Journal of Business & Economic Statistics, October 1995, Vol. 13, No. 4

Duration Dependence and Dispersion in Count-Data Models

Rainer Winkelmann
Department of Economics, University of Canterbury, Christchurch, New Zealand

This article explores the relation between nonexponential waiting times between events and the distribution of the number of events in a fixed time interval. It is shown that within this framework the frequently observed phenomenon of overdispersion—that is, a variance that exceeds the mean—is caused by a decreasing hazard function of the waiting times, whereas an increasing hazard function leads to underdispersion. Using the assumption of iid gamma-distributed waiting times, a new count-data model is derived. Its use is illustrated in two applications, the number of births and the number of doctor consultations.

KEY WORDS: Gamma distribution; Negative binomial distribution; Overdispersion; Poisson process; Renewal theory.

1. INTRODUCTION

The basic regression model for count data (number of events in a given interval of time) is the Poisson model, where $Y \mid x \sim \text{Poisson}(E(Y \mid x) = \exp(x'\beta))$. In econometric applications this model is usually inadequate. In particular, the Poisson model imposes the restriction that the conditional variance equals the conditional mean, but typically the conditional variance exceeds the conditional mean (overdispersion). Occasionally, the conditional mean exceeds the conditional variance (underdispersion). In either case, estimation based on the Poisson model is inefficient and leads to biased inference (Winkelmann 1994). Potential solutions to this problem include the following:

1.1 Quasi-likelihood Methods

Estimation is based on the first or the first two moments only. Usually $E(Y \mid x) = \exp(x'\beta)$ and $\text{var}(Y \mid x) = \phi \exp(x'\beta)$, where $0 < \phi < 1$ indicates underdispersion and $\phi > 1$ overdispersion. Standard references are Gourieroux, Monfort, and Trognon (1984) and, in the statistical literature on generalized linear models, McCullagh and Nelder (1989).

1.2 Mixture Models for Heterogeneity

The Poisson mean is itself a random variable. If this is gamma, we obtain $Y \mid x \sim$ NB with $E(Y \mid x) = \exp(x'\beta)$, $\text{var}(Y \mid x) = \exp(x'\beta) + \phi \exp(x'\beta)^{k+1}$, where NB stands for the negative binomial distribution. For $k = 0$ and $k = 1$, this gives the negative binomial model denoted by Cameron and Trivedi (1986) as NEGBIN I and NEGBIN II, respectively. Winkelmann and Zimmermann (1995) left k unrestricted. Unobserved heterogeneity always causes overdispersion.

The following three approaches relax the assumption of independent and stationary increments of the Poisson process.

1.3 Nonhomogeneous Poisson Process

Successive events are independent and the process intensity varies as a function of (calendar) time. Lawless (1987) formulated a proportional intensity model with $\lambda_x(t) = \lambda_0(t) \exp(x'\beta)$ and

$$Y \mid x \sim \text{Poisson}\left(E(Y \mid x) = \int_0^T \lambda_0(t)\, dt \exp(x'\beta)\right),$$

where $\lambda_x(t)$ is the instantaneous risk of an occurrence. In this approach, the current intensity is a function of calendar time only and, in particular, independent of the previous history of the process. In general, estimation of all the model parameters requires information on occurrence times. Lawless (1987) gave conditions under which β is identified from a sample of counts.

1.4 Occurrence Dependence

Successive events are dependent: The current probability for an occurrence depends on the *number* of previous event occurrences. Such models are said to display true contagion. They have been intensively discussed in the literature on accident proneness (Arbous and Kerrich 1951; Feller 1943). Gurland (1959) gave a contagious discrete-time model due to Polya (in which an occurrence *increases*, and a nonoccurrence *decreases*, the probability of a future occurrence) that leads to the negative binomial distribution. This provides an alternative derivation to Section 1.2—a third derivation of the negative binomial model can be based on a model for random colonies (Gurland 1959). The variety of circumstances in which the negative binomial model arises reflects an intrinsic identification problem: This probabilistic model cannot distinguish between unobserved heterogeneity ("apparent contagion") and occurrence dependence ("true contagion").

Other models for occurrence dependence have been developed. In particular, the hurdle model of Mullahy (1986) and the spell-specific heterogeneity model by Gourieroux and Visser (1992) can also be viewed as such models.

1.5 Duration Dependence

Waiting times between events are independent but not exponential (which would lead to the Poisson distribution for counts). Instead they have some other distribution with a nonconstant hazard function. If the hazard function is a decreasing function of time, the distribution displays negative duration dependence. If the hazard function is an increasing function of time, the distribution displays positive duration dependence. In both cases, the conditional probability of a current occurrence depends on the time since the last occurrence rather than on the number of previous events. Events are "dependent" in the sense that the occurrence of at least one event (in contrast to none) up to time t influences the probability of a further occurrence in $t + \Delta t$.

In regression analysis the first two approaches are the standard approaches, and the third and the fourth are used to some extent. This article takes the fifth approach, using renewal theory (for iid point processes that are not necessarily Poisson) to study the link between duration dependence and dispersion. It is shown that negative duration dependence (asymptotically) causes overdispersion and positive duration dependence underdispersion.

