The Stata Journal

The *Stata Journal*, electronic version (ISSN 1536-8734) is a publication of Stata Press, and Stata is a registered trademark of StataCorp LP.

The Stata Journal (2006)
6, Number 2, pp. 246–255

Maximum simulated likelihood estimation of a negative binomial regression model with multinomial endogenous treatment

Partha Deb
Hunter College, City University of New York
New York, NY
partha.deb@hunter.cuny.edu

Pravin K. Trivedi
Indiana University
Bloomington, IN
trivedi@indiana.edu

Abstract. We describe specification and estimation of a multinomial treatment effects negative binomial regression model. A latent factor structure is used to accommodate selection into treatment, and a simulated likelihood method is used for estimation. We describe its implementation via the `mtreatnb` command.

Keywords: st0105, mtreatnb, multinomial treatment effects, latent factors, count data, negative binomial, multinomial logit, multinomial logistic, Halton sequences, maximum simulated likelihood

1 Introduction

We develop a treatment-effects model that can be used to analyze the effects of an endogenous multinomial treatment (when exactly one treatment is chosen from a set of more than two choices) on a nonnegative integer–valued outcome. Although econometric models for count data are well developed and have many uses (Cameron and Trivedi 1998), there are few extensions of such models to accommodate endogenous regressors. Extensions of treatment-effects models to multinomial treatment indicators are also not well developed for nonlinear models in general. We specify the model with a latent factor structure that allows for idiosyncratic influences on treatment choice to affect outcomes, thus enabling us to make a distinction between selection on unobservables and selection on observables. The multinomial treatment variable is assumed to have a multinomial logit structure, and the outcome is assumed to follow a negative binomial distribution conditional on treatment. We use a negative binomial distribution to accommodate overdispersion, which is a typical feature of count outcomes.

In this context, introducing latent factors into the equations for treatment and outcome has two main advantages over other ways of generating correlated errors. First, the appropriately normalized latent factors have a natural interpretation as proxies for unobserved covariates since they enter into the equations in the same way as observed covariates, and the associated factor loadings can be interpreted in much the same way as coefficients on observed covariates can. Second, the latent factors can be used generally to combine conditional and marginal distributions to generate joint distributions despite the fact that these joint distributions typically do not have a closed-form representation. Thus if you wanted to use distributions other than the multinomial logit

and/or the negative binomial, the same structure and principles would apply. A potential disadvantage is that correlation between endogenous variables induced by latent factors is subject to an upper bound less than one.

We (forthcoming) develop such a model in which the multinomial treatment variable arises from the choice of type of health insurance plan and the outcome measures medical care usage, e.g., number of visits to the doctor and number of hospital stays. We (2005) and Deb et al. (forthcoming) describe other applications of this model. Lee (1983) proposed a two-step method for a model with multinomial treatment and outcome with exponential mean, which is an alternative, computationally simpler approach, but at the cost of inefficiency. Also, whereas the latent factor approach can easily be adapted to models with alternative statistical structures for treatment and outcome, Lee's approach requires that the formulas be worked out case by case.

This paper is organized as follows. In section 2, we describe the simultaneous-equations model and the estimation methods. Section 3 describes the syntax for the command `mtreatnb`. We describe an empirical example in section 4 and provide some computational guidance based on our experience in section 5.

2 Methods

2.1 Model specification

Each individual i chooses one treatment from a set of three or more choices, which typically includes a control group, implying a multinomial choice model. Let EV^*_{ij} denote the indirect utility that we would obtain by selecting the jth treatment, $j = 0, 1, 2, \ldots, J$ and

$$\text{EV}^*_{ij} = \mathbf{z}_i'\alpha_j + \delta_j l_{ij} + \eta_{ij}$$

where $\mathbf{z}_i$ denotes exogenous covariates with associated parameters α_j and η_{ij}, which are independently and identically distributed error terms. Also EV^*_{ij} includes a latent factor l_{ij} that incorporates unobserved characteristics common to individual i's treatment choice and outcome. The l_{ij} are assumed to be independent of η_{ij}. Without loss of generality, let $j = 0$ denote the control group and $\text{EV}^*_{i0} = 0$.

