DISCUSSION PAPER SERIES

No. 7463

MULTIVARIATE SARMANOV COUNT DATA MODELS

Eugenio J. Miravete

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP7463.asp

ISSN 0265-8003

MULTIVARIATE SARMANOV COUNT DATA MODELS

Eugenio J. Miravete, University of Texas, Austin and CEPR

Discussion Paper No. 7463
September 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Eugenio J. Miravete

CEPR Discussion Paper No. 7463

September 2009

ABSTRACT

Multivariate Sarmanov Count Data Models

I present two flexible models of multivariate, count data regression that make use of the Sarmanov family of distributions. This approach overcomes several existing difficulties to extend Poisson regressions to the multivariate case, namely: i) it is able to account for both over and underdispersion, ii) it allows for correlations of any sign among the counts, iii) correlation and dispersion depend on different parameters, and iv) constrained maximum likelihood estimation is computationally feasible. Models can be extended beyond the bivariate case. I estimate the bivariate versions of two of these models to address whether the pricing strategies of competing duopolists in the early U.S. cellular telephone industry can be considered strategic complements or substitutes. I show that a Sarmanov model with double Poisson marginals outperforms the alternative count data model based on a multivariate renewal process with gamma distributed arrival times because the latter imposes very restrictive constraints on the valid range of the correlation coefficients.

JEL Classification: C16, C35 and L11
Keywords: double Poisson, gamma, multivariate count data models, Sarmanov distributions and tariff options

Eugenio J. Miravete
Department of Economics
University of Texas at Austin
1 University Station C3100, BRB
2.126
Austin, TX 78712
USA

Email: miravete@eco.utexas.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=139116

Submitted 09 September 2009

Thanks are due to Jason Abrevaya, Stephen Donald, and Jerry Hausman for their comments. This paper was presented at the Texas Camp Econometrics XIV.

1 Introduction

The absence of a sufficiently flexible multivariate distribution of counts has so far prevented the full information estimation of multivariate count data models. Simple analysis such as seemingly unrelated regressions, simultaneous equations models, or models with endogenous regressors remain mostly intractable whenever several endogenous count data variables are involved.[1] In the absence of such flexible joint distribution of counts, applied econometricians have turned their attention to computing intensive models such as orthogonal polynomial series expansions, *e.g.*, Cameron and Trivedi (1998, §8.5), or moment-based estimation methods, *e.g.*, Gourieroux, Monfort, and Trognon (1984). However, even the best models available are difficult to extend beyond the bivariate case, neither they consider any effect of unobserved heterogeneity other than the well known overdispersion of the distribution of counts. Furthermore, when correlation of counts is considered, it is generally assumed to depend on the same parameter identifying the overdispersion effects, therefore restricting correlations among counts to be necessarily positive.

This paper makes use of multivariate Sarmanov distributions with flexible marginals to build two multivariate count data regression models. The assumed marginal distributions are flexible in the sense that they can accommodate both over and underdispersion independently of any correlation pattern among the counts. I build these models using the distributions of the two existing unidimensional count data models that can accommodate both over and underdispersion: the double Poisson and the gamma count regression models. I also show that only the double Poisson-Sarmanov model can successfully accommodate multivariate counts with non-negligible correlation as the gamma-Sarmanov reduces the viable correlation to a singleton as soon as one single observation includes a zero count.

It has long been recognized that the Poisson model is generally too restrictive when estimating univariate count data regressions. Implicit to the Poisson model is the assumption of equidispersion of the distribution of counts, which is customarily rejected in the data. If the mean does not equal the variance of the distribution of counts estimates of a Poisson regression are still consistent but inference is no longer robust, *e.g.*, Cameron and Trivedi (1998, 3.1-3.4). Many models, such as the Negative Binomial regression, have been suggested to address the existence of unobserved heterogeneity in the data that could explain the commonly observed overdispersion but not the less frequent underdispersion of the distribution of counts. Hausman, Hall, and Griliches (1984) even deal successfully with overdispersion in a univariate panel count data model. However only Efron (1986) and Winkelmann (1995) suggest flexible enough models to address both overdispersion and underdispersion of the distribution of counts. In particular, Efron (1986) introduces the double Poisson distribution while Winkelmann (1995) suggests a model based

[1] There has been some success in dealing with one count and a continuous variable, *e.g.*, Windmeijer and Santos-Silva (1997), or with a count and a dichotomous variable, Hausman, Leonard, and McFadden (1995).

on a gamma distribution of waiting times that exploits the one-to-one relationship between the properties of the hazard rate of the distribution of waiting times and the over and underdispersion of the distribution of events that take place within a given time interval. The two Sarmanov models of this paper assume either double Poisson marginal distributions of the counts or gamma marginal distributions of the waiting time underneath a multivariate renewal process.

Dealing with over and underdispersion is the main goal of most developments in the single-equation count data regression models. Moving towards the multivariate case, the difficulty consists in building a framework that allows for the most flexible correlation pattern possible among counts. Cameron and Trivedi (1998, §8) best summarize the difficulties of estimating a multivariate count data regression model. Kocherlakota and Kocherlakota (1993) present what is perhaps the best known approach to deal with multiple and potentially correlated counts. They derive a bivariate Poisson distribution resulting from the addition of a common Poisson component to two independently distributed Poisson variables. The advantage of this approach —known as trivariate reduction— is that it allows for Poisson marginal distributions, but it also includes a necessarily positive correlation coefficient with severely restricted range that characterizes the dependence structure of the count variables. Similarly, Marshall and Olkin (1990) generate a multivariate count data distribution from mixtures and convolutions of distributions of count events. The advantage of this second approach is that it allows for the simultaneous existence of unobserved heterogeneity that leads only to overdispersion and positive correlation of counts. Still, *a serious limitation of all these models* is that correlation among counts are necessarily positive because there is a single source of heterogeneity that explains *simultaneously* the overdispersion of the marginal distributions of counts and their correlation.[2] Finally, the method of moment approach of Gourieroux et al. (1984) does suffer from the same shortcomings and in addition it does not offer a joint density function that allows econometricians to integrate a conditional distribution in order to deal with endogeneity and simultaneity of count dependent variables.

The Sarmanov count data regression models of the present paper have some remarkable features that overcome most difficulties in extending the existing single-dimensional count data regression models to multivariate environments. First, it can accommodate both over and underdispersion of the distribution of counts, therefore addressing a far larger pattern of behavior induced by the existence of unobserved individual heterogeneity. Second, any correlation sign is allowed, including the possibility of negative correlation among counts, thus reducing the possibility of misspecification.[3] Third, dispersion and correlation

[2] The multivariate Poisson-gamma mixture model of the random effects model of Hausman et al. (1984, §3) is a restricted version of the model of Marshall and Olkin (1990), while Gurmu and Elder (2000) and Winkelmann (2000) suggest multivariate negative binomial models. In all these works, only overdispersion is allowed but in the latter two cases, correlation is independent from dispersion, although necessarily positive.

[3] There are other models that allow for correlation among counts of any sign based on the bivariate Poisson-lognormal distribution of Aitchison and Ho (1989) such as those of Hellström (2006), Munkin and Trivedi (1999), or Riphahn, Wambach, and

of counts depend on different parameters of the model. This is an important property because it adds flexibility to the model by separating the effect of unobserved individual heterogeneity and correlation among counts (again reducing the likelihood of misspecification). Fourth, the model can easily be extended beyond the bivariate case and even serve as a basis for the specification of multivariate count, panel data models. And finally, the estimation is not particularly time consuming: the likelihood function can always be written in closed form and there is no need to use simulation methods to obtain the parameter estimates.

The only drawback of the Sarmanov count data regression model is that in order to have a properly defined multivariate distribution of counts, the range of the estimates of correlations may be smaller than $[-1, 1]$. Indeed, the range of correlation is effectively bounded by the value of the rest of the parameters of the model. The estimation thus requires the use of constrained maximum likelihood methods. In order to deal with the possibility that some of the parameters are on the boundary of these constraints I make use of rescaled bootstrapping to obtain robust confidence intervals.

Joint demand of countable products or services arise in many individual decision problems such as medical service, *e.g.*, Munkin and Trivedi (1999) or Riphahn et al. (2003); job changes, Jung and Winkelmann (1993); types of food, Meghir and Robin (1992); and recreational trips, Hausman et al. (1995), Hellström (2006), or Terza and Wilson (1990). In this paper I study the pricing strategies of competing duopolists in the early U.S. cellular telephone industry. Back in the 1980s, as it is still today common in many industries, firms implemented nonlinear tariffs by means of a menu of self-selecting tariff options. In the absence of strategic considerations, firms will offer the number of tariff plans that is optimal to screen a customer base with some degree of heterogeneity while compensating for the costs of design and commercialization. In such a situation, estimating two independent count data regression models would be appropriate. However, offering numerous rather than few tariff options might carry some strategic value and thus competitors may respond by offering a similar number of tariff plans in order to match the strategy of competitors. The significant positive estimates of the correlation among the count number of tariff plans offered supports the view that the number of tariff options in the early U.S. cellular industry are strategic complements. Alternatively, a negative correlation would arise if firms use the number of tariff plans as a device to segment markets and differentiate themselves from each other, but results do not favor such interpretation.

The paper is organized as follows. Section 2 describes in detail the properties of the Sarmanov family of bivariate distributions. Section 3 presents the bivariate count data regression model based on a bivariate Sarmanov distribution with double Poisson marginals. Section 4 builds a regression model using the bivariate counting process generated by a bivariate gamma-Sarmanov distribution of waiting

Million (2003). These models can only address the case of overdispersed counts while estimation has to resort to simulation methods as the Poisson-lognormal mixture does not have a closed form expression. Gurmu and Elder (2008) obtain a closed form expression only after considering a first order Laguerre polynomial approximation to the bivariate distribution of unobservables. The Sarmanov regression models presented in this paper can address both over and underdispersion and are based on a well defined family of multivariate distributions.

times. Section 5 discusses the multivariate extensions of these models based on the generalization of the Sarmanov distribution proposed by Lee (1996, §8). I then show that in the case of the gamma-Sarmanov model, the likelihood of the multivariate count data regression model can be written as sums of products of single-dimensional integrals. This section presents the trivariate case in detail while the full multivariate generalization of these models is left to Appendix A. Section 6 briefly discusses the estimation of these models while details are left to Appendix B. Appendix B also evaluates the robustness of the double Poisson-Sarmanov estimates with respect to the chosen accuracy of the approximation to its distribution function, as well as the effect of re-scaling the endogenous variable in the case of the gamma-Sarmanov model. Section 7 estimates the bivariate versions of these models to analyze the number of tariff plans offered by competing cellular telephone carriers in the U.S. during the mid-1980s. Finally Section 8 concludes.