Furthermore, specific parametric assumptions are used to derive a generalized count-data model that nests the Poisson regression model through a single parametric restriction and allows for both over- and underdispersion. As a starting point, note that the Poisson process can be thought of as a sequence of independently and identically exponentially distributed waiting times (Cox 1962). To derive a generalized model, I replace the exponential distribution with a less restrictive nonnegative distribution. Possible candidates known from the duration literature are the Weibull, the gamma (including generalized gamma), and the lognormal distributions. Both Weibull and gamma nest the exponential distribution, and both allow for a (monotone) nonconstant hazard—that is, duration dependence. Although the Weibull distribution is preferred in duration analysis for its closed-form hazard function, the gamma distribution is preferred here for its reproductive property: Sums of independent gamma distributions are again gamma distributed. As a result, the probability function of the corresponding gamma count distribution takes a rather simple form.

A corresponding regression model is formulated. Advantages of this generalized model are the following: First, it provides a count-data model of substantially higher flexibility than the Poisson model at the cost of one additional parameter. The Poisson restriction can be tested using a standard Wald test. Second, it provides an interpretation of over- and underdispersion in terms of an underlying sequence of waiting times. Third, the model is easy to implement on any computer software that calculates the incomplete gamma function and has a numerical maximization routine. Fourth, the parametric nature of the approach allows one to calculate, and draw inference on, single probabilities (although it is, of course, susceptible to the critique of being not robust if the model is misspecified).

Areas of potential applicability include the analysis of accident proneness (for instance, airline accidents), labor mobility (the number of changes of employer), the demand for health-care services (as measured by the number of doctor consultations in a given time interval), and, in economic demography, total fertility (the number of births by a woman).

Finally, a cautionary remark toward the distinction between unobserved heterogeneity and duration dependence is in order. The required identifying conditions have been thoroughly studied in the duration literature. In the count-data literature, it is well known that the negative binomial distribution cannot distinguish between unobserved heterogeneity and positive occurrence dependence. There is no reason to assume that the situation is more favorable in the present context. The presence of over- or underdispersion in the gamma count model should not be interpreted as evidence for, but rather as compatible with, duration dependence in the underlying waiting times.

2. SOME DEFINITIONS

I consider a sequence of events for which the occurrence times of (and thus the waiting times between) events are unobserved. We observe the number of events before a fixed point in time. Elementary probability arguments establish that the distributions of the arrival times determine the distribution of the number of events.

Let $\{\tau_k, k \in \mathbf{N}\}$ denote a sequence of *waiting times* between the $(k-1)$th and the kth event. Then, the arrival time of the nth event is given by

$$\vartheta_n = \sum_{k=1}^{n} \tau_k, \qquad n = 1, 2, \ldots. \tag{1}$$

Let N_T represent the total number of events in the open interval between 0 and T. For fixed T, N_T is a *count variable*. It follows from the definitions of N_T and ϑ_n that

$$N_T < n \quad \text{iff} \quad \vartheta_n \geq T. \tag{2}$$

Thus,

$$\begin{aligned} P(N_T < n) &= P(\vartheta_n \geq T) \\ &= 1 - F_n(T), \end{aligned} \tag{3}$$

where $F_n(T)$ is the cumulative distribution function of ϑ_n. Furthermore,

$$\begin{aligned} P(N_T = n) &= P(N_T < n + 1) - P(N_T < n) \\ &= F_n(T) - F_{n+1}(T). \end{aligned} \tag{4}$$

Equation (4) provides the fundamental relation between the distribution of arrival times and the distribution of counts. The probability distribution of N_T can be obtained explicitly for all n from knowledge of the distributions of ϑ_n.

If $\{\tau_i\}$ are iid with common density function $f(\tau)$, the process is called a *renewal process* (see Cox 1962; Feller 1971). N_T gives the number of renewals in $(0, T)$, and $E(N_T)$ is the renewal function. The probability function of N_T was given in (4). But $\vartheta_n = \sum_{i=1}^{n} \tau_i$. Given the renewal density $f(\tau)$, the explicit density of ϑ_n can be derived in several cases using the calculus of Laplace transforms.

The link between $f(\tau)$ and $P_1(N_t)$ is used in Section 3 to establish that negative (positive) duration dependence causes over(under)dispersion. Section 4 provides an explicit parametric model for gamma-distributed waiting times and develops a specification for regression analysis.

3. DURATION DEPENDENCE AND DISPERSION

Without making assumptions on the exact distribution of τ, a limiting result can be obtained. Denote the mean and the variance of the waiting-time distribution by $E(\tau) = \mu$ and $\text{var}(\tau) = \sigma^2$ and the coefficient of variation by $\upsilon = \sigma/\mu$. Define the hazard function

$$\lambda(t) = \frac{f(t)}{1 - F(t)},$$

where $f(t)$ and $F(t)$ are the density function and the cumulative probability function of τ, respectively. The distribution displays negative duration dependence for $d\lambda(t)/dt < 0$ and positive duration dependence for $d\lambda(t)/dt > 0$. Assume that the hazard function is monotonic. Then

$$\frac{d\lambda(t)}{dt} \left.\begin{matrix} < \\ = \\ > \end{matrix}\right\} 0 \Longrightarrow \upsilon \left.\begin{matrix} > \\ = \\ < \end{matrix}\right\} 1$$

(see Barlow and Proschan 1965, p. 33).