Let d_j be binary variables representing the observed treatment choice and $\mathbf{d}_i = (d_{i1}, d_{i2}, \ldots, d_{iJ})$. Also let $\mathbf{l}_i = (l_{i1}, l_{i2}, \ldots, l_{iJ})$. Then the probability of treatment can be represented as

$$\Pr(\mathbf{d}_i|\mathbf{z}_i, \mathbf{l}_i) = \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 l_{i1}, \mathbf{z}_i'\alpha_2 + \delta_2 l_{i2}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J l_{iJ})$$

where $\mathbf{g}$ is an appropriate multinomial probability distribution. Specifically, we assume that $\mathbf{g}$ has a mixed multinomial logit (MMNL) structure, defined as

$$\Pr(\mathbf{d}_i|\mathbf{z}_i, \mathbf{l}_i) = \frac{\exp(\mathbf{z}_i'\alpha_j + \delta_j l_{ij})}{1 + \sum_{k=1}^{J} \exp(\mathbf{z}_i'\alpha_k + \delta_k l_{ik})}$$

The outcome is a count variable; i.e., $y_i = 0, 1, 2, \ldots$. The expected outcome equation for individual i, $i = 1, \ldots, N$, is formulated as

$$E(y_i|\mathbf{d}_i, \mathbf{x}_i, \mathbf{l}_i) = \mathbf{x}_i'\beta + \sum_{j=1}^{J} \gamma_j d_{ij} + \sum_{j=1}^{J} \lambda_j l_{ij}$$

where $\mathbf{x}_i$ is a set of exogenous covariates with associated parameter vectors β and γ_j denoting the treatment effects relative to the control. $E(y_i|\mathbf{d}_i, \mathbf{x}_i, \mathbf{l}_i)$ is a function of each of the latent factors l_{ij}; i.e., the outcome is affected by unobserved characteristics that also affect selection into treatment. When λ_j, the factor-loading parameter, is positive (negative), treatment and outcome are positively (negatively) correlated through unobserved characteristics; i.e., there is positive (negative) selection, with γ and λ the associated parameter vectors, respectively.

We assume that f is the negative binomial-2 density,

$$f(y_i|\mathbf{d}_i, \mathbf{x}_i, \mathbf{l}_i) = \frac{\Gamma(y_i + \psi)}{\Gamma(\psi)\Gamma(y_i + 1)} \left(\frac{\psi}{\mu_i + \psi}\right)^{\psi} \left(\frac{\mu_i}{\mu_i + \psi}\right)^{y_i}$$

where $\mu_i = E(y_i|\mathbf{d}_i, \mathbf{x}_i, \mathbf{l}_i) = \exp(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \mathbf{l}_i'\lambda)$ and $\psi \equiv 1/\alpha$ $(\alpha > 0)$ is the overdispersion parameter.

As in the standard multinomial logit model, the parameters in the MMNL are identified only up to a scale. Therefore, a normalization for the scale of the latent factors without loss of generality is required. We assume $\delta_j = 1$ for each j but allow the user to change this constant in `mtreatnb`. Also, although the model is identified when $\mathbf{z}_i = \mathbf{x}_i$, including some variables in $\mathbf{z}_i$ that are not included in $\mathbf{x}_i$ is usually preferable; i.e., identification via exclusion restrictions is the preferred approach.

2.2 Estimation

The joint distribution of treatment and outcome variables, conditional on the common latent factors, can be written as the product of the marginal density of treatment and the conditional density of

$$\begin{aligned}\Pr(y_i, \mathbf{d}_i|\mathbf{x}_i, \mathbf{z}_i, \mathbf{l}_i) &= f(y_i|\mathbf{d}_i, \mathbf{x}_i, \mathbf{l}_i) \times \Pr(\mathbf{d}_i|\mathbf{z}_i, \mathbf{l}_i) \\ &= f(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \mathbf{l}_i'\lambda) \times \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 l_{i1}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J l_{iJ})\end{aligned}$$