2 The Sarmanov Family of Distributions

Let y_k, $k=1,2$ denote two random variables with univariate probability density function $f_k(y_k)$. In general, these marginal distributions have support on $A_k \subseteqq \mathbb{R}$ and have mean and variance:[4]

$$\mu_k = \int_{-\infty}^{\infty} y_k f_k(y_k) dy_k \qquad \text{and} \qquad \sigma_k^2 = \int_{-\infty}^{\infty} (y_k - \mu_k)^2 f_k(y_k) dy_k \,. \tag{1}$$

The Sarmanov count data regression models of this paper make use of particular specifications of the family of bivariate distributions introduced by Sarmanov (1966). This bivariate probability density function takes the following general form:

$$f_{12}(y_1, y_2) = f_1(y_1) f_2(y_2) \times [1 + \omega_{12} \psi_1(y_1) \psi_2(y_2)] \;, \tag{2}$$

where the mixing functions $\psi_k(y_k)$, $k = 1,2$ are bounded and nonconstant functions such as:

$$\int_{-\infty}^{\infty} \psi_k(y_k) f_k(y_k) dy_k = 0 \,. \tag{3}$$

Lee (1996, §4) explores a general approach for finding the mixing function $\psi_k(y_k)$, and proves that for marginal distributions with support in $\mathbb{R}_+$ the mixing function is given by:

[4] It is straightforward to write the results of this section for discrete variables. The same characterization could be obtained using probability frequency functions rather than probability density functions. Notice that $A_k \subseteqq \mathbb{R}$ can define the set of natural numbers. See Examples 6.2 and 6.3 in Lee (1996).

$$\psi_k(y_k) = \exp(-y_k) - L_k(1), \qquad \forall y_k \geq 0\,, \tag{4}$$

where $L_k(1)$ is the value of the Laplace transform of the marginal distribution evaluated at $\zeta = 1$, that is:

$$L_k(\zeta) = \int_0^\infty \exp(-\zeta y_k) f_k(y_k) dy_k \qquad \text{at } \zeta = 1\,. \tag{5}$$

For expression (2) to properly define a bivariate density function the value of ω_{12} needs to fulfill the following constraint:

$$\omega_{12} \in \mathbb{R} :\ 1 + \omega_{12}\psi_1(y_1)\psi_2(y_2) \geq 0 \qquad \forall y_1, y_2\,. \tag{6}$$

The Sarmanov family contains the Farlie-Gumbel-Morgenstern family of distributions as noticed by Johnson, Balakrishnan, and Kotz (2000, §44.13). The Sarmanov family shows wider range for correlation coefficients. Indeed, Lee (1996) shows that restriction (6) holds when ω_{12} falls within the following bounds:

$$\underline{\omega}_{12} = \frac{-1}{\max\{L_1(1)L_2(1), [1 - L_1(1)][1 - L_2(1)]\}} \leq \omega_{12} \leq \frac{1}{\max\{L_1(1)[1 - L_2(1)], [1 - L_1(1)]L_2(1)\}} = \overline{\omega}_{12}\,, \tag{7}$$

which needs to be fulfilled for every observation in the sample. Finally, notice that integrating $y_1 y_2$ with respect to (2) and making use of (9), the product moment is:

$$E\left[y_1 y_2\right] = \mu_1\mu_2 + \omega_{12}\nu_1\nu_2\,, \tag{8}$$

where ν_k indicates the mixing function weighted mean:

$$\nu_k = \int_{-\infty}^{\infty} y_k \psi_k(y_k) f_k(y_k) dy_k = -L_k'(1) - L_k(1)\mu_k\,, \tag{9}$$

which makes use of the definition of $L_k(1)$ given in (5) and where $L_k'(1)$ denotes the value of the derivative of the Laplace transform of the assumed marginal distribution also evaluated at $\zeta = 1$. Thus, the correlation coefficient of a well defined Sarmanov distribution can be written as:

$$\rho_{12} = \frac{\omega_{12}\nu_1\nu_2}{\sigma_1\sigma_2}\,. \tag{10}$$

Thus, when $\omega_{12} = 0$ the correlation parameter is $\rho_{12} = 0$, and variables y_1 and y_2 are independent.[5]

[5] Although most econometric studies focus on the estimation of correlation coefficients, other measures of associations could also be computed. Shubina and Lee (2004) study how condition (7) constrains the valid range of correlation (10) for different marginal distributions. These authors also show that the maximum ranges of other association measures such as Kendall's *tau* and Spearman's rank correlation coefficient are independent of the assumed marginal distributions.

3 A Bivariate Double Poisson-Sarmanov Count Data Regression Model

Let $y_k = 0, 1, 2, \ldots$ be distributed according to a double Poisson distribution with parameters μ_k and θ_k, conditional on a set of regressors $\boldsymbol{x_k}$ in a sample with $i = 1, 2, \ldots, n$ observations. Efron (1986) shows that the probability frequency function of a double Poisson distribution is:

$$\tilde{f}_k(y_k|\mu_k, \theta_k) = c(\mu_k, \theta_k) f_k(y_k|\mu_k, \theta_k) \,, \tag{11a}$$

$$f_k(y_k|\mu_k, \theta_k) = \sqrt{\theta_k} \exp(-\theta_k \mu_k) \exp(-y_k) \frac{{y_k}^{y_k}}{y_k!} \left(\frac{e\mu_k}{y_k} \right)^{\theta_k y_k} \,, \tag{11b}$$

$$\frac{1}{c(\mu_k, \theta_k)} = \sum_{y_k=0}^{\infty} f_k(y_k|\mu_k, \theta_k) \simeq 1 + \frac{1-\theta_k}{12\theta_k \mu_k} \left(1 + \frac{1}{\theta_k \mu_k} \right) \,, \tag{11c}$$

where $e = \exp(1)$, $\tilde{f}_k(y_k|\mu_k, \theta_k)$ denotes the exact double Poisson density, and $f_k(y_k|\mu_k, \theta_k)$ is the approximate probability mass function for the double Poisson family. The constant $c(\mu_k, \theta_k)$ makes $\tilde{f}(y_k|\mu_k, \theta_k)$ integrate to 1. Efron (1986) shows that $c(\mu_k, \theta_k)$ nearly equals 1 and thus he concludes that $f_k(y_k|\mu_k, \theta_k)$ is a good approximation for $\tilde{f}_k(y_k|\mu_k, \theta_k)$. The obvious advantage of the double Poisson over the standard Poisson distribution is that the mean and variance do not depend on the same single parameter. Thus, Efron (1986) also shows that conditional on a set of regressors $\boldsymbol{x_k}$, the expected count corresponding to observation i and its variance are:

$$\mathrm{E}[y_{ki}|\boldsymbol{x_{ki}}] \simeq \mu_{ki} \,, \tag{12a}$$

$$\sigma_{ki}^2 = \mathrm{Var}[y_{ki}|\boldsymbol{x_{ki}}] \simeq \frac{\mu_{ki}}{\theta_k} \,. \tag{12b}$$

Hence, the double Poisson includes the standard Poisson as a particular case when $\theta_k = 1$ but it allows for overdispersion when $\theta_k < 1$ as well as for underdispersion if $\theta_k > 1$. As it is commonly the case for count data regression models , I will specify an exponential mean function relating the observable characteristics to the expected number of counts:

$$\mu_{ki} = \exp\left(\boldsymbol{x'_{ki}} \boldsymbol{\beta_k} \right) \,. \tag{13}$$

Using Stirling's formula $z! \simeq \sqrt{2\pi z} \cdot z^z \cdot \exp(-z)$ for $z = y_k$ and $z = \theta_k y_k$, the frequency function (11b) is accurately approximated by:[6]

$$f_k(y_k|\mu_k, \theta_k) \simeq \theta_k \exp(-\theta_k \mu_k) \frac{(\theta_k \mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k + 1)} \,. \tag{14}$$

[6] The use of Stirling's formula eases the estimation considerably and explains the fact that many of the relationships presented below to build the double Poisson-Sarmanov model are indeed approximations rather than equalities.

This expression is useful to obtain an approximation to the Laplace transform of the double Poisson distribution, which evaluated at $\zeta = 1$ becomes:

$$L_k(1|\mu_k, \theta_k) \simeq c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} . \tag{15}$$

Thus, according to (4), the approximate mixing function of the double Poisson-Sarmanov distribution is:

$$\psi_k(y_k|\mu_k, \theta_k) \simeq \exp(-y_k) - c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} . \tag{16}$$

Next, approximating (9) for the case of the double Poisson distribution we get:

$$\nu_k(\mu_k, \theta_k) \simeq c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} (y_k - \mu_k) , \tag{17}$$

so that the approximate correlation coefficient of two double Poisson distributed counts corresponding to (10) is:

$$\rho_{12} \simeq \omega_{12} \prod_{k=1}^{2} \left\{ \frac{c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k)}{\sqrt{\mu_k/\theta_k}} \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_i)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} (y_k - \mu_k) \right\} = \omega_{12} \prod_{k=1}^{2} Q(\mu_k, \theta_k) . \tag{18}$$

Combining all these elements into (2) we obtain the probability of observing simultaneously a pair of counts $\{y_1, y_2\}$ according to the double Poisson-Sarmanov distribution:

$$f_{12}(y_1, y_2) \simeq \left(\prod_{k=1}^{2} \left\{ c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \frac{(\theta_k \mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k + 1)} \right\} \right) \times \left(1 + \rho_{12} \frac{\prod_{m=1}^{2} \left\{ \exp(-y_m) - c(\mu_m, \theta_m)\theta_m \exp(-\theta_m \mu_m) \sum_{y_m=0}^{\infty} \frac{(\theta_m \mu_m)^{\theta_m y_m} \exp(-y_m)}{\Gamma(\theta_m y_m + 1)} \right\}}{\prod_{m=1}^{2} Q(\mu_m, \theta_m)} \right) . \tag{19}$$

In addition, for this bivariate double Poisson-Sarmanov to be coherent, the constraint corresponding to the general case (7) needs to hold. Making use of (15) and (18), the constraint can be written in terms of ρ_{12} as follows:

$$\underline{\omega}_{12} \prod_{k=1}^{2} Q(\mu_k, \theta_k) \ \leq \rho_{12} \leq \ \overline{\omega}_{12} \prod_{k=1}^{2} Q(\mu_k, \theta_k) \qquad \forall\, i. \tag{20}$$

4 A Bivariate Gamma-Sarmanov Count Data Regression Model

In this section I focus on the bivariate version of the Sarmanov count data regression model assuming gamma distributed arrival times that define the counting process.[7] For convenience, let $t_k \geq 0$ denote the waiting time for events of type k to occur. The joint distribution of waiting times of events of type 1 and 2 is represented by $g_{12}(t_1, t_2)$ while $g_k(t_k)$, $k = 1, 2$ are the corresponding marginal distributions. I assume that these marginal distributions of waiting time have the following gamma probability density functions with shape parameter $\theta_k > 0$ and inverse scale parameter $\alpha_k > 0$:

$$g_k(t_k|\theta_k, \alpha_k) = \frac{\alpha_k^{\theta_k} t_k^{\theta_k - 1} \exp(-\alpha_k t_k)}{\Gamma(\theta_k)} , \tag{21}$$

with mean $\mu_k = \theta_k / \alpha_k$ and variance $\sigma_k^2 = \theta_k / \alpha_k^2$ and where $\Gamma(\theta_k)$ denotes the gamma function:

$$\Gamma(\theta_k) = \int_0^\infty \eta^{\theta_k - 1} \exp(-\eta) d\eta \, . \tag{22}$$

Let consider a renewal process where $\{t_{q_k} > 0\}$, for $q_k \in \mathbb{N}$ denotes the waiting times between the $(q_k - 1)^{th}$ and the q_k^{th} event of type k. Waiting times between events of the same type are assumed to be independently and identically distributed with gamma density function (21). Notice that in any given time interval $(0, T)$ the number of observed events of type k, N_k^T, (*i.e.*, renewals of type k) is a count variable that is determined by the independent realizations of the corresponding waiting times. Cox (1962, p.40) shows that the asymptotic distribution of N_k^T is:

$$N_k^T \sim \boldsymbol{N} \left[\frac{T}{\mu_k}, \frac{\sigma_k^2 T}{\mu_k^3} \right] = \boldsymbol{N} \left[\frac{\alpha_k T}{\theta_k}, \frac{\alpha_k T}{\theta_k^2} \right] . \tag{23}$$