Theorem. Let $\{\tau_i\}$ be a sequence of independent, positive, identically distributed random variables and $N(t)$ the number of renewals between 0 and t. Assume that the densities of the waiting times τ_i have a monotonic hazard function. Then negative (positive) duration dependence of the waiting time densities causes over(under)dispersion of the distribution of $N(t)$ as $t \to \infty$.

Proof. $N(t)$ is asymptotically normal distributed with

$$N(t) \overset{\text{asy}}{\sim} \text{normal}\left(\frac{t}{\mu}, \frac{\sigma^2 t}{\mu^3}\right) \tag{5}$$

as $t \to \infty$ (Cox 1962, p. 40).

The ratio of variance to mean of the limiting distribution is given by

$$\frac{\text{variance}}{\text{mean}} \sim \frac{\sigma^2 t}{\mu^3}\frac{\mu}{t} = \frac{\sigma^2}{\mu^2}. \tag{6}$$

Thus, the variance mean ratio is greater (less) than 1 iff the coefficient of variation of the waiting times $\upsilon = \sigma/\mu$ is greater (less) than 1. For positive duration, dependence $\upsilon < 1$ and the count distribution is underdispersed. For negative duration, dependence $\upsilon > 1$ and the count distribution is overdispersed.

The exponential distribution has coefficient of variation $\upsilon = 1$, leading to equidispersion. This result is exact, whereas (6) is only a limiting result.

4. A GAMMA COUNT MODEL

It will be assumed that the waiting times τ_k are identically and independently gamma distributed. Dropping the index k the density can be written as

$$f(\tau; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\tau^{\alpha-1}e^{-\beta\tau}, \qquad \alpha, \beta \in \mathbf{R}^+ \tag{7}$$

for $\tau > 0$. It has mean $E(\tau) = \alpha/\beta$ and variance $\text{var}(\tau) = \alpha/\beta^2$. The hazard function $\lambda(\tau)$ obeys the equation

$$\frac{1}{\lambda(\tau)} = \int_0^\infty e^{-\beta u}\left(1 + \frac{u}{\tau}\right)^{\alpha-1} du. \tag{8}$$

The gamma distribution admits no closed-form expression for the tail probabilities and thus no simple formula for the hazard function. From (8), however, it follows that $\lambda(\tau)$ is (monotonically) increasing for $\alpha > 1$, decreasing for $\alpha < 1$, and constant (and equal to β) for $\alpha = 1$.

Now, consider the arrival time of the nth event,

$$\vartheta_n = \tau_1 + \cdots + \tau_n, \qquad n = 1, 2, \ldots, \tag{9}$$

where $\{\tau_i\}$ are iid gamma distributed. The reproductive property of the gamma distribution (Johnson and Kotz 1970, p. 170) implies that ϑ_n is gamma distributed with density

$$f_n(\vartheta; \alpha, \beta) = \frac{\beta^{n\alpha}}{\Gamma(n\alpha)}\vartheta^{n\alpha-1}e^{-\beta\vartheta}. \tag{10}$$

To derive the new count-data distribution, we have to evaluate the cumulative distribution function

$$F_n(T) = \frac{1}{\Gamma(n\alpha)}\int_0^{\beta T} u^{n\alpha-1}e^{-u}du, \qquad n = 1, 2, \ldots, \tag{11}$$

where the integral is the incomplete gamma function. The right side will be denoted as $G(\alpha n, \beta T)$. Note that $F_0(T) = 1$. The number of event occurrences during the time interval $(0, T)$ has the two-parameter distribution function

$$P\{N = n\} = G(\alpha n, \beta T) - G(\alpha n + \alpha, \beta T) \tag{12}$$

for $n = 0, 1, 2, \ldots$, where $\alpha, \beta \in \mathbf{R}^+$ and $G(0, \beta T) \equiv 1$.

For α taking integer values, (10) coincides with a distribution known in the statistical literature as the *Erlangian* distribution (Cox 1962, p. 15). Integrating (11) by parts gives

$$G(n, \beta T) = 1 - e^{-\beta T}\left(1 + \beta T + \frac{(\beta T)^2}{2!} + \cdots + \frac{(\beta T)^{\alpha n-1}}{(\alpha n - 1)!}\right). \tag{13}$$

Hence

$$P(N = n) = G(\alpha n, \beta T) - G(\alpha n + \alpha, \beta T) = e^{-\beta T}\sum_{i=0}^{\alpha-1}\frac{(\beta T)^{\alpha n+i}}{(\alpha n + i)!}, \qquad n = 0, 1, 2, \ldots. \tag{14}$$

For $\alpha = 1$, $f(\tau)$ is the exponential density and (14) simplifies to the Poisson distribution. It was noted previously that the Poisson distribution is characterized by independently exponential distributed waiting times.