The problem in estimation arises because the l_{ij} are unknown. We assume that the l_{ij} are independently and identically distributed draws from the standard normal distribution so their joint distribution $\mathbf{h}$ can be integrated out of the joint density; i.e.,

$$\begin{aligned}&\Pr(y_i, \mathbf{d}_i|\mathbf{x}_i, \mathbf{z}_i) = \\ &\quad \int \left\{ f(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \mathbf{l}_i'\lambda) \times \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 l_{i1}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J l_{iJ}) \right\} \mathbf{h}(\mathbf{l}_i)\, d\mathbf{l}_i\end{aligned} \tag{1}$$

The main computational problem, given suitable specifications for $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}_j$, is that the integral (1) does not have, in general, a closed-form solution. But this difficulty can be addressed by using simulation-based estimation (Gouriéroux and Monfont 1996) by noting that

$$\begin{aligned}\Pr(y_i, \mathbf{d}_i|\mathbf{x}_i, \mathbf{z}_i) &= \mathrm{E}\big\{f(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \mathbf{l}_i'\lambda) \times \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 l_{i1}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J l_{iJ})\big\} \\ &\approx \frac{1}{S}\sum_{s=1}^{S}\Big\{f(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \tilde{\mathbf{l}}_{is}'\lambda) \times \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 \tilde{l}_{i1s}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J \tilde{l}_{iJs})\Big\}\end{aligned}$$

where $\tilde{\mathbf{l}}_{is}$ is the sth draw (from a total of S draws) of a pseudorandom number from the density $\mathbf{h}$. The simulated log-likelihood function for the data is given by

$$\begin{aligned}\ln l(y_i, \mathbf{d}_i|\mathbf{x}_i, \mathbf{z}_i) \approx \sum_{i=1}^{N}\ln\Bigg[\frac{1}{S}\sum_{s=1}^{S}\Big\{f(\mathbf{x}_i'\beta + \mathbf{d}_i'\gamma + \tilde{\mathbf{l}}_{is}'\lambda)\times \\ \mathbf{g}(\mathbf{z}_i'\alpha_1 + \delta_1 \tilde{l}_{i1s}, \ldots, \mathbf{z}_i'\alpha_J + \delta_J \tilde{l}_{iJs})\Big\}\Bigg]\end{aligned}$$

Provided that S is sufficiently large, maximization of the simulated log likelihood is equivalent to maximizing the log likelihood.

We have found that standard simulation methods produce extremely slow convergence of the simulated likelihood function. Therefore, we adapt an acceleration technique that uses quasirandom draws based on Halton sequences described in Bhat (2001) and Train (2003). Gates (2006) describes an implementation in Mata. Halton sequences have two desirable properties vis-à-vis pseudorandom draws. First, they are designed to give more even coverage over the domain of the mixing distribution. Second, the simulated probabilities are negatively correlated over observations. This negative correlation reduces the variance in the simulated likelihood function. Under suitable regularity conditions, the integration error using pseudorandom sequences is in the order of N^{-1} compared with pseudorandom sequences where the convergence rate is $N^{-1/2}$ (Bhat 2001). A discussion of our experience with different choices of S is given in section 5.

We recommend that the covariance of the MSL estimates be obtained by using the robust *sandwich* formula; i.e., we recommend using the `robust` option. The sandwich formula appropriately accounts for uncertainty due to simulation chatter for finite S (McFadden and Train 2000). Hessian and outer-product formulas for the covariance are only asymptotically (in S) appropriate.

(*Continued on next page*)

3 The mtreatnb command

3.1 Syntax

`mtreatnb` *depvar* [*indepvars*] [*if*] [*in*] [*weight*], `mtreatment(`*depvar_mt indepvars_mt*`)` `simulationdraws(`#`)` [`basecategory(`# or *string*`)` `prefix(`*string*`)` `robust` `cluster(`*varname*`)` `scale(`#`)` `startpoint(`#`)` `altfactors(`*string*`)` `altstart(`*string*`)` *maximize_options* `verbose`]

`fweights`, `pweights`, `iweights`, and `aweights` are allowed.