Thus, the variance-to-mean ratio of the asymptotic distribution of N_k^T is $1/\theta_k$, which leads to the result that the distribution of counts of type k is asymptotically underdispersed whenever $\theta_k > 1$ and *vice versa*. As in (13), I parameterize the expected number of events as the following exponential mean function:

$$\frac{\alpha_{ki}}{\theta_k} = \exp\left(\boldsymbol{x}'_{\boldsymbol{ki}} \boldsymbol{\beta_k} \right) , \tag{24}$$

where $\boldsymbol{x_{ki}}$ is the vector of k regressors of observation i that enter the specification of the mean and variance of the distribution of N_k^T. The set of parameters to be estimated includes the scalar θ_k and the vector $\boldsymbol{\beta_k}$

[7] Statisticians call this model renewal process with unobservable gamma interarrival times. To my knowledge, this is the first model that addresses the estimation of a bivariate renewal process. For univariate renewal processes. See Miller and Bhat (1997).

so that equation (24) coincides with the common parameterization of the Poisson regression model when $\theta_k = 1$. The arrival time of the q_k^{th} event $(q_k = 1, 2, 3, \ldots)$ is just the sum of the waiting times of q_k renewals:

$$\tau_{q_k} = \sum_{\vartheta_k=1}^{q_k} t_{\vartheta_k}. \tag{25}$$

Due to the reproductive property of the gamma distribution, the sum (25) of independent and identically distributed gamma variables is also gamma distributed with the following probability density function:

$$g_k(\tau_{q_k}|q_k\theta_k, \alpha_k) = \frac{\alpha_k^{q_k\theta_k}\tau_{q_k}^{q_k\theta_k-1}\exp(-\alpha_k\tau_{q_k})}{\Gamma(q_k\theta_k)}\,, \tag{26}$$

with mean $\mu_{q_k} = q_k\theta_k/\alpha_k$ and variance $\sigma^2_{q_k} = q_k\theta_k/\alpha_k^2$. It follows that the (marginal) probability distribution function of τ_{q_k} is simply the incomplete gamma function ratio:

$$\text{Prob}(\tau_{q_k} \leq T|\alpha_k, \theta_k) = G_k(q_k\theta_k, \alpha_k T) = \int_0^T \frac{\alpha_k^{q_k\theta_k}}{\Gamma(q_k\theta_k)} s^{q_k\theta_k-1}\exp(-\alpha_k s)ds = \int_0^{\alpha_k T} \frac{z^{q_k\theta_k-1}}{\Gamma(q_k\theta_k)}\exp(-z)dz\,, \tag{27}$$

which indicates the probability that arrival time of the q_k^{th} event of type k does not exceeds T. Thus, within the fixed time interval $(0, T)$, the total number of events of type k, $q_k = N_k^T$, that occur are intimately related to the total waiting time of those renewals to take place:

$$y_k < q_k \iff \tau_{q_k} \geq T\,. \tag{28}$$

Therefore:

$$\text{Prob}(y_k < q_k) = \text{Prob}(\tau_{q_k} \geq T|\alpha_k, \theta_k) = 1 - G_k(q_k\theta_k, \alpha_k T)\,. \tag{29}$$

Noticing that $G(0, \alpha_k T) = 1$, it is then straightforward to show that the probability that an exact number of events of type i takes place during the interval $(0, T)$ is:

$$\text{Prob}(y_k = q_k) = \text{Prob}(y_k < q_k + 1) - \text{Prob}(y_k < q_k) = G_k(q_k\theta_k, \alpha_k T) - G_k((q_k + 1)\theta_k, \alpha_k T)\,. \tag{30}$$

Proceeding similarly, we can account for the realization of two types of events in a given time interval $(0, T)$. The joint probability of observing at least N_1 and N_2 events of type 1 and 2, respectively is:

$$\text{Prob}(y_1 < q_1, y_2 < q_2) = \text{Prob}(\tau_{q_1} \geq T, \tau_{q_2} \geq T)\,, \tag{31}$$

where this bivariate survival function can be written as:

$$\text{Prob}(\tau_{q_1} \geq T, \tau_{q_2} \geq T) = 1 - \text{Prob}(\tau_{q_1} \leq T) - \text{Prob}(\tau_{q_2} \leq T) + \text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T)\,. \tag{32}$$

The probability of observing an exact number of events of each type can then be written as follows:

$$\begin{aligned}
\text{Prob}(y_1 = q_1, y_2 = q_2) = {} & [\text{Prob}(y_1 < q_1 + 1, y_2 < q_2 + 1) - \text{Prob}(y_1 < q_1, y_2 < q_2 + 1)] \\
& - [\text{Prob}(y_1 < q_1 + 1, y_2 < q_2) - \text{Prob}(y_1 < q_1, y_2 < q_2)] \\
= {} & \left[\text{Prob}(\tau_{q_1+1} \geq T, \tau_{q_2+1} \geq T) - \text{Prob}(\tau_{q_1} \geq T, \tau_{q_2+1} \geq T)\right] \\
& - \left[\text{Prob}(\tau_{q_1+1} \geq T, \tau_{q_2} \geq T) - \text{Prob}(\tau_{q_1} \geq T, \tau_{q_2} \geq T)\right]\,.
\end{aligned} \tag{33}$$

After applying (32) repeatedly to every element of (33) the general expression for the joint probability of encountering a particular pair of count realizations can be rewritten as:

$$\begin{aligned}
\text{Prob}(y_1 = q_1, y_2 = q_2) = {} & \left[\text{Prob}(\tau_{q_1+1} \leq T, \tau_{q_2+1} \leq T) - \text{Prob}(\tau_{q_1} \leq T, \tau_{q_2+1} \leq T)\right] \\
& - \left[\text{Prob}(\tau_{q_1+1} \leq T, \tau_{q_2} \leq T) - \text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T)\right]\,.
\end{aligned} \tag{34}$$

In order to write the bivariate gamma-Sarmanov distribution I first apply (4) to obtain the mixing function, which is:

$$\psi_k(\tau_{q_k} | q_k\theta_k, \alpha_k) = \exp(-\tau_{q_k}) - \left(\frac{\alpha_k}{1+\alpha_k}\right)^{q_k\theta_k}\,, \tag{35}$$

since the Laplace transform of the gamma distribution evaluated at $\zeta = 1$ is:

$$L_k(1 | q_k\theta_k, \alpha_k) = \left(\frac{\alpha_k}{1+\alpha_k}\right)^{q_k\theta_k}\,. \tag{36}$$

Second, the mixing function weighted mean (9) of a counting process generated by a gamma distributed arrival time is:

$$\nu_{q_k} = \frac{-q_k\theta_k}{\alpha_k}\left(\frac{1}{1+\alpha_k}\right)\left(\frac{\alpha_k}{1+\alpha_k}\right)^{q_k\theta_k}\,, \tag{37}$$

so that after computing (10) the correlation between arrival times t_1 and t_2 is:

$$\rho_{12} = \omega_{12}\prod_{i=1}^{2}\sqrt{q_k\theta_k}\left(\frac{1}{1+\alpha_k}\right)\left(\frac{\alpha_k}{1+\alpha_k}\right)^{q_k\theta_k}\,. \tag{38}$$

Finally, I use the following property of the gamma function repeatedly to simplify the expression of the gamma-Sarmanov distribution:

$$g_k(\tau_{q_k}|q_k\theta_k, \alpha_k + 1) = \left(\frac{\alpha_k}{1+\alpha_k}\right)^{-q_k\theta_k} \exp(-\tau_{q_k}) g_k(\tau_{q_k}|q_k\theta_k, \alpha_k)\,. \tag{39}$$

Combining (2), (35), (38), and (39), the joint probability density function of the gamma-Sarmanov distributed arrival times τ_{q_1} and τ_{q_2} can be written as follows:

$$\begin{aligned}
&g_{12}(\tau_{q_1}, \tau_{q_2}|q_1\theta_1, q_2\theta_2, \alpha_1, \alpha_2, \rho_{12}) = \\
&= \left[\prod_{i=1}^{2} \frac{\alpha_k^{q_k\theta_k} \tau_{q_k}^{q_k\theta_k - 1} \exp(-\alpha_k \tau_{q_k})}{\Gamma(q_k\theta_k)}\right] \times \left[1 + \omega_{12} \prod_{i=1}^{2} \left(\exp(-\tau_{q_k}) - \left(\frac{\alpha_k}{1+\alpha_k}\right)^{q_k\theta_k}\right)\right] \\
&= \left[\prod_{i=1}^{2} \frac{\alpha_k^{q_k\theta_k} \tau_{q_k}^{q_k\theta_k - 1} \exp(-\alpha_k \tau_{q_k})}{\Gamma(q_k\theta_k)}\right] \times \Bigg[1 + \omega_{12}\Bigg(\exp(-\tau_{q_1} - \tau_{q_2}) - \exp(-\tau_{q_1}) \left(\frac{\alpha_2}{1+\alpha_2}\right)^{q_2\theta_2} \\
&\qquad - \exp(-\tau_{q_2}) \left(\frac{\alpha_1}{1+\alpha_1}\right)^{q_1\theta_1} + \left(\frac{\alpha_1}{1+\alpha_1}\right)^{q_1\theta_1} \left(\frac{\alpha_2}{1+\alpha_2}\right)^{q_2\theta_2} \Bigg)\Bigg] \qquad (40) \\
&= g_1(\tau_{q_1}|q_1\theta_1, \alpha_1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2) + \rho_{12} \frac{(1+\alpha_1)(1+\alpha_2)}{\sqrt{q_1\theta_1 q_2\theta_2}} \times \\
&\qquad \Big[g_1(\tau_{q_1}|q_1\theta_1, \alpha_1 + 1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2 + 1) - g_1(\tau_{q_1}|q_1\theta_1, \alpha_1 + 1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2) \\
&\qquad - g_1(\tau_{q_1}|q_1\theta_1, \alpha_1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2 + 1) + g_1(\tau_{q_1}|q_1\theta_1, \alpha_1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2)\Big] \\
&= g_1(\tau_{q_1}|q_1\theta_1, \alpha_1) g_2(\tau_{q_2}|q_2\theta_2, \alpha_2) + \rho_{12} h_1(\tau_{q_1}|q_1\theta_1, \alpha_1) h_2(\tau_{q_2}|q_2\theta_2, \alpha_2)\,,
\end{aligned}$$

where for the sake of a more concise notation I made use of the following two definitions:

$$h_k(\tau_{q_k}|q_k\theta_k, \alpha_k) = \Big[g_k(\tau_{q_k}|q_k\theta_k, \alpha_k + 1) - g_k(\tau_{q_k}|q_k\theta_k, \alpha_k)\Big] \frac{1+\alpha_k}{\sqrt{q_k\theta_k}}\,, \tag{41a}$$

$$H_k(q_k\theta_k, \alpha_k) = \int_0^T h_k(s_k|q_k\theta_k, \alpha_k) ds_k\,. \tag{41b}$$

A most useful feature of the gamma-Sarmanov family is that due to the reproductive property of the gamma distribution (26), the evaluation of the bivariate distribution of counts reduces to a linear combination of products of single dimensional and easy to compute integrals (incomplete gamma functions). To see this, notice that the join distribution function of τ_{q_1} and τ_{q_2} can be written as:[8]

[8] This expression could be use to estimate a multivariate duration model such as the analysis of multiple spells between births of Heckman and Walker (1990) using Swedish data and the multiple transition models reviewed by Van Den Berg (2001).