For noninteger α, no closed-form expression is available for $G(\alpha n, \beta T)$ (and thus for $P\{N = n\}$). Numerical evaluations of the integral can be based on asymptotic expansions (see Abramowitz and Stegun 1964; Bowman and Shenton 1988). Figures 1 and 2 compare the probability functions of the gamma count distribution with a Poisson distribution of identical mean $[E(N) = 2]$ for two values of α. Depending on the value of α, the new model is more concentrated ($\alpha = 1.5$) or more dispersed ($\alpha = .5$) than the reference distribution.

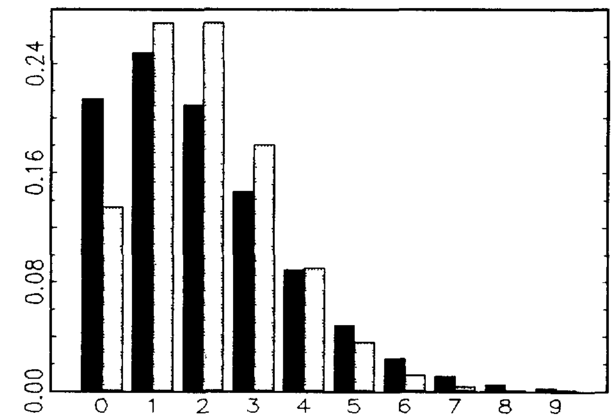

Figure 1. Probability Functions for Gamma Count and Poisson Distributions; α = .5 (overdispersion): Heavily Shaded, Gamma Count; Lightly Shaded, Poisson.

The expected value is given by

$$E(N) = \sum_{i=1}^{\infty} G(\alpha i, \beta T). \tag{15}$$

For increasing T it holds that [see (5)]

$$N(T) \overset{\text{asy}}{\sim} \text{normal}\left(\frac{\beta T}{\alpha}, \frac{\beta T}{\alpha^2}\right). \tag{16}$$

The limiting variance–mean ratio equals a constant $1/\alpha$. It follows that the gamma count-distribution function (12) displays overdispersion for $0 < \alpha < 1$ and underdispersion for $\alpha > 1$. The same holds true for a unit time period in which the moments are evaluated numerically. Figures 3 and 4 show the variance–mean ratio for various values of α and β.

The underling waiting times have a decreasing (increasing) hazard for $0 < \alpha < 1$ ($\alpha > 1$). Thus, negative duration dependence is associated with overdispersion, positive duration dependence with underdispersion. The intermediate case of no duration dependence—that is, exponentially distributed waiting times—leads to the Poisson distribution with equal mean and variance. Note also that the negative binomial distribution permits more general types of overdispersion.

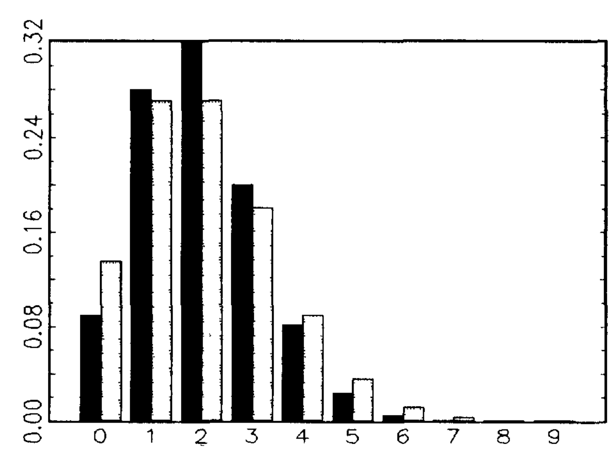

Figure 2. Probability Functions for Gamma Count and Poisson Distributions; α = 1.5 (underdispersion): Heavily Shaded, Gamma Count; Lightly Shaded, Poisson.

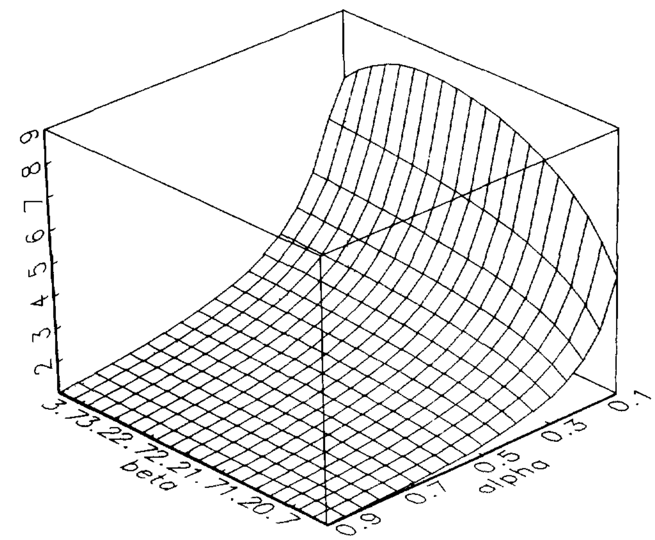

Figure 3. Variance to Mean Ratio for Gamma Count Distribution; $0 < \alpha < 1$.