3.2 Options

`mtreatment(`*depvar_mt indepvars_mt*`)` specifies the variables for the multinomial treatment equation. *depvar_mt* must have more than two and less than 10 categories. This option is an integral part of specifying the treatment-effects model and is required.

`simulationdraws(`#`)` specifies the number of simulation draws per observation and is required. These draws are based on Halton sequences.

`basecategory(`# or *string*`)` is the value or label of *depvar* that will be the base category in the multinomial treatment equations.

`prefix(`*string*`)` lets you choose a prefix other than `_I` for the indicator variables created from the multinomial treatment variable. The default is a set of indicator variables starting with `_I`. When you use `mtreatnb`, it drops all previously created indicator variables starting with the prefix specified in the `prefix()` option or with `_I` by default.

`robust` uses the robust or sandwich estimator of variance. The default is the traditional calculation based on the information matrix.

`cluster(`*varname*`)` adjusts standard errors for intragroup correlation.

`scale(`#`)` lets you choose the standard deviation of the normally distributed quasirandom variables. The default is `scale(1)`.

`startpoint(`#`)` lets you choose the starting point in the Halton sequence from which the quasirandom variates are generated. The default is `startpoint(20)`.

`altfactors(`*string*`)` lets you choose the starting values for the parameters associated with the latent factors. Specify these values as comma-separated numbers. The default starting values are zeros.

`altstart(`*string*`)` lets you choose the starting values for all parameters. Specify these values as comma-separated numbers.

maximize_options control the maximization process. Because latent class models have complicated likelihood functions, `difficult` may be a useful option if the default

setup is unsatisfactory. The other *maximize_options* are seldom used. `altfactors` and `altstart` may also be useful to generate alternative starting values if the default setup is unsatisfactory.

`verbose` lets you display iteration logs and estimates tables for the mixed multinomial logit and negative binomial regressions, which are estimated to create starting values. The default is that this output is not displayed.

4 Example

To illustrate the method, we use data from the 2001 Medical Expenditure Panel Survey, a representative survey of the noninstitutionalized population in the United States with wide scope and excellent information on demographic characteristics, health status, employment status, and earnings, and a wide variety of measures of health care usage. Our sample consists of 5,033 persons who are aged between 25 and 59 years, privately insured, and employed but not self-employed. The outcome variable is the number of doctor visits in a year (`docvis`) and the multinomial treatment variable describes the type of health insurance plan (`instype`) and takes three values: fee-for-service (`ffs`)—the control, health maintenance organizations (`hmo`), and other managed-care (`omc`) organizations. Exogenous covariates include age, gender, race, education, and health status. Characteristics of the person's employer (`firmsize`, `govtjob`) enter only the treatment equations and serve as exclusion restrictions or instruments. Maximum simulated likelihood estimates using `mtreatnb` with $S = 400$ are given below.