$$
\begin{aligned}
\text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T) &= \int_0^T \int_0^T g_{12}(s_1, s_2 | q_1\theta_1, q_2\theta_2, \alpha_1, \alpha_1, \rho_{12}) ds_1 ds_2 \\
&= G_1(q_1\theta_1, \alpha_1) G_2(q_2\theta_2, \alpha_2) + \rho_{12} H_1(q_1\theta_1, \alpha_1) H_2(q_2\theta_2, \alpha_2) \, .
\end{aligned}
\tag{42}
$$

Finally, applying (42) to each element of equation (34) and after some algebra we can obtain the probability of observing a particular pair of count events (q_1, q_2):

$$
\begin{aligned}
\text{Prob}(y_1 = q_1, y_2 = q_2) &= \left[G_1\left(q_1\theta_1, \alpha_1\right) - G_1\left((q_1+1)\theta_1, \alpha_1\right) \right] \left[G_2\left(q_2\theta_2, \alpha_2\right) - G_2\left((q_2+1)\theta_2, \alpha_2\right) \right] \\
&\quad + \rho_{12} \Big[\left[H_1\left(q_1\theta_1, \alpha_1\right) - H_1\left((q_1+1)\theta_1, \alpha_1\right) \right] \left[H_2\left(q_2\theta_2, \alpha_2\right) - H_2\left((q_2+1)\theta_2, \alpha_2\right) \right] \Big] \\
&= \prod_{k=1}^{2} \left[(1 - L_{q_k}^{-1}) G_k\left(q_k\theta_k, \alpha_k\right) \right] + \rho_{12} \prod_{k=1}^{2} \left[(1 - L_{q_k}^{-1}) H_k\left(q_k\theta_k, \alpha_k\right) \right] \\
&= \left(\prod_{k=1}^{2} \left[(1 - L_{q_k}^{-1}) G_k\left(q_k\theta_k, \alpha_k\right) \right] \right) \times \left(1 + \rho_{12} \prod_{k=1}^{2} \frac{\left[(1 - L_{q_k}^{-1}) H_k\left(q_k\theta_k, \alpha_k\right) \right]}{\left[(1 - L_{q_k}^{-1}) G_k\left(q_k\theta_k, \alpha_k\right) \right]} \right) ,
\end{aligned}
\tag{43}
$$

where $L_{q_k}^{-1}$ is the lead operator applied on the q_k-component of functions $G(\cdot)$ and $H(\cdot)$.

Thus, contrary to all existing multivariate count data models, the bivariate gamma-Sarmanov count data regression model reduces the evaluation of a multidimensional probability density function to a linear combination of single dimensional integrals. This feature of the model survives the generalization of the family of bivariate gamma-Sarmanov distributions to higher dimensions. The only limitation of the present approach is that while the correlation coefficient can take positive or negative values independently of the dispersion pattern of the marginal distributions, its range is constrained by the estimates of the rest of parameters through equation (7). After substituting (36) into (7) and combining the result with (38) the constraint of the gamma-Sarmanov distribution can be written in terms of ρ_{12} as follows :

$$
\begin{aligned}
&\frac{- \prod_{k=1}^{2} \sqrt{q_k\theta_k} \left(\frac{1}{1+\alpha_k} \right) \left(\frac{\alpha_k}{1+\alpha_k} \right)^{q_k\theta_k}}{\max \left\{ \left(\frac{\alpha_1}{1+\alpha_1} \right)^{q_1\theta_1} \left(\frac{\alpha_2}{1+\alpha_2} \right)^{q_2\theta_2}, \left[1 - \left(\frac{\alpha_1}{1+\alpha_1} \right)^{q_1\theta_1} \right] \left[1 - \left(\frac{\alpha_2}{1+\alpha_2} \right)^{q_2\theta_2} \right] \right\}} \leq \rho_{12} \leq \\
&\frac{\prod_{k=1}^{2} \sqrt{q_k\theta_k} \left(\frac{1}{1+\alpha_k} \right) \left(\frac{\alpha_k}{1+\alpha_k} \right)^{q_k\theta_k}}{\max \left\{ \left(\frac{\alpha_1}{1+\alpha_1} \right)^{q_1\theta_1} \left[1 - \left(\frac{\alpha_2}{1+\alpha_2} \right)^{q_2\theta_2} \right], \left[1 - \left(\frac{\alpha_1}{1+\alpha_1} \right)^{q_1\theta_1} \right] \left(\frac{\alpha_2}{1+\alpha_2} \right)^{q_2\theta_2} \right\}}, \qquad \forall i \, .
\end{aligned}
\tag{44}
$$

5 Multivariate Extensions

Multivariate extensions of this model are of clear interest for practical purposes. Lee (1996, §8) suggests the following generalization of the joint density function of a multivariate Sarmanov distribution:

$$f_{1,2,\ldots,n}(y_1,\ldots,y_n) = \left[\prod_{k=1}^{n} f_k(y_k)\right] \times \left[1 + R_{\psi_1,\ldots,\psi_n,\boldsymbol{\Omega_n}}(y_1,\ldots,y_n)\right] , \tag{45}$$

where:

$$\begin{aligned} 1 + R_{\psi_1,\ldots,\psi_n,\boldsymbol{\Omega_n}}(y_1,\ldots,y_n) &= 1 + \sum\sum_{1\le l_1\le l_2\le n} \omega_{l_1 l_2} \prod_{m=1}^{2} \psi_{l_m}(y_{l_m}) \\ &+ \sum\sum\sum_{1\le l_1\le l_2\le l_3\le n} \omega_{l_1 l_2 l_3} \prod_{m=1}^{3} \psi_{l_m}(y_{l_m}) + \cdots + \omega_{12\ldots n} \prod_{m=1}^{n} \psi_m(y_m) \ge 0, \quad \forall y_1,\ldots,y_n . \end{aligned} \tag{46}$$

In this section I generalize both models to deal with three simultaneous count variables and I leave the multivariate generalization for Appendix A.

5.1 The Trivariate Double Poisson-Sarmanov Distribution

Extending the double Poisson-Sarmanov to more than two dimensions reduces to repeating the analysis of Section 3 and substituting the probability frequency function (14) and mixing function (16) into (45) and (46). The trivariate double Poisson-Sarmanov distribution follows easily from combining (19) and (45):

$$\begin{aligned} f_{123}(y_1,y_2,y_3) &= \left[\prod_{k=1}^{3}\left\{c(\mu_k,\theta_k)\theta_k \exp(-\theta_k\mu_k)\frac{(\theta_k\mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k+1)}\right\}\right] \times \\ &\left[1 + \sum\sum_{1\le l_1\le l_2\le n} \rho_{l_1 l_2} \frac{\prod_{m=1}^{2}\left\{\exp(-y_{l_m}) - c(\mu_{l_m},\theta_{l_m})\theta_{l_m}\exp(-\theta_{l_m}\mu_{l_m})\sum_{y_{l_m}=0}^{\infty}\frac{(\theta_{l_m}\mu_{l_m})^{\theta_{l_m}y_{l_m}}\exp(-y_{l_m})}{\Gamma(\theta_{l_m}y_{l_m}+1)}\right\}}{\prod_{m=1}^{2} Q(\mu_{l_m},\theta_{l_m})}\right. \\ &\left. + \rho_{123}\frac{\prod_{m=1}^{3}\left\{\exp(-y_m) - c(\mu_m,\theta_m)\theta_m\exp(-\theta_m\mu_m)\sum_{y_m=0}^{\infty}\frac{(\theta_m\mu_m)^{\theta_m y_m}\exp(-y_m)}{\Gamma(\theta_m y_m+1)}\right\}}{\prod_{m=1}^{3} Q(\mu_m,\theta_m)}\right] , \end{aligned} \tag{47}$$

where, similarly to equation (10), the higher order correlation coefficient is given by:

$$\rho_{123} = \frac{\omega_{123}\nu_1\nu_2\nu_3}{\sigma_1\sigma_2\sigma_3}\,, \tag{48}$$

which in turn makes use of the product moment:

$$E\left[y_1y_2y_3\right] = \mu_1\mu_2\mu_3 + \omega_{123}\nu_1\nu_2\nu_3\,. \tag{49}$$

5.2 The Trivariate Gamma-Sarmanov Distribution

Extending the bivariate gamma-Sarmanov distribution of Section 4 to several dimensions requires obtaining an expression for $\text{Prob}(y_1\!=\!q_1,\ldots,y_n\!=\!q_n)$ similar to (34) using the equivalence between the probability of a multivariate count event and the corresponding multivariate renewal process as stated in equation (31). The only added difficulty of extending the gamma-Sarmanov model to multiple dimensions has to do with the multivariate equivalent of constraint (44), which becomes increasingly more complicated as the number of dimensions, n, increases.

Deriving the probability of encountering a particular combination of three counts according to the gamma-Sarmanov distribution is more involved although it has the advantage that it can be expressed as a linear combination of single dimensional integrals and that it can be easily generalized to higher dimensional cases. To begin with, the trivariate survival function can be written as follows:

$$\begin{aligned}\text{Prob}(\tau_{q_1} \geq T, \tau_{q_2} \geq T, \tau_{q_3} \geq T) = 1 &- \text{Prob}(\tau_{q_1} \leq T) - \text{Prob}(\tau_{q_2} \leq T) - \text{Prob}(\tau_{q_3} \leq T)\\ &+\text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T) + \text{Prob}(\tau_{q_1} \leq T, \tau_{q_3} \leq T)\\ &+\text{Prob}(\tau_{q_2} \leq T, \tau_{q_3} \leq T) - \text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T, \tau_{q_3} \leq T)\,.\end{aligned} \tag{50}$$

Following the same logic than in deriving equation (34), the probability of observing a particular combination of events of each type is:

$$\begin{aligned}&\text{Prob}(y_1 = q_1, y_2 = q_2, y_3 = q_3) =\\ &\qquad\Big(\left[\text{Prob}(y_1 \leq q_1+1, y_2 \leq q_2+1, y_3 \leq q_3+1) - \text{Prob}(y_1 \leq q_1, y_2 \leq q_2+1, y_3 \leq q_3+1)\right]\\ &\qquad-\left[\text{Prob}(y_1 \leq q_1+1, y_2 \leq q_2, y_3 \leq q_3+1) - \text{Prob}(y_1 \leq q_1, y_2 \leq q_2, y_3 \leq q_3+1)\right]\Big)\\ &\quad-\Big(\left[\text{Prob}(y_1 \leq q_1+1, y_2 \leq q_2+1, y_3 \leq q_3) - \text{Prob}(y_1 \leq q_1, y_2 \leq q_2+1, y_3 \leq q_3)\right]\\ &\qquad-\left[\text{Prob}(y_1 \leq q_1+1, y_2 \leq q_2, y_3 \leq q_3) - \text{Prob}(y_1 \leq q_1, y_2 \leq q_2, y_3 \leq q_3)\right]\Big)\,,\end{aligned} \tag{51}$$

so that:

$$
\begin{aligned}
&\text{Prob}(y_1 = q_1, y_2 = q_2, y_3 = q_3) = \\
&\qquad \Big(\big[\text{Prob}(\tau_{q_1+1} \geq 1, \tau_{q_2+1} \geq 1, \tau_{q_3+1} \geq 1) - \text{Prob}(\tau_{q_1} \geq 1, \tau_{q_2+1} \geq 1, \tau_{q_3+1} \geq 1)\big] \\
&\qquad\quad - \big[\text{Prob}(\tau_{q_1+1} \geq 1, \tau_{q_2} \geq 1, \tau_{q_3+1} \geq 1) - \text{Prob}(\tau_{q_1} \geq 1, \tau_{q_2} \geq 1, \tau_{q_3+1} \geq 1)\big] \Big) \\
&\qquad - \Big(\big[\text{Prob}(\tau_{q_1+1} \geq 1, \tau_{q_2+1} \geq 1, \tau_{q_3} \geq 1) - \text{Prob}(\tau_{q_1} \geq 1, \tau_{q_2+1} \geq 1, \tau_{q_3} \geq 1)\big] \\
&\qquad\quad - \big[\text{Prob}(\tau_{q_1+1} \geq 1, \tau_{q_2} \geq 1, \tau_{q_3} \geq 1) - \text{Prob}(\tau_{q_1} \geq 1, \tau_{q_2} \geq 1, \tau_{q_3} \geq 1)\big] \Big) .
\end{aligned}
\tag{52}
$$