To obtain a gamma count regression model, the assumption of a homogeneous population is relaxed by formulating a conditional model in which the parameters depend on a vector of individual covariates x_i. Assuming that the period at risk (i.e., the length of the time interval during which event occurrences are counted) is identical for all observations, T may be set to unity without loss of generality as long as an intercept is included. Assume that

$$\frac{\beta}{\alpha} = \exp(x_i'\gamma). \tag{17}$$

This parameterization yields the regression

$$E(\tau_i \mid x_i) = \exp(-x_i'\gamma). \tag{18}$$

The regression is for the waiting times τ_i and not directly for the counts N_i because, unless $\alpha = 1$, it does not hold that $E(N_i \mid x_i) = [E(\tau_i \mid x_i)]^{-1}$. The estimated parameters $\widehat{\gamma}$

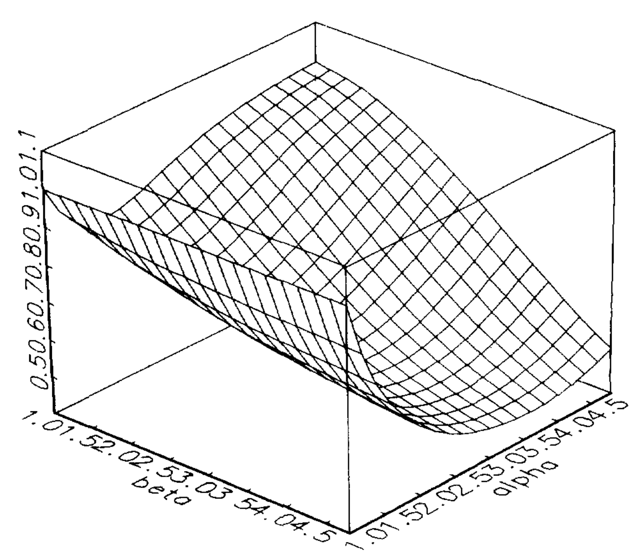

Figure 4. Variance to Mean Ratio for Gamma Count Distribution; $\alpha > 1$.

have to be interpreted accordingly. $-\gamma$ measures the percentage change in the expected waiting time caused by a unit increase in x_i. If x_i is in logarithms, $-\gamma$ is a partial elasticity.

Given a sample of independent observations $(y_1, \ldots, y_n)$, estimates $\widehat{\alpha}$ and $\widehat{\gamma}$ can be obtained by maximizing the log-likelihood ℓ, which is given by

$$\ell(y_i; x_i, \alpha, \gamma) = \sum_{i=1}^{n} \ln\{G(\alpha y_i, \alpha \exp(x_i'\gamma)) - G(\alpha y_i + \alpha, \alpha \exp(x_i'\gamma))\}. \quad (19)$$

Because the function is nonlinear in $\widehat{\alpha}$ and $\widehat{\gamma}$, an iterative solution algorithm is needed. In the following application, the algorithm described by Berndt, Hall, Hall, and Hausman (1974) with numerical forward-differenced gradients was used.

To make the estimated coefficients comparable to those obtained from the negative binomial or Poisson regression models, a simple strategy is to hold all explanatory variables constant at their means, and to compute $\Delta\widehat{Y}/\Delta x$, where x is the remaining explanatory variable and the change is defined by a unit increase at the mean value (in the case of continuous variables) or by a change from 0 to 1 in the case of dummy variables. This measures literally the estimated effect of that explanatory variable on the dependent variable, holding other variables constant.

5. ESTIMATION AND TESTING: TWO ILLUSTRATIONS

In this section, I illustrate the use of the new count-data model in two applications, the number of births by a woman and the number of consultations with a doctor or specialist. Is duration dependence a plausible feature of such processes? For the case of fertility, one can consult a rich empirical literature on the timing and spacing of births. Examples are Newman and McCulloch (1984) and Heckman and Walker (1990).

Heckman and Walker (1990) used a multistate duration model. Correlation across spells was introduced through spell-persistent unobservables (like fecundity differences). Using modern Swedish fertility data, Heckman and Walker found (a) that unobservables correlated across spells are not an important feature of the data and (b) that the waiting times between births display positive duration dependence. This previous finding suggests that the assumptions of the gamma count model might hold in the case of fertility data. Furthermore, the presence of positive duration dependence implies underdispersion on the level of counts that is in fact typically found in such data (e.g., Schultz 1990). I am unaware of similar evidence for the timing of doctor consultations. Assuming that a doctor consultation initiates a treatment and leads to a gradual improvement of the health condition, however, one might conjecture that the risk of further visits decreases with elapsed time since the initial visit, giving rise to negative duration dependence.

