	-.15938867	.14118341	.04898526
HANDPER	.00811821	.00687895	.00436759	.01053000
MARRIED	.06801734	-.11266046	.09394893	-.03039990
EDUC	-.02141476	-.05611756	.00669951	-.02245035
HHNINC	-.09317739	.28561716	-.10306991	.45290684
HHKIDS	-.05833037	.06532961	-.18274059	.02830548
SELF	-.27162867	-.07241379	-.27422854	-.12053254
CIVIL	.03028441	-.15292197	-.00966680	.16538415
BLUEC	-.05264588	.19299252	-.09417967	-.31982329
WORKING	.00747418	-.29732586	.01608792	.01324764
PUBLIC	.04537642	-.25102523	.08274879	.07493233
ADDON	.15634672	.61629605	.11353610	.28858560
YEAR1984	.00000000	.00000000	.00000000	.00000000
YEAR1985	.01555697	.38790097	-.02109554	.18039499
YEAR1986	.14371463	-.03957923	.11756445	.28653104
YEAR1987	.16606851	.06845882	-.10021643	.12299139
YEAR1988	.04018257	-.05038515	-.16888748	.43457292
YEAR1991	.02525603	-.06140680	-.06445527	.44076348
YEAR1994	.28195404	.07614490	.26471138	.13759322
σ	1.23777937	1.78190925	1.04809696	1.47269236
σ^2	1.53209777	3.17228488	1.09850724	2.16882279
σ_ε	.96737031	1.28460990	.79772504	1.00736894
ω_u	.77219975	1.23371888	.67981026	1.07425817
ρ_{DW}	0.276		0.201	
Average Correlation	0.06938		0.05795	
Sample Corr(Doc,Hosp)	0.1477		0.1255	
RWM Reported Results				
$\sigma(\varepsilon_{it})$	0.996	1.244	0.822	1.053
$\omega(u_i)$	0.795	1.195	0.701	1.123
$\sigma^2(\varepsilon_{it})$	0.992	1.548	0.676	1.109
$\omega^2(u_i)$	0.632	1.428	0.491	1.261
$p=\sigma^2(\varepsilon_{it})/[\sigma^2(\varepsilon_{it})+\omega^2(u_i)]$	0.6108	0.5202	0.5793	0.4679
$\tau=[\sigma^2(\varepsilon_{it})+\omega^2(u_i)]^{1/2}$	1.274	1.725	1.080	1.540
ρ_{DH}	0.276		0.201	
ρ	0.490		0.386	

Table 5. RWM Estimated Bivariate Poisson Models

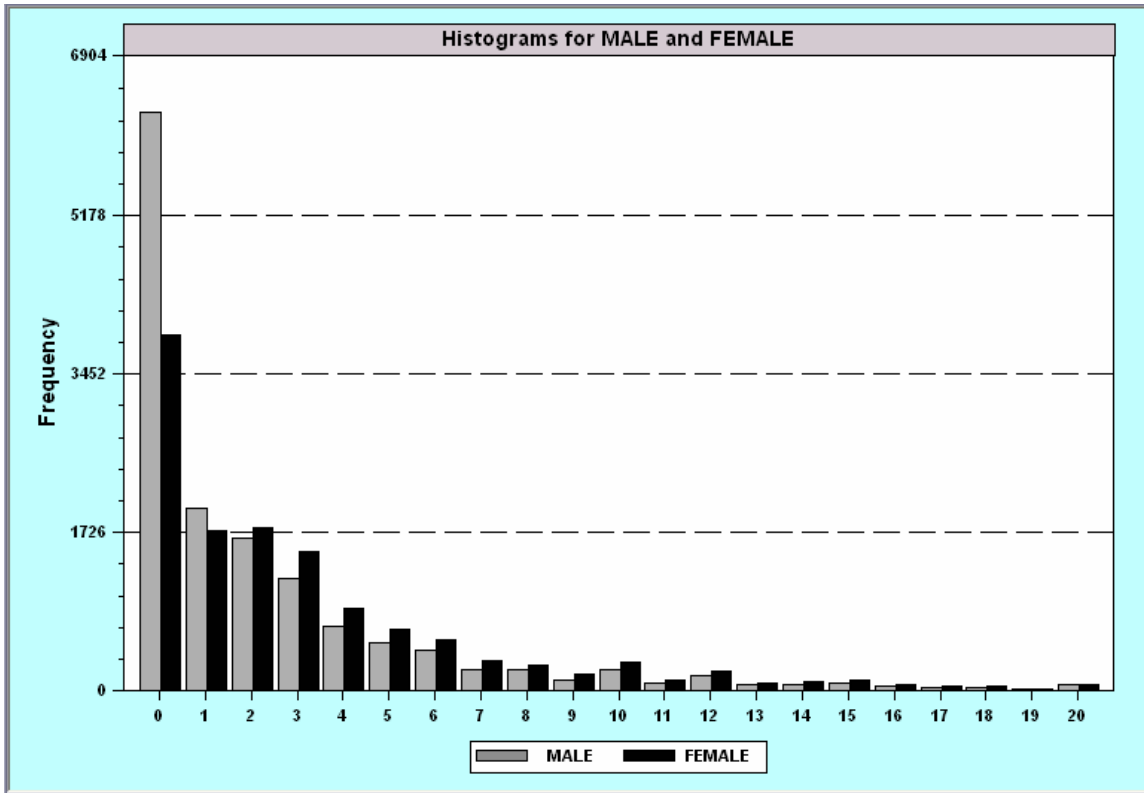| Variable	Males		Females	
	DocVis	**HospVis**	**DocVis**	**HospVis**
Constant	2.563	−0.206	2.423	−1.567
AGE	−0.060	−0.077	−0.040	−0.032
AGESQ	0.823	0.942	0.499	0.234
HSAT	−0.237	−0.243	−0.191	−0.196
HANDDUM	−0.029	−0.086	0.063	0.039
HANDPER	0.007	0.008	0.004	0.010
MARRIED	0.085	−0.054	0.009	−0.044
EDUC	−0.022	−0.051	0.014	−0.015
HHNINC	−0.090	0.375	−0.107	0.407
HHKIDS	−0.059	0.103	−0.117	0.073
SELF	−0.356	−0.196	−0.256	−0.117
CIVIL	−0.011	−0.086	−0.069	0.281
BLUEC	−0.029	0.173	−0.034	−0.320
WORKING	0.041	−0.026	0.002	−0.014
PUBLIC	0.075	−0.136	0.058	0.246
ADDON	0.090	0.549	0.096	0.219
σ (ε_{it})	0.996	1.244	0.822	1.053
ω (u_i)	0.795	1.195	0.701	1.123
ρ	0.490		0.386	