Substituting (50) repeatedly into (52), the general expression for the probability of a particular combination of these three counts is:

$$
\begin{aligned}
&\text{Prob}(y_1 = q_1, y_2 = q_2, y_3 = q_3) = \\
&\qquad \Big(\big[\text{Prob}(\tau_{q_1+1} \leq 1, \tau_{q_2+1} \leq 1, \tau_{q_3+1} \leq 1) - \text{Prob}(\tau_{q_1} \leq 1, \tau_{q_2+1} \leq 1, \tau_{q_3+1} \leq 1)\big] \\
&\qquad\quad - \big[\text{Prob}(\tau_{q_1+1} \leq 1, \tau_{q_2} \leq 1, \tau_{q_3+1} \leq 1) - \text{Prob}(\tau_{q_1} \leq 1, \tau_{q_2} \leq 1, \tau_{q_3+1} \leq 1)\big] \Big) \\
&\qquad - \Big(\big[\text{Prob}(\tau_{q_1+1} \leq 1, \tau_{q_2+1} \leq 1, \tau_{q_3} \leq 1) - \text{Prob}(\tau_{q_1} \leq 1, \tau_{q_2+1} \leq 1, \tau_{q_3} \leq 1)\big] \\
&\qquad\quad - \big[\text{Prob}(\tau_{q_1+1} \leq 1, \tau_{q_2} \leq 1, \tau_{q_3} \leq 1) - \text{Prob}(\tau_{q_1} \leq 1, \tau_{q_2} \leq 1, \tau_{q_3} \leq 1)\big] \Big) .
\end{aligned}
\tag{53}
$$

The trivariate gamma-Sarmanov probability density function is then written similarly to equation (40):

$$
\begin{aligned}
&g_{123}(\tau_{q_1}, \tau_{q_2}, \tau_{q_3} | q_1\theta_1, q_2\theta_2, q_3\theta_3, \alpha_1, \alpha_2, \alpha_3, \rho_{12}, \rho_{13}, \rho_{23}, \rho_{123}) = \left(\prod_{k=1}^{3} \frac{\alpha_k \tau_{q_k}^{q_k\theta_k - 1} \exp(-\alpha_k \tau_{q_k})}{\Gamma(q_k\theta_k)} \right) \times \\
&\left(1 + \sum\sum_{1 \leq l_1 \leq l_2 \leq 3} \omega_{l_1 l_2} \prod_{m=1}^{2} \left[\exp(-\tau_{q_{l_m}}) - \left(\frac{\alpha_{l_m}}{1 + \alpha_{l_m}} \right)^{q_{l_m}\theta_{l_m}} \right] + \omega_{123} \prod_{m=1}^{3} \left[\exp(-\tau_{q_m}) - \left(\frac{\alpha_m}{1 + \alpha_m} \right)^{q_m\theta_m} \right] \right) ,
\end{aligned}
\tag{54}
$$

where correlation coefficients ρ_{ij} are defined in terms of ω_{ij} as in equations (48)-(49), while the higher order correlation coefficient ρ_{123} is defined as in (48)-(49). In both cases, these correlations do not refer to y_1, y_2 and y_3 but to the arrival times τ_1, τ_2 and τ_3. After making extensive use of these relationships, definitions (41a)-(41b), and some algebra, the trivariate gamma-Sarmanov probability function becomes:

$$
\begin{aligned}
&g(\tau_{q_1}, \tau_{q_2}, \tau_{q_3} | q_1\theta_1, q_2\theta_2, q_3\theta_3, \alpha_1, \alpha_2, \alpha_3, \rho_{12}, \rho_{13}, \rho_{23}, \rho_{123}) = g_1(\tau_{q_1} | q_1\theta_1, \alpha_1) g_2(\tau_{q_2} | q_2\theta_2, \alpha_2) g_3(\tau_{q_3} | q_3\theta_3, \alpha_3) \\
&\quad + \rho_{12} h_1(\tau_{q_1} | q_1\theta_1, \alpha_1) h_2(\tau_{q_2} | q_2\theta_2, \alpha_2) g_3(\tau_{q_3} | q_3\theta_3, \alpha_3) + \rho_{13} h_1(\tau_{q_1} | q_1\theta_1, \alpha_1) h_3(\tau_{q_3} | q_3\theta_3, \alpha_3) g_2(\tau_{q_2} | q_2\theta_2, \alpha_2) \\
&\quad + \rho_{23} h_2(\tau_{q_2} | q_2\theta_2, \alpha_2) h_3(\tau_{q_3} | q_3\theta_3, \alpha_3) g_1(\tau_{q_1} | q_1\theta_1, \alpha_1) + \rho_{123} h_1(\tau_{q_1} | q_1\theta_1, \alpha_1) h_2(\tau_{q_2} | q_2\theta_2, \alpha_2) h_3(\tau_{q_3} | q_3\theta_3, \alpha_3) .
\end{aligned}
\tag{55}
$$

and thus:

$$\begin{aligned}\text{Prob}(\tau_{q_1} \leq T, \tau_{q_2} \leq T, \tau_{q_3} \leq T) = \int_0^T\int_0^T\int_0^T g_{123}(s_1, s_2, s_3|\cdot)ds_1 ds_2 ds_3 = G_1(\tau_{q_1}|\cdot)G_2(\tau_{q_2}|\cdot)G_3(\tau_{q_3}|\cdot) \\ + \rho_{12}H_1(\tau_{q_1}|\cdot)H_2(\tau_{q_2}|\cdot)G_3(\tau_{q_3}|\cdot) + \rho_{13}H_1(\tau_{q_1}|\cdot)G_2(\tau_{q_2}|\cdot)H_3(\tau_{q_3}|\cdot) \\ + \rho_{23}G_1(\tau_{q_1}|\cdot)H_2(\tau_{q_2}|\cdot)H_3(\tau_{q_3}|\cdot) + \rho_{123}H_1(\tau_{q_1}|\cdot)H_2(\tau_{q_2}|\cdot)H_3(\tau_{q_3}|\cdot)\,.\end{aligned} \tag{56}$$

Substituting this expression into each element of (53) and grouping them conveniently leads to the following operational expression to evaluate the probability of encountering a particular combination of three count events:

$$\begin{aligned}\text{Prob}(y_1 = q_1, y_2 = q_2, y_3 = q_3) = \left(\prod_{k=1}^{3}\left[(1 - L_{q_k}^{-1})G_k(q_k\theta_k, \alpha_k)\right]\right) \times \\ \left(1 + \sum\sum_{1 \leq l_1 \leq l_2 \leq 3} \rho_{l_1 l_2} \prod_{m=1}^{2} \frac{\left[(1 - L_{q_{l_m}}^{-1})H_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]}{\left[(1 - L_{q_{l_m}}^{-1})G_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]} + \rho_{123} \prod_{m=1}^{3} \frac{\left[(1 - L_{q_m}^{-1})H_q(q_m\theta_m, \alpha_m)\right]}{\left[(1 - L_{q_m}^{-1})G_q(q_m\theta_m, \alpha_m)\right]}\right)\,.\end{aligned} \tag{57}$$

6 Estimation

Despite the apparently cumbersome notation, estimation of the two proposed models is relatively straightforward. Take for instance the scalar realizations y_{1i} and y_{2i} of two count random variables given two vectors of regressors $\boldsymbol{x_{1i}}$ and $\boldsymbol{x_{2i}}$, parameter vectors $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$, as well as parameter scalar ω_{12}. Estimation by maximum likelihood maximizes the probability of jointly observing $\{y_{11}, y_{21}\}, \{y_{12}, y_{22}\}, \ldots, \{y_{1n}, y_{2n}\}$ in the sample. Using the general Sarmanov distribution (2), the log-likelihood function can be written as:

$$\mathcal{L}\left(\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}, \omega_{12}\right) = \sum_{i=1}^{n}\sum_{k=1}^{2} \ln f_k\left(y_{ki}|\boldsymbol{x_{ki}}, \boldsymbol{\gamma_{ki}}\right) + \sum_{i=1}^{n} \ln\left[1 + \omega_{12}\prod_{k=1}^{2}\psi_k\left(y_{ki}|\boldsymbol{x_{ki}}, \boldsymbol{\gamma_{ki}}\right)\right]\,. \tag{58}$$

Notice that ω_{12} only enters the term between brackets. The estimation thus proceeds iteratively, alternatively fixing the value of ω_{12} or $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ until we achieve convergence. Initial values $\hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\boldsymbol{\gamma}}_{\mathbf{2}}^{(\mathbf{0})}$ are obtained under the assumption of independence, *i.e.*, setting $\omega_{12} = 0$ and estimating two separate count data regression models. The initial estimate of ω_{12} is obtained by grid search, evaluating (58) over the interval defined by the constraint (7) while holding the estimated $\hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\boldsymbol{\gamma}}_{\mathbf{2}}^{(\mathbf{0})}$ constant. With this new value of $\hat{\omega}_{12}^{(0)}$, new estimates $\hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{(\mathbf{1})}$ and $\boldsymbol{\gamma}_{\mathbf{2}}^{(\mathbf{1})}$ are obtained by maximizing (58) while holding ω_{12} constant at the estimated value $\hat{\omega}_{12}^{(0)}$. The process is repeated until convergence is achieved.

Estimating a trivariate or multivariate model is slightly more convoluted because in principle equation (46) would allow for multiple combinations of correlations coefficients that fulfill such constraint. However, the solution to this maximization problem is unique because Lee (1996, Theorem 5a) states that if $\{y_1, y_2, \ldots, y_n\}$ are jointly distributed according to a n-variate Sarmanov distribution, then any subset of $\{y_1, y_2, \ldots, y_n\}$ will also be distributed as a Sarmanov distribution. To see how this helps estimating the different correlation coefficients of a multivariate Sarmanov distribution, consider the trivariate case. The log-likelihood function of the trivariate case is:

$$\begin{aligned}\mathcal{L}\left(\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}, \boldsymbol{\gamma_3}, \omega_{12}, \omega_{13}, \omega_{23}, \omega_{123}\right) = \sum_{i=1}^{n}\sum_{k=1}^{3} \ln f_k\left(y_{ki}|\boldsymbol{x_{ki}}, \boldsymbol{\gamma_{ki}}\right) \\ + \sum_{i=1}^{n} \ln\left[1 + \sum\sum_{1\le l_1\le l_2\le 3} \omega_{l_1 l_2} \prod_{m=1}^{2} \psi_{l_m}\left(y_{l_m i}|\boldsymbol{x_{l_m i}}, \boldsymbol{\gamma_{l_m i}}\right) + \omega_{123}\prod_{k=1}^{3}\psi_k\left(y_{ki}|\boldsymbol{x_{ki}}, \boldsymbol{\gamma_{ki}}\right)\right].\end{aligned} \tag{59}$$

Under the assumption of independence, single dimensional count data regressions produce initial estimates for $\boldsymbol{\gamma_1}$, $\boldsymbol{\gamma_2}$, and $\boldsymbol{\gamma_3}$. Conditioning on $\hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\boldsymbol{\gamma}}_{\mathbf{2}}^{(\mathbf{0})}$ in (58) we obtain the estimate $\omega_{12}^{(0)}$ by the grid search procedure described above. The same approach can be used to obtain estimates $\omega_{13}^{(0)}$ and $\omega_{23}^{(0)}$ while conditioning the likelihood function (58) on $\{\hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{(\mathbf{0})}, \hat{\boldsymbol{\gamma}}_{\mathbf{3}}^{(\mathbf{0})}\}$ and $\{\hat{\boldsymbol{\gamma}}_{\mathbf{2}}^{(\mathbf{0})}, \hat{\boldsymbol{\gamma}}_{\mathbf{3}}^{(\mathbf{0})}\}$, respectively. Then, maximizing (59) produces an estimate of ω_{123} while holding all the other parameters constant. Once $\hat{\omega}_{123}^{(0)}$ has been obtained, new estimates $\boldsymbol{\gamma}_{\mathbf{1}}^{(\mathbf{1})}$, $\boldsymbol{\gamma}_{\mathbf{2}}^{(\mathbf{1})}$, and $\boldsymbol{\gamma}_{\mathbf{3}}^{(\mathbf{1})}$ are estimated by maximizing (59) while holding $\{\hat{\omega}_{12}^{(0)}, \hat{\omega}_{13}^{(0)}, \hat{\omega}_{23}^{(0)}, \hat{\omega}_{123}^{(0)}\}$ constant. This procedure is then repeated until convergence is achieved.