I start with a discussion of the fertility example. In economic demography, fertility is commonly measured as a rate that gives the number of births per 1,000 of population. Microlevel studies, in contrast, measure fertility by the number of births per woman. I follow here the second approach and estimate a model for *completed* fertility—that is, the number of births for women past childbearing age or the "children ever born." The "period at risk" is defined as age 15–44 in official statistics. A discussion of the competing theories of fertility choice is beyond the scope of this section [Olsen (1994) is a useful reference]. Behavioral factors are assumed to affect parental goals, which, in turn, determine the risk of a birth. Among the main contributing factors are the women's labor-market opportunities as measured by education level and previous labor-market attachment.

I use data from the second (1985) wave of the *German Socio-Economic Panel*. The sample consists of 1,243 women over 44 in 1985, who are in first marriages and who answered the questions relevant to the analysis. In Table 1, I present the results from a regression for the number of children. The average number of children is 2.4, and the variance is 2.3. Note that equality of mean and variance for the *marginal* distribution is compatible with underdispersion for the *conditional* distribution of the number of children. The explanatory variables include general education (measured as years of schooling); dummies for post-secondary education, either vocational training or university; nationality (German); religious denomination (Catholic, Protestant, and Muslim, with other or none as reference group); year of birth; and age at marriage.

The two estimated models are the Poisson and the gamma models. The negative binomial estimation failed because the

Table 1. Regression Results for Total Marital Fertility

	Poisson		Gamma	
Variable	*Coeff.*	$\Delta\hat{Y}/\Delta x$	*Coeff.*	$\Delta\hat{Y}/\Delta x$
Constant	1.147		1.193	
	(4.054)		(4.805)	
German	−.200	−.490	−.190	−.489
	(−2.943)		(−3.189)	
Years of schooling	.034	080	032	.079
	(1.138)		(1.198)	
Vocational training	−.153	−.349	−.144	−.346
	(−3.932)		(−3.994)	
University	−.155	−.332	− 146	−.331
	(−1.157)		(−1.137)	
Catholic	.218	.514	.206	.511
	(3.596)		(3.527)	
Protestant	.113	.262	.107	.261
	(1.770)		(1.703)	
Muslim	.548	1.627	.523	1.616
	(7.179)		(7.475)	
Rural	.059	.135	.055	.133
	(1.719)		(1.712)	
Year of birth	−.002	−.005	−.002	−.005
	(1.138)		(1.202)	
Age at marriage	−.030	−.069	−.029	−.068
	(−5.124)		(−5.377)	
α			1.439	
			(6.183)	
Log-*L*	−2,101.8		−2,078.2	
Restricted log-*L*	−2,186.8		−2,182.5	
Observations	1,243		1,243	

NOTE Asymptotic *t* values are in parentheses, the data are from the German Socio-Economic Panel Wave, 1985

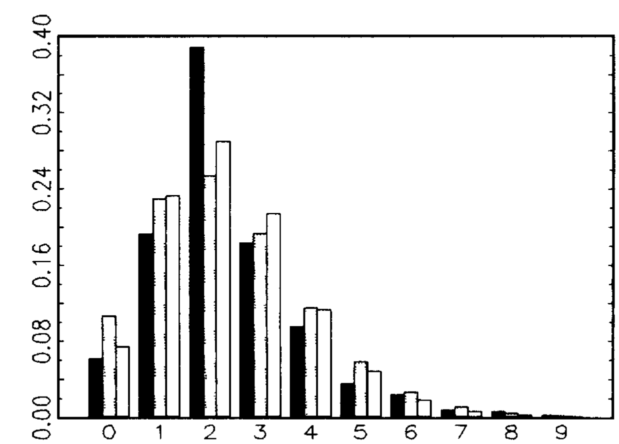

Figure 5. Predicted Relative Frequencies for Number of Children: Heavily Shaded, Sample; Lightly Shaded, Poisson; White, Gamma.

data are *underdispersed*. The gamma model estimates an α of 1.44, which is significantly greater than 1 (H_0: $\alpha \leq 1$; t = 6.183). Moreover, a likelihood ratio test clearly rejects the Poisson specification (-2LR = 47.2) at any conventional level of significance. The presence of underdispersion causes the estimated Poisson standard errors to be upwardly and the t values to be downwardly biased. I conclude that the empirical evidence is compatible with positive duration dependence for the birth process and thus with the Heckman and Walker (1990) finding.

The estimated coefficients are very stable for the two specifications, both in sign and in order of magnitude. The same holds true for the marginal effect $\Delta\widehat{Y}/\Delta x$ evaluated at the sample mean of x. The "cultural" variables have the expected signs and are significantly different from 0: Germans have lower fertility, whereas women with rural background, Catholics, and, in particular, Muslims have higher fertility. A Muslim woman has, on average, 1.6 more children than a women without religious denomination. The effect of education is less straightforward. Although the presence of vocational training (most of which is apprenticeship training) has a significant negative impact on completed fertility, both general education and academic training do not. One interpretation is that vocational training (which is a major source of training in Germany—the sample mean is .43) signals a close attachment to the labor force and thereby measures opportunity costs as well as preferences.