Figure 1   Histograms for DocVis
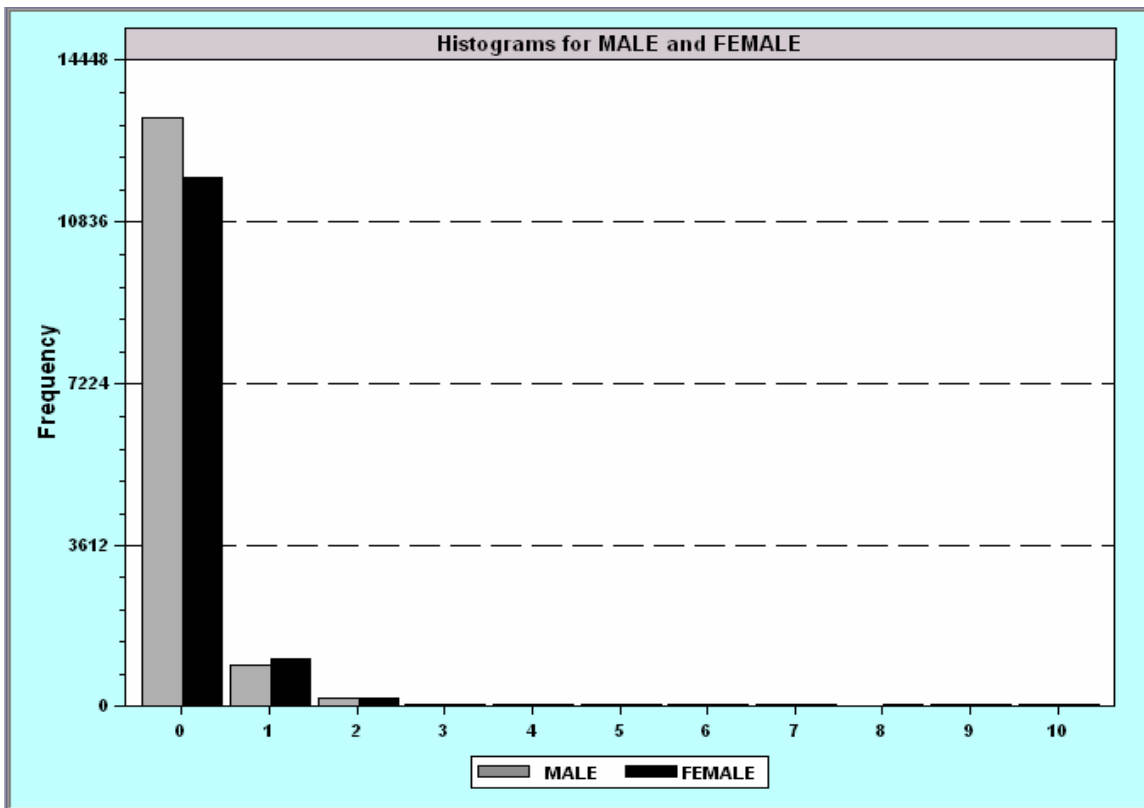
Figure 2   Histograms for HospVis