At this point we need to be concerned with two issues. First, can we consistently estimate parameters that may lie on the boundary generically defined by condition (6)? Notice that the Sarmanov model only imposes a constraint on the correlation coefficient. All other parameters, although they define the range of variation of the correlation coefficient, remain unrestricted. Furthermore, the range defined by (7) is a compact convex set so that any estimate of the correlation coefficient includes its neighborhood, thus fulfilling the requirements of Andrews (2000, §4.2) who studies the asymptotic distributions of estimators when the true parameter lies on the boundary of the parameter space.[9]

The second issue refers to how we should obtain consistent inference for these parameter estimates. We need to address the possibility that estimated parameters may lie on the boundary defined by the corresponding constraint, *i.e.*, (20) and (44), respectively. Andrews (1999) shows that standard bootstrapping does not produce consistent inference when a parameter is on the boundary of the parameter space defined by a nonlinear inequality such as general conditions (6) and (46) for the bivariate and multivariate case, respectively. Rather than computing common bootstrap standard errors Andrews (1999, §6.4) suggests

[9] All estimates of correlation coefficients reported below fall in the interior of the compact convex set defined by $[\underline{\omega}_{12}, \overline{\omega}_{12}]$.

the use of a rescaled bootstrap method in which bootstrap samples of size $b < n$ are employed.[10] Andrews (2000, §4) shows that this modified bootstrapping approach produces consistent standard errors estimates regardless of whether the true parameter is on a boundary of the parameter space or not. I thus employ rescaled bootstrapping to produce consistent inference for these constrained maximum-likelihood estimates. To speed up the process of obtaining robust inference I follow Andrews (2002) and compute a 10-step version of the rescaled bootstrap.

Two additional issues regarding the econometric implementation of these two models are discussed at length in Appendix B. First, for the double Poisson-Sarmanov model we have to decide how many terms of the infinite sums in equations (15)-(19) to account for in the estimation. In general, longer series are needed the more overdispersed the distribution of counts is. Second, the realizations of the endogenous count variables critically condition the range of the correlation parameter in the gamma-Sarmanov case. Since q_k enters the numerator of constraint (44), the range of ρ_{12} becomes nil as soon as a single realization of q_1 and q_2 are both zero. This is not the case in the application of the next section as the minimum value of the endogenous variables is always one. One potential alternative to address this problem of the gamma-Sarmanov case is to re-scale the endogenous variables. However, Appendix B shows the allowable range of the correlation parameter of the corresponding waiting times widens non-monotonically as the minimum of the endogenous counts takes higher values. Therefore, in practice, the gamma-Sarmanov model is less capable of handling multivariate counts than the double Poisson-Sarmanov model, whose range of viable correlation coefficients is affected by the estimates of the rest of parameters of the model only, and not by the particular values of the endogenous counts.

7 Number of Tariff Options in Duopoly Competition

Firms engaging in price discrimination commonly offer a few tariff options to screen a heterogeneous customer base. In principle the set of fully nonlinear tariffs offered by two competing firms are the best response to each other's tariffs given the distribution of consumer heterogeneity. The few existing theoretical results on this area show that equilibrium in nonlinear tariffs exists both in common agency or in exclusive agency environments (see Stole (2005) for an overview). However, such results refer to fully nonlinear tariffs rather than how tariffs are commonly implemented, *i.e.*, , through a menu of tariff options.

[10] Horowitz (2001, §2.2) calls the rescaled bootstrapping *replacement subsampling* and discusses how it outperforms bootstrapping when the bootstrap is inconsistent. In the case of rescaled bootstrapping we use a given number of bootstrap samples of size b where some of these samples might be repeated. This differentiates rescaled bootstrapping from subsampling where some or all $n!/[b!(n-b)!]$ samples of size b (always without repetition) are employed in estimating each replication, *e.g.*, see Politis, Romano, and Wolf (1999, §2.1).

Table 1: Frequency Distributions of Number of Tariff Options

	1984–1988				*1992*			
	Incumbent		*Entrant*		*Incumbent*		*Entrant*	
Tariff Options	Cases	Rel.Freq.	Cases	Rel.Freq.	Cases	Rel.Freq.	Cases	Rel.Freq.
1	14	0.0269	3	0.0423	51	0.0979	5	0.0704
2	71	0.1363	7	0.0986	76	0.1459	3	0.0423
3	198	0.3800	5	0.0704	122	0.2342	13	0.1831
4	128	0.2457	16	0.2254	162	0.3109	18	0.2535
5	63	0.1209	40	0.5634	55	0.1056	32	0.4507
6	47	0.0902	0	0.0000	55	0.1056	0	0.0000
Mean, (Var.)	3.5681	(1.4651)	4.1690	(1.3996)	3.4971	(1.9774)	3.9718	(1.4563)

Absolute and relative frequency distributions of the number of tariff options offered by each active firm.

The use of few tariff options to screen consumer might be due to the existence of some commercialization costs or other marketing consideration. Thus, the foregone profits of an additional tariff will eventually not compensate such cost, as foregone profits decline rapidly with the number of tariff options, *e.g.*, Wilson (1993, §8.3). Commercialization costs may refer not only to the cost of designing and selling this additional tariff option, but also the money value of the reputation effect that such strategy may have with customers who might value tariff complexity negatively. In any case, if this were the only reason determining the number of tariffs options offered by each carrier, we should expect that the number of tariffs offered by the first competitor (conditional on available firm and market characteristics) were uncorrelated with the number of tariff options offered by the second firm in the absence of synergies across commercialization costs of different firms.

Alternatively, with non-zero correlations the number of tariff options offered becomes strategically relevant. If correlation among the conditional distribution of counts is positive firms tend to offer a similar number of tariff options and their numbers are strategic complements. This environment responds to equilibrium models of nonlinear pricing of Armstrong and Vickers (2001) and Rochet and Stole (2002) where firms end up offering similar, if not identical, two-part tariffs. On the contrary, if correlation among the number of tariff options offered is negative, firms might attempt to differentiate their products through pricing and therefore use the number of tariff options as strategic substitute. Yang and Ye (2008) show that this situation could arise in mature markets where business stealing, rather than expanding the base of active customer, is the main effect of price discrimination.

Data, which has been described at length elsewhere,[11] contains a complete description of the tariff options offered by any of the two firms present in the largest markets of the U.S. between 1984 and 1988.

[11] See for instance Busse (2000), Miravete and Röller (2004), and Parker and Röller (1997).

Table 2: Correlation Among Number of Tariff Options

Plans	1984–1988							1992						
	1	2	3	4	5	6	All	1	2	3	4	5	6	All
1	9	0	1	4	0	0	14	0	0	1	1	1	0	3
2	20	35	11	4	0	1	71	2	1	1	2	1	0	7
3	9	15	55	68	26	25	198	1	0	2	1	1	0	5
4	8	19	42	36	9	14	128	0	0	4	3	9	0	16
5	5	7	9	34	7	1	63	2	2	5	11	20	0	40
6	0	0	4	16	13	14	47	0	0	0	0	0	0	0
All	15	76	122	162	55	55	521	5	3	14	18	32	0	72
Kendall's τ		0.2928			(9.99)					0.1836			(2.26)	

Total cases for each combination of tariff options offered by the incumbent and entrant firm. Rows indicate the number of options of the entrant and columns those of the incumbent. Kendall's τ measures the association among the number of tariff options. The corresponding absolute value t-statistics are shown in parentheses. There are 521 pairs of tariff strategies in the 1984–1988 sample and 72 pairs in the 1992 sample.

I thus can compute the number of tariff options of each firm, PLANS. This information was collected by *Economic and Management Consultants International, Inc.* and reported in *Cellular Price and Marketing Letter*, Information Enterprises, various issues, 1984–1988. For year 1992, YEAR92, Marciano (2000) combined information from *Cellular Directions, Inc.*, the *Cellular Telephone Industry Association*, and direct interviews with managers.

Table 1 presents the histogram of the actual and effective number of tariff options by type of cellular carrier. Incumbents offer 3.5 and entrants 4 tariff options on average. When comparing pricing over time, it appears that there is a very slight reduction of options from 1984–1988 to 1992. Notice also that the unconditional distribution of tariff plans is always underdispersed, *i.e.*, the variance of the distribution of number of plans never exceeds the mean, which is the opposite of what most count data regression models address as the consequence of unobserved heterogeneity.

Table 2 indicates that the number of tariff options appear to be strategic complements when we measure the association between these strategies unconditionally of any firm or market observed heterogeneity. It is clear from this table that in the 1984–1988 period, firms frequently offer either the same or very similar number of tariff options. Between 1984 and 1988, firms offered the same number of tariff options in 30% of cases while in 71% of cases, the difference between the number of tariff plans offered by the incumbent and the entrant does not exceed one. In the 1992 sample these percentages increase up to 71% and 75% of cases, respectively.

Table 3: Descriptive Statistics

	Incumbent		*Entrant*	
Variables	Mean	Std.Dev.	Mean	Std.Dev.
PLANS	3.6402	1.2219	3.5541	1.3915
YEAR92	0.1199	0.3252	0.1199	0.3252
COMMUTING	3.1428	0.1512	3.1428	0.1512
POPULATION	0.0793	0.9583	0.0793	0.9583
EDUCATION	2.5752	0.0352	2.5752	0.0352
BUSINESS	3.2840	0.8876	3.2840	0.8876
GROWTH	0.9361	1.0274	0.9361	1.0274
INCOME	3.6406	0.1318	3.6406	0.1318
MULTIMARKET	3.1824	2.2808	3.1824	2.2808
REGULATED	0.5270	0.4997	0.5270	0.4997
AMERITECH	0.1554	0.3626	0.0942	0.2206
BELLATL	0.0574	0.2329	0.0671	0.1725
BELLSTH	0.0878	0.2833	0.0600	0.1652
CENTEL	0.0895	0.2857	0.0541	0.1623
CONTEL	0.0507	0.2195	0.0270	0.1204
GTE	0.1436	0.3510	0.0777	0.1970
MCCAW			0.2782	0.2473
NYNEX	0.0963	0.2952	0.0550	0.1734
PACTEL	0.0220	0.1467	0.0388	0.1354
SWBELL	0.1334	0.3403	0.0802	0.2174
USWEST	0.0895	0.2857	0.0566	0.1638

All variables are defined in the text. The number of observations is 592.