In Figure 5, I plot the sample relative frequencies and the predicted relative frequencies evaluated at the individual X's

Table 2. Regression Results for Number of Doctor Consultations

	Poisson		$NEGBIN_k$		Gamma	
Variable	*Coeff.*	$\Delta\hat{Y}/\Delta x$	*Coeff.*	$\Delta\hat{Y}/\Delta x$	*Coeff.*	$\Delta\hat{Y}/\Delta x$
One	−2.224		−2.172		−8.281	
	(−11.716)		(−9.290)		(−3.634)	
Sex	.157	.035	224	.048	.470	.030
	(2.795)		(3.229)		(2.141)	
Age	1.056	.427	−.379	−.068	3.606	.497
	(1.055)		(−.293)		(1.028)	
Agesq	−.849	−.130	.805	.270	−3.190	−.130
	(−.787)		(564)		(−.870)	
Income	−.205	−.042	−.138	−.028	−.591	−.034
	(−2.323)		(−1.273)		(−1.737)	
Levyplus	.123	.028	.106	.023	.353	.022
	(1.720)		(1.240)		(1.301)	
Freepoor	−.440	−.082	−.495	−.087	−1.396	−.074
	(−2.447)		(−2.430)		(−1 816)	
Freerepa	.080	.018	.141	.032	.215	.014
	(.867)		(1.200)		(.680)	
Illness	.187	.046	.216	.052	.510	.035
	(10.227)		(8.804)		(4.106)	
Actdays	.127	.030	.150	.035	.323	.022
	(25.198)		(15.054)		(4.414)	
Hscore	.030	.006	.039	.008	.068	.004
	(2.979)		(2.751)		(2.048)	
Chcond1	.114	.026	.094	.020	.523	.034
	(1.712)		(1.200)		(1.803)	
Chcond2	.141	.033	.199	.047	.552	.038
	(1.698)		(1.894)		(1.680)	
ϕ; α			1.130		.235	
			(9.960)		(−12.439)*	
k			1.138			
			(10.194)			
Log-L	−3,355.5		−3198.0		−3,261.3	
Restricted log-L	−3,989.2		−3585.9		−3,668.1	
Observations	5,190		5190		5,190	

NOTE Asymptotic t values are in parentheses, *H_0 α = 1, Data are from Australian Health Survey 1977–1978 for the Poisson and the gamma count models. The Poisson model overpredicts tail outcomes and underpredicts middle outcomes. The gamma count model, in contrast, puts more weight in the middle, which leads to an improved fit in the fertility case.

I now turn to the estimates for the number of doctor consultations. The data set is identical to the one used by Cameron and Trivedi (1986). The sample of 5,190 individuals is generated from the Australian Health Survey 1977–1978. The dependent variable is the number of consultations with a doctor or specialist in the two-week period prior to the interview, with mean .302 and variance .637. Eighty percent of the respondents had zero consultations. Further details and a motivation of the selection of explanatory variables, were given by Cameron and Trivedi (1986) and the references that they quoted.

Regressors include demographics (sex, age, age squared), income, various measures of health status [number of reduced activity days (actdays), general health questionnaire score (hscore), recent illness, two types of chronic conditions (chcond1, chcond2)], and three types of health-insurance coverage (levyplus, freepoor, freerepat—the former representing a higher level of coverage and the latter two a basic level supplied free of charge).

Table 2 contains the regression results for the Poisson, the negative binomial, and the gamma models. The negative binomial model is estimated using the Winkelmann and Zimmermann (1995) specification with unrestricted k. The estimated value for k is 1.14. Estimation of the general model allows for a formal model selection between the NEGBIN I (H_0: $k = 0$; $t = 10.194$) and the NEGBIN II (H_0: $k = 1$; $t = 1.236$) models used by Cameron and Trivedi (1986), favoring the latter parameterization.

Both negative binomial and gamma models are superior to the simple Poisson model. The respective parametric restrictions (H_0: $\phi = 0$; $t = 9.960$) and (H_0: $\alpha = 1$; $t = 12.439$) can be rejected in both cases. The same holds true for the likelihood ratio tests with $-2\text{LR}_{\text{NB}} = 315.0$ and $-2\text{LR}_{\text{gam}} = 188.4$. Although it is clear that the rejection of the Poisson model is caused by overdispersion, the causes of this misspecification are not identified. The negative binomial model, which is compatible with both occurrence dependence and unobserved heterogeneity, is superior in this case. For the NEGBIN II, which has the same number of parameters as the gamma model, the Log L is $-3{,}198.8$ (Cameron and Trivedi 1986) versus $-3{,}261.3$ for the gamma model. With the gamma model as maintained hypothesis, the finding would indicate negative duration dependence. Figure 6 gives the predicted probabilities, evaluated at the individual regressors, for the three models. Typically, overdispersion is associated with excess zeros—that is, more zeros in the sample than the Poisson model would predict. Here, zeros are predicted well, whereas the improvement provided by the two alternative models stems from the strictly positive counts.