(*Continued on next page*)

```
. use http://urban.hunter.cuny.edu/~deb/Stata/mepssmall.dta
. mtreatnb docvis age female minority education nchroniccond,
> mtreat(instype age female minority education nchroniccond firmsize govtjob)
> sim(400) basecat(ffs) robust
Fitting mixed multinomial logit regression for treatments:

Fitting negative binomial regression for outcome:

Fitting full model for treatments and outcome:
Iteration 0:   log pseudolikelihood = -17454.622  (not concave)
  (output omitted)
Iteration 14:  log pseudolikelihood =  -17235.09
Multinomial treatment-effects NB regression       Number of obs   =       5033
                                                  Wald chi2(21)   =    1868.79
Log pseudolikelihood =  -17235.09                 Prob > chi2     =     0.0000
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
omc						
age	-.0108416	.0493525	-0.22	0.826	-.1075707	.0858875
female	.0240614	.0900778	0.27	0.789	-.1524879	.2006107
minority	.0889349	.1052808	0.84	0.398	-.1174117	.2952816
education	.0345643	.0183949	1.88	0.060	-.001489	.0706176
nchroniccond	.0557137	.0581656	0.96	0.338	-.0582888	.1697161
firmsize	.0081583	.0023237	3.51	0.000	.0036041	.0127126
govtjob	-.4257763	.1176419	-3.62	0.000	-.6563502	-.1952024
_cons	-1.460597	.3354336	-4.35	0.000	-2.118035	-.8031592
hmo						
age	-.0614867	.0409451	-1.50	0.133	-.1417376	.0187642
female	.1171071	.0748347	1.56	0.118	-.0295662	.2637805
minority	.4006916	.0854049	4.69	0.000	.2333011	.5680822
education	.0123691	.0142317	0.87	0.385	-.0155244	.0402626
nchroniccond	-.0497685	.0518889	-0.96	0.337	-.1514689	.0519319
firmsize	.0087538	.0019555	4.48	0.000	.004921	.0125865
govtjob	-.0055707	.0927711	-0.06	0.952	-.1873987	.1762573
_cons	-.0542389	.266907	-0.20	0.839	-.577367	.4688893
docvis						
_Iomc	.9623102	.0677789	14.20	0.000	.8294659	1.095154
_Ihmo	.4865105	.0672454	7.23	0.000	.354712	.618309
age	.1487909	.0212606	7.00	0.000	.1071208	.190461
female	.6417384	.0401551	15.98	0.000	.563036	.7204409
minority	-.3204898	.0467591	-6.85	0.000	-.4121359	-.2288436
education	.0429977	.0077763	5.53	0.000	.0277564	.0582391
nchroniccond	.6808642	.024907	27.34	0.000	.6320474	.7296809
_cons	-1.409899	.154162	-9.15	0.000	-1.712051	-1.107747
/lnalpha	-1.547348	.1941742	-7.97	0.000	-1.927923	-1.166774
/lambda_omc	-1.012026	.0433514	-23.34	0.000	-1.096993	-.927059
/lambda_hmo	-.4127126	.0632442	-6.53	0.000	-.536669	-.2887562
alpha	.2128115	.0413225			.14545	.3113698

```
Notes:
1. ffs is the base outcome
2. 400 Halton sequence-based quasirandom draws per observation
```

As in the multinomial logit model, there is one equation for each treatment relative to the control (base). The results show that `firmsize`, which is one of the variables excluded from the outcome equation, is significant in both treatment equations. Individuals who work in larger firms are more likely to choose OMC plans than FFS plans and HMO plans relative to FFS plans. Also individuals who work in government jobs are significantly less likely to choose OMC plans than FFS plans. On the other hand, health status is not a significant determinant of plan choice. The parameter estimates of the outcome equation can be interpreted in the same fashion as those from a negative binomial regression. The results show significant treatment effects (as read off the coefficients on `_Iomc` and `_Ihmo`). Because the conditional mean for the outcome is exponential, parameter estimates can be interpreted directly in percent changes in the mean outcome. Therefore, individuals in OMC plans have 96% more visits than those in FFS plans, whereas those in HMOs have 49% more visits than those in FFS plans. There is also significant evidence of selection on unobservables. The coefficients on the latent factors, `/lambda_omc` and `/lambda_hmo`, are both negative, suggesting that individuals who are more likely to choose either type of managed-care plan relative to FFS, on the basis of their unobserved characteristics, visit the doctor less often. Other individual characteristics are also statistically significant in the outcome equation.

It is also often useful to construct the likelihood-ratio test for exogeneity of treatment, which is a test for the joint hypothesis that the λs are equal to zero; i.e., `/lambda_omc` $= 0$ and `/lambda_hmo` $= 0$. The constrained log likelihood can be calculated as the sum of the log-likelihood values of the MMNL and the negative binomial regressions. This log likelihood, and that of the unconstrained model, is available via `ereturn`. Stata's `lrtest` is not appropriate for the likelihood-ratio test. In general, the likelihood-ratio statistic for exogeneity follows a $\chi^2(q)$ distribution, where q is the number of λ parameters or, equivalently, the number of treatment equations. In our application, $q = 2$.