Table 3 presents the market and firm specific characteristics used in the estimation.[12] COMMUTING refers to the average daily commuting time in minutes in each city; POPULATION represents the number of inhabitants in each market measured in millions; EDUCATION is the median number of years of schooling; GROWTH is the average percent growth of the population in the 1980's; INCOME measures the median income in thousands of dollars; and BUSINESS accounts for the number of business in sectors with high demand for cellular services of each market and measured in thousands of firms.[13] With the exception of GROWTH, all these variables are measured in logarithms. Two other interesting market indicators are MULTIMARKET and REGULATED. The former is the number of markets in which a particular couple of firms

[12] Other regressors are available but they are not significant neither equation.

[13] Businesses with potential high cellular demand include service firms, health care, professional, and legal services, contract construction, transportation, finance, insurance, and real estate. The source of all these demographics is the 1989 *Statistical Abstracts of the United States*; U.S. Department of Commerce, Bureau of the Census, using the Federal Communication Commission (*FCC*) Cellular Boundary Notices, 1982–1987, available in *The Cellular Market Data Book*, EMCI, Inc., as well as the 1990 U.S. Decennial Census.

Table 4: Double Poisson – Sarmanov Regression

	Independent Regressions				*Sarmanov Regression*			
Variables	*Incumbent*		*Entrant*		*Incumbent*		*Entrant*	
CONSTANT	2.3986	(0.47)	–12.5831	(1.79)	2.3967	(25.87)	–12.6123	(27.21)
YEAR92	0.7212	(6.87)	0.6151	(4.18)	0.7115	(2.69)	0.6216	(2.98)
COMMUTING	–1.1548	(1.92)	1.5559	(2.06)	–1.1798	(11.88)	1.5674	(13.16)
POPULATION	–0.0489	(0.40)	0.0743	(0.59)	–0.0475	(0.41)	0.0652	(0.48)
EDUCATION	0.1227	(0.06)	2.8608	(0.97)	0.1284	(1.66)	2.8691	(20.19)
BUSINESS	0.0333	(0.26)	–0.2681	(2.01)	0.0237	(0.19)	–0.2694	(1.85)
GROWTH	0.0891	(1.51)	–0.4534	(6.76)	0.0874	(0.98)	–0.4500	(4.51)
INCOME	1.4633	(2.32)	1.3248	(1.64)	1.4941	(14.12)	1.3347	(13.57)
MULTIMARKET	0.0409	(1.80)	0.1082	(4.06)	0.0394	(0.80)	0.1037	(1.71)
REGULATED	0.0928	(0.87)	0.6520	(4.66)	0.0815	(0.54)	0.6342	(4.31)
AMERITECH	–0.2183	(0.87)	0.3169	(0.63)	–0.2299	(1.76)	0.2406	(1.80)
BELLATL	1.0770	(4.97)	0.1317	(0.31)	1.0957	(5.78)	0.0848	(0.54)
BELLSTH	–1.2825	(6.09)	–0.9200	(2.07)	–1.2362	(5.91)	–0.9414	(5.77)
CENTEL	–0.2719	(1.26)	1.3981	(2.94)	–0.2827	(1.58)	1.3036	(8.43)
CONTEL	–0.8500	(3.70)	–0.7116	(1.42)	–0.8524	(3.71)	–0.7340	(5.68)
GTE	–1.1022	(6.38)	–0.1997	(0.52)	–1.0929	(5.81)	–0.2429	(1.41)
MCCAW			0.8311	(2.76)			0.8508	(4.67)
NYNEX	0.9543	(5.30)	0.9591	(2.37)	0.9632	(5.53)	0.8880	(4.76)
PACTEL	–1.2295	(4.07)	–0.0734	(0.11)	–1.1967	(6.29)	–0.0748	(0.76)
SWBELL	–0.5886	(2.53)	0.0341	(0.06)	–0.5839	(3.49)	–0.0037	(0.03)
USWEST	–0.0150	(0.08)	0.7996	(1.69)	–0.0048	(0.03)	0.7491	(4.27)
θ	3.6324	(16.37)	2.3895	(15.24)	3.6585	(15.60)	2.4113	(13.82)
ρ					0.0396		(3.46)	
$-\ln \mathcal{L}$	830.16		942.49		1,766.60			

Marginal effects evaluated at the sample mean of regressors. Endogenous variables are the number of tariff options of of each competing firm. The absolute value of a 2,000 replication, scaled, 10-step, bootsrapped t-statistics are reported between parentheses.

compete against each other.[14] The latter is a dummy variable that indicates whether new tariffs need to be approved by the regulator.[15] Lastly, in order to control for potential firm effects, I include firm dummies to identify the largest shareholder of each cellular carrier (available from the *FCC*). Only those carriers with at least 4% of licenses in this sample are identified. They are: *Ameritech Mobile* (AMERITECH), *Bell Atlantic Mobile* (BELLATL), *Bell South Mobile* (BELLSTH), *Century Cellular* (CENTEL), *Contel Cellular* (CONTEL), *GTE Mobilnet* (GTE), *McCaw Communications* (MCCAW), *Nynex Mobile* (NYNEX), *PacTel Mobile Access* (PACTEL), *SouthWest Bell* (SWBELL), and *US West Cellular* (USWEST). Lastly, YEAR92 identifies those observations from 1992, when arguably, the cellular market had matured.

[14] Busse (2000) addresses the relationship between multimarket contact on collusion so that the offering of certain tariff features allow firms to coordinate pricing.

[15] Shew (1994) documents that the possibility of having request approval of new tariffs in the future prompts firms in this industry to offer an "excessive" number of options when they enter the market, the only time when they do not have to seek such approval as cost data is not yet available.

Table 5: Gamma – Sarmanov Regression

	Independent Regressions				*Sarmanov Regression*			
Variables	*Incumbent*		*Entrant*		*Incumbent*		*Entrant*	
CONSTANT	3.3596	(0.67)	–12.0638	(1.75)	13.6188	(16.22)	–39.2506	(17.74)
YEAR92	0.7200	(6.86)	0.6095	(4.21)	3.8871	(5.21)	2.0284	(4.47)
COMMUTING	–1.1582	(1.96)	1.6326	(2.21)	–5.3870	(10.28)	4.7878	(11.37)
POPULATION	–0.0467	(0.39)	0.0702	(0.57)	–0.2773	(0.55)	0.1628	(0.42)
EDUCATION	0.1294	(0.06)	2.9608	(1.03)	0.9545	(2.76)	9.7543	(17.20)
BUSINESS	0.0297	(0.23)	–0.2715	(2.09)	0.0318	(0.06)	–0.8568	(1.80)
GROWTH	0.0875	(1.49)	–0.4535	(6.86)	0.4239	(1.52)	–1.3151	(5.10)
INCOME	1.4560	(2.34)	1.2625	(1.59)	7.4807	(14.01)	4.2245	(11.48)
MULTIMARKET	0.0386	(1.69)	0.1066	(4.07)	0.1805	(1.35)	0.3038	(3.08)
REGULATED	0.0858	(0.82)	0.6436	(4.67)	0.4665	(0.98)	1.9622	(5.64)
AMERITECH	–0.2346	(0.94)	0.2848	(0.57)	–1.3512	(2.69)	0.4139	(1.07)
BELLATL	1.0929	(5.15)	0.1097	(0.26)	5.0187	(5.17)	–0.2938	(0.53)
BELLSTH	–1.2607	(5.96)	–0.9010	(2.06)	–6.4693	(6.64)	–3.3252	(6.25)
CENTEL	–0.2849	(1.34)	1.3797	(2.96)	–1.8101	(2.33)	3.3430	(7.80)
CONTEL	–0.8488	(3.71)	–0.7085	(1.43)	–4.6046	(6.19)	–2.5703	(6.47)
GTE	–1.0911	(6.39)	–0.2198	(0.58)	–5.5656	(7.29)	–1.2687	(2.42)
MCCAW			0.8317	(2.82)			2.1798	(3.95)
NYNEX	0.9610	(5.44)	0.9366	(2.34)	4.5916	(6.44)	2.4639	(4.80)
PACTEL	–1.2207	(4.08)	–0.0894	(0.13)	–6.0404	(5.91)	–0.7025	(2.13)
SWBELL	–0.5889	(2.57)	0.0191	(0.04)	–3.1432	(4.37)	–0.5999	(1.71)
USWEST	–0.0152	(0.08)	0.8032	(1.72)	–0.2628	(0.43)	1.8433	(4.20)
θ	4.4532	(13.86)	2.8243	(13.31)	4.7869	(17.00)	2.9505	(18.98)
ρ					0.0115		(2.77)	
$-\ln \mathcal{L}$	828.53		940.49		1,765.43			

Marginal effects evaluated at the sample mean of regressors. Endogenous variables are the number of tariff options of of each competing firm. The absolute value of a 2,000 replication, scaled, 10-step, bootsrapped t-statistics are reported between parentheses.

Tables 4 and 5 present the results of the estimation of the bivariate double-Poisson and gamma Sarmanov count data regressions model, respectively. In all cases, estimates capture the fact that the distribution of number of tariffs are positively correlated and underdispersed. These tables also report the estimates of the corresponding, restricted, independent, count data regression models. Table 4 shows that marginal effects are very similar but only when we compare the independent double Poisson and the double Poisson-Sarmanov regressions. However the estimation of the correlation coefficient ρ is significant and thus the estimates of the model under independence are no longer consistent. The correlation coefficient of the gamma-Sarmanov model of Table 5 is also positive and significant. Furthermore, notice that independent double Poisson and independent gamma count estimates are also very similar. Only the gamma-Sarmanov estimates with non-zero correlation are, in absolute value, an order of magnitude larger than any of the other estimates. Table 6 in Appendix B shows that this is because the estimates of the correlated gamma-Sarmanov model are not robust to the re-scaling of the endogenous counts.

Estimates of correlation between the counts are always small but positive and significant, therefore supporting the view that the number of tariff options offered by competing cellular carriers are strategic complements.[16] Thus, it appears that firms do not use more options to differentiate from each other and grow at the expense of the competitor. This result is reasonable in an early market where there is room for both firms to grow by signing up new customers. The likelihood ratio test rejects the independent regression model in favor of the Sarmanov regression in both cases, independently of the assumed marginal distribution. For the double Poisson case, the value of the test is 6.05 (0.01 p-value) while for the gamma case, this test is 7.17 (0.01 p-value). Estimates vary in non-systematic ways regardless of whether we consider the double Poisson-Sarmanov or the gamma-Sarmanov model, *i.e.*, the estimates obtained under assumption of independence do not always overestimate or underestimate those from the Sarmanov regression.

Ownership fixed effects are generally significant and market characteristics tend to have the same sign both as determinant of the number of tariffs of the incumbent and the entrant (although sometimes in one of the equations they fail to be significant). INCOME and YEAR92 have a positive effect on the number of tariffs offered by both competing carriers while MULTIMARKET only has a positive effect on the number of tariff options offered by the entrant. GROWTH and BUSINESS have a negative effect on the number of tariff options offered by the entrant only, the latter being only marginally significant. These negative signs could be reconciled with situations where dynamic pricing considerations and switching cost are present. In the absence of a fast growing economy or if business customers are not numerous, the entrant has to offer several tariff options to segment the market of smaller users in order to induce them to subscribe. Offering just one or few tariffs that are less expensive than those offered by the incumbent will not secure the bulk of high valuation customers as the incumbent has previously targeted them and locked them in long term contracts. Finally, COMMUTING is the only variable that shows a significant opposite effect for each firm: in markets with longer commuting times entrants offer more tariff options than incumbents.