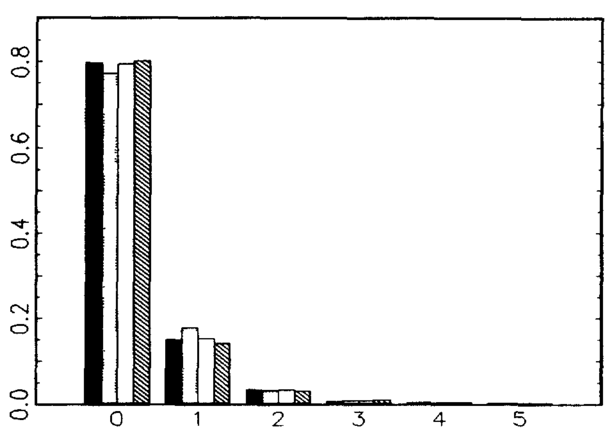

Figure 6. Predicted Relative Frequencies for Number of Doctor Consultations: Heavily Shaded, Sample; Lightly Shaded, Poisson; White, Gamma Count; Diagonally Shaded, Negative Binomial.

6. CONCLUSIONS

Why is the rejection of the Poisson model so universal in econometric applications? Existing explanations have stressed the importance of unobserved heterogeneity and occurrence dependence. This article adds a further explanation—duration dependence for the distribution of the waiting times between event occurrences. In this framework, negative duration dependence causes overdispersion and positive duration dependence underdispersion. I use parametric assumptions, independently and identically gamma-distributed waiting times, to derive a corresponding count-data model. The use of the new model is demonstrated in two applications and yields results compatible with previous research using duration data. One important advantage of the new model is that it removes the artificial asymmetry between overdispersion and underdispersion. Rather, they are two sides of the same coin, a violation of the constant hazard assumption underlying the Poisson model. Additional work remains to be done to extend the model to situations of nonmonotonic and semiparametric hazard functions and to study the effects of unobserved heterogeneity in this framework.

ACKNOWLEDGMENTS

I am grateful to James J. Heckman, Pravin K. Trivedi, and anonymous referees for their valuable comments and suggestions.

[*Received January 1994 Revised January 1995.*]

REFERENCES

Abramowitz, M., and Stegun, I A (1964), *Handbook of Mathematical Functions, Bureau of Standards* (Applied Mathematics Series No. 55), Washington, DC: National Bureau of Standards

Arbous, A. G., and Kerrich, J. C (1951), "Accident Statistics and the Concept of Accident Proneness," *Biometrics*, 7, 340–432

Barlow, R. E., and Proschan, F (1965), *Mathematical Theory of Reliability*, New York: John Wiley

Berndt, E., Hall, B., Hall, R., and Hausman, J. (1974), "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, 3/4, 653–666

Bowman, K O., and Shenton, L R (1988), *Properties of Estimators for the Gamma Distribution*, New York Marcel Dekker

Cameron, A. C., and Trivedi, P. K. (1986), "Economic Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests," *Journal of Applied Econometrics*, 1, 29–53.

Cox, D. R. (1962), *Renewal Theory*, New York: John Wiley

Feller, W. (1943), "On a General Class of Contagious Distributions," *The Annals of Mathematical Statistics*, 14, 389–400

——— (1971), *An Introduction to Probability Theory and Its Applications* (Vol. 2, 2nd ed.), New York: John Wiley.

Gourieroux, C., Monfort, A., and Trognon, A. (1984), "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica*, 52, 701–721

Gourieroux, C., and Visser, M. (1992), "A Count Data Model With Unobserved Heterogeneity," unpublished manuscript.

Gurland, J. (1959), "Some Applications of the Negative Binomial and Other Contagious Distributions," *American Journal of Public Health*, 49, 1388–1399

Heckman, J. J., and Walker, J. R. (1990), "The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence From Swedish Longitudinal Data," *Econometrica*, 58, 1411–1442.

Johnson, N. L., and Kotz, S. (1970), *Continuous Univariate Distributions* (Vol. 1), New York: John Wiley.

Lawless, J. F. (1987), "Regression Methods for Poisson Process Data," *Journal of The American Statistical Association*, 82, 808–815.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman and Hall

Mullahy, J. (1986), "Specification and Testing in Some Modified Count Data Models," *Journal of Econometrics*, 33, 341–365.

Newman, J. L., and McCulloch, C. E. (1984), "A Hazard Rate Approach to the Timing of Births," *Econometrica*, 52, 939–962.

Olsen, R. J. (1994), "Fertility and the Size of the U.S. Labor Force," *Journal of Economic Literature*, 32, 60–100

Schultz, T. P. (1990), "Testing the Neoclassical Model of Family Labor Supply and Fertility," *The Journal of Human Resources*, 25, 599–534

Winkelmann, R. (1994), *Count Data Models—Econometric Theory and an Application to Labor Mobility* (Lecture Notes in Economics and Mathematical Systems Nr. 410), Berlin: Springer-Verlag

Winkelmann, R., and Zimmermann, K. F. (1995), "Recent Developments in Count Data Modeling: Theory and Applications," *Journal of Economic Surveys*, 9, 1–24