```
. *** likelihood-ratio test of exogeneity ***
. scalar LR = 2*(e(ll) - e(ll_exog))
. scalar p = 1-chi2(2,LR)
. display LR
439.06408
. display p
0
```

The result shows that the null hypothesis of exogeneity is overwhelmingly rejected.

5 Using mtreatnb: More remarks

Estimation using MSL is computationally intensive. Iteration times in `mtreatnb` increase linearly in the number of simulation draws and multiplicatively in the number of equations. The algorithm is relatively insensitive to the number of covariates, assuming that they are numerically well behaved. Estimation of the model described above took 22 minutes on a PC with an Intel Pentium 1.8-GHz processor and 512 MB of RAM.

The literature indicates that S should increase faster than $\sqrt{N}$, but this assertion does not give explicit guidance in choosing S. In practice, a small number of draws (often 50–100) works well for models such as the mixed multinomial logit and multinomial probit. However, our experience with models with endogenous regressors is that many more draws are required for similar precision (typically one order of magnitude more draws). Except for syntax checking and perhaps preliminary trial runs, we recommend using as large an S as appears computationally reasonable.

In principle, the parameters of the model are identified even if the regressors in the treatment equations are identical to those used in the outcome equation. In practice, however, we recommend using exclusion restrictions; i.e., include regressors in the treatment equations that do not enter the outcome equation.

In the default `mtreatnb` setup, an MMNL model and a negative binomial regression (using `nbreg`), which assumes exogenous treatment, are fitted first. Parameter estimates from these models along with zeros for the λ parameters are used as starting values. As a by-product, the sum of the two maximized log likelihoods is available via `ereturn`. Also, although we have found this setup to be generally reliable, users may sometimes wish to specify their own starting values for λs or for the entire parameter vector. When the user specifies the entire parameter vector, the preliminary models are not fitted and `ereturn` produces a missing value for the log likelihood under exogeneity.

Finally, although the standard deviation of the distribution of each of the latent factors is set equal to one by default for a scale normalization, there may be occasions when it may need to be decreased using the `scale` option. Such situations will occur when the overdispersion parameter α is very close to zero, i.e., when there is no overdispersion conditional on the latent factors. In such situations, the maximization algorithm may take a long time to converge. Because overdispersion in the unconditional negative binomial process is a function of α, the dispersion of the distributions of the latent factors and the magnitudes of the factor-loading parameters $|\lambda|$, decreasing the dispersion of the distributions of the latent factors will typically resolve the issue. A corollary of this issue is that `mtreatnb` may not work well if the outcome variable is not significantly overdispersed.

6 Acknowledgments

We thank David Drukker for invaluable help with the code for `mtreatnb` and for many suggestions to improve the paper. We also thank Colin Cameron and an anonymous referee for suggestions to improve the paper.

7 References

Bhat, C. R. 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research, Part B* 35: 677–693.

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

Deb, P., C. Li, P. K. Trivedi, and D. Zimmer. Forthcoming. The effect of managed care on use of health care services: Results from two contemporaneous household surveys. *Health Economics*.

Deb, P., and P. K. Trivedi. 2005. Provider networks and primary care signups: Do they restrict the use of medical services? Hunter College, Working Paper.

———. Forthcoming. Specification and simulated likelihood estimation of a non-normal treatment-outcome model with selection: Application to health care utilization. *Econometrics Journal*.

Gates, R. 2006. A Mata Geweke–Hajivassiliou–Keane multivariate normal simulator. *Stata Journal* 6: 190–213.

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

Lee, L.-F. 1983. Generalized econometric models with selectivity. *Econometrica* 51: 507–512.

McFadden, D., and K. Train. 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15: 447–470.

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

About the authors

Partha Deb is associate professor of economics at Hunter College and the Graduate Center, City University of New York.

Pravin K. Trivedi is Rudy Professor of Economics at Indiana University. He is the coauthor of *Regression Analysis of Count Data* and *Microeconometrics: Methods and Applications* (both written with A. Colin Cameron).