8 Concluding Remarks

The Sarmanov count data models presented in this paper can accommodate both over and underdispersion and allows for the possibility that counts are not only positively but also negatively correlated. Furthermore, these two features of the joint distribution of counts are not driven by a common unobserved factor to all univariate marginal distributions and the parameterization of the likelihood function allows for all possible combinations of over or underdispersed marginals and correlation of any sign. The Sarmanov

[16] Correlation coefficient bounds given by (20) and (44) are $[-0.0014, 0.0480]$ and $[-0.0084, 0.0136]$ for the double Poisson-Sarmanov and the gamma-Sarmanov models, respectively. Although tight, they always include the estimate of the correlation coefficient and its neighborhood as interior points.

model thus overcomes most existing impediments to estimate a full multivariate count data regression model, thus reducing the risk of misspecification bias.

The paper has presented and estimated two particular versions of the Sarmanov model. The double Poisson-Sarmanov assumes that the marginal distributions of the counts are double-Poisson. This distribution defines the correlation directly on the counts and its implementation of arbitrarily close approximations to its Laplace transform. The gamma-Sarmanov model does not define the correlation among counts directly but rather on the waiting times that defines a multivariate renewal process. The advantage is that the multivariate likelihood function of the gamma-Sarmanov model takes advantage of the reproductive property of gamma distributions, and thus it can be expressed as a product of single-dimensional integrals. The major drawback of this second model though is that it is not appropriate to address situations where the endogenous counts take zero values as it turns the constraint on the correlation coefficient into a singleton.

Results from a particular application to study the pricing strategies of cellular carriers in the U.S. during the mid-1980s show that ignoring the possibility of correlation between the number of tariffs offered by the competing duopolists leads to biased estimates of the effects of explanatory variables included in the model. Positive correlations among the counts, even after controlling for market and firm characteristics indicate that firms do not attempt to differentiate from each other by offering a significantly different number of tariff plans. We can thus conclude that the number of tariff options offered by these firms can be considered strategic complements.

References

AITCHISON, J. AND C. H. HO (1989): "The Multivariate Poisson-Lognormal Distribution." *Biometrika*, 76, 643–653.

ANDREWS, D. W. K. (1999): "Estimation When a Parameter Is on a Boundary." *Econometrica*, 67, 1341–1383.

ANDREWS, D. W. K. (2000): "Inconsistency of the Bootstrap When a Parameter Is on a Boundary of the Parameter Space." *Econometrica*, 68, 399–405.

ANDREWS, D. W. K. (2002): "Higher-Order Improvements of a Computationally Attractive *k*-Step Bootstrap for Extremum Estimators." *Econometrica*, 70, 119–162.

ARMSTRONG, C. M. AND J. S. VICKERS (2001): "Competitive Price Discrimination." *RAND Journal of Economics*, 32, 579–605.

BUSSE, M. R. (2000): "Multimarket Contact and Price Coordination in the Cellular Telephone Industry." *Journal of Economics & Management Strategy*, 9, 287–320.

CAMERON, A. C. AND P. K. TRIVEDI (1998): *Regression Analysis of Count Data*. New York, NY: Cambridge University Press.

COX, D. R. (1962): *Renewal Theory*. New York, NY: John Wiley & Sons.

EFRON, B. (1986): "Double Exponential Families and Their Use in Generalized Linear Regression." *Journal of the American Statistical Association*, 81, 709–721.

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1984): "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica*, 52, 701–720.

GURMU, S. AND J. ELDER (2000): "Generalized Bivariate Count Data Regression Models." *Economics Letters*, 68, 31–36.

GURMU, S. AND J. ELDER (2008): "A Bivariate Zero-Inflated Count Data Regression Model with Unrestricted Correlation." *Economics Letters*, 100, 245–248.

HAUSMAN, J. A., B. H. HALL, AND Z. GRILICHES (1984): "Econometric Models for Count Data with an Application to the Patents—R & D Relationship." *Econometrica*, 52, 909–938.

HAUSMAN, J. A., G. LEONARD, AND D. L. MCFADDEN (1995): "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses due to Natural Resource Damage." *Journal of Public Economics*, 56, 1–30.

HECKMAN, J. J. AND J. R. WALKER (1990): "The Relationship between Wages and Income and the Timing and Spacing of Birthts." *Econometrica*, 58, 1411–1441.

HELLSTRÖM, J. (2006): "A Bivariate Count Data Model for Household Tourism Demand." *Journal of Applied Econometrics*, 21, 213–226.

HOROWITZ, J. L. (2001): "The Bootstrap." In J. J. Heckman and E. Leamer (eds.): *Handbook of Econometrics,* Vol. 5. New York, NY: North-Holland.

JOHNSON, N. L., N. BALAKRISHNAN, AND S. KOTZ (2000): *Continuous Multivariate Distributions,* Vol. 1, 2nd edn. New York, NY: John Wiley & Sons.

JUNG, C. J. AND R. WINKELMANN (1993): "Two Aspects of Labor Mobility: A Bivariate Poisson Regression Approach." *Empirical Economics*, 18, 543–556.

KOCHERLAKOTA, S. AND K. KOCHERLAKOTA (1993): *Bivariate Discrete Distributions*. New York, NY: Marcel Dekker.

LEE, M.-L. T. (1996): "Properties and Aplications of the Sarmanov Family of Bivariate Distributions." *Communications in Statistics, Part A — Theory and Methods*, 25, 1207–1222.

MARCIANO, S. (2000): "Pricing Policies in Oligopoly with Product Differentiation: The Case of Cellular Telephony." Ph.D. Thesis, University of Chicago Graduate School of Business.

MARSHALL, A. W. AND I. OLKIN (1990): "Multivariate Distributions Generated from Mixtures of Convolution and Product Families." In H. A. Block, A. R. Sampson, and T. H. Savits (eds.): *Topics in Statistical Dependence, IMS Lecture Notes Monograph*. Hayward, CA: Institute of Mathematical Statistics.

MEGHIR, C. AND J.-M. ROBIN (1992): "Frequency of Purchase and Estimation of Demand Equations." *Journal of Econometrics*, 53, 53–86.

MILLER, G. K. AND U. N. BHAT (1997): "Estimation for Renewal Processes with Unobservable Gamma or Erlang Interarrival Times." *Journal of Statistical Planning and Inference*, 61, 355–372.

MIRAVETE, E. J. AND L.-H. RÖLLER (2004): "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early U.S. Cellular Telephone Industry." Mimeo, University of Pennsylvania and Wissenschaftszentrum Berlin für Sozialforschung (WZB).

MUNKIN, M. AND P. K. TRIVEDI (1999): "Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models." *Econometrics Journal*, 2, 29–48.

PARKER, P. M. AND L.-H. RÖLLER (1997): "Collusive Conduct in Duopolies: Multimarket Contact and Cross-Ownership in the Mobile Telephone Industry." *RAND Journal of Economics*, 28, 304–322.

POLITIS, D. M., J. P. ROMANO, AND M. WOLF (1999): *Subsampling*. New York, NY: Springer-Verlag.

RIPHAHN, R. T., A. WAMBACH, AND A. MILLION (2003): "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation." *Journal of Applied Econometrics*, 18, 387–405.

ROCHET, J.-C. AND L. A. STOLE (2002): "Nonlinear Pricing with Random Participation." *The Review of Economic Studies*, 69, 277–311.

SARMANOV, O. V. (1966): "Generalized Normal Correlation and Two-Dimensional Fréchet Classes." *Soviet Mathematics — Doklady*, 168, 596–599.

SHEW, W. B. (1994): "Regulation, Competition, and Prices in the U.S. Cellular Telephone Industry." Mimeo, The American Enterprise Institute.

SHUBINA, M. AND M.-L. T. LEE (2004): "On Maximum Attainable Correlation and Other Measures of Dependence for the Sarmanov Family of Bivariate Distributions." *Communications in Statistics, Part A — Theory and Methods*, 33, 1031–1052.

STOLE, L. A. (2005): "Price Discrimination and Imperfect Competition." In C. M. Armstrong and R. H. Porter (eds.): *Handbook of Industrial Organization,* Vol. 3. New York, NY: North-Holland.

TERZA, J. V. AND P. W. WILSON (1990): "Analyzing Frequencies of Several Types of Events: A Mixed Multinomial-Poisson Approach." *Review of Economics and Statistics*, 72, 108–115.

VAN DEN BERG, G. J. (2001): "Duration Models: Specification, Identification and Multiple Durations." In J. J. Heckman and E. Leamer (eds.): *Handbook of Econometrics,* Vol. 5. New York, NY: North-Holland.

WILSON, R. B. (1993): *Nonlinear Pricing*. New York, NY: Oxford University Press.

WINDMEIJER, F. A. G. AND J. M. C. SANTOS-SILVA (1997): "Endogeneity in Count Data Models: An Application to Demand for Health Care." *Journal of Applied Econometrics*, 12, 281–294.

WINKELMANN, R. (1995): "Duration Dependence and Dispersion in Count-Data Models." *Journal of Business & Economics Statistics*, 13, 467–474.

WINKELMANN, R. (2000): "Seemingly Unrelated Negative Binomial Regression." *Oxford Bulletin of Economics and Statistics*, 62, 553–560.

YANG, H. AND L. YE (2008): "Nonlinear Pricing, Market Coverage, and Competition." *Theoretical Economics*, 3, 123–153.

Appendix

A The Multivariate Case

The multivariate double Poisson-Sarmanov distribution is obtained by extending equation (47) along the lines of the generalization (45) suggested by Lee (1996, §8):

$$
\begin{aligned}
f_{123}(y_1, y_2, y_3) &= \left[\prod_{k=1}^{3}\left\{c(\mu_k, \theta_k)\theta_k \exp(-\theta_k\mu_k)\frac{(\theta_k\mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k + 1)}\right\}\right] \times \\
&\left[1 + \sum\sum_{1\le l_1 \le l_2 \le n} \rho_{l_1 l_2} \frac{\prod_{m=1}^{2}\left\{\exp(-y_{l_m}) - c(\mu_{l_m}, \theta_{l_m})\theta_{l_m}\exp(-\theta_{l_m}\mu_{l_m})\sum_{y_{l_m}=0}^{\infty}\frac{(\theta_{l_m}\mu_{l_m})^{\theta_{l_m} y_{l_m}}\exp(-y_{l_m})}{\Gamma(\theta_{l_m} y_{l_m}+1)}\right\}}{\prod_{m=1}^{2} Q(\mu_{l_m}, \theta_{l_m})}\right. \\
&+ \sum\sum\sum_{1\le l_1 \le l_2 \le l_3 \le n} \rho_{l_1 l_2 l_3} \frac{\prod_{m=1}^{3}\left\{\exp(-y_{l_m}) - c(\mu_{l_m}, \theta_{l_m})\theta_{l_m}\exp(-\theta_{l_m}\mu_{l_m})\sum_{y_{l_m}=0}^{\infty}\frac{(\theta_{l_m}\mu_{l_m})^{\theta_{l_m} y_{l_m}}\exp(-y_{l_m})}{\Gamma(\theta_{l_m} y_{l_m}+1)}\right\}}{\prod_{m=1}^{3} Q(\mu_{l_m}, \theta_{l_m})} \\
&\left. + \cdots + \rho_{12\ldots n} \frac{\prod_{m=1}^{n}\left\{\exp(-y_m) - c(\mu_m, \theta_m)\theta_m\exp(-\theta_m\mu_m)\sum_{y_m=0}^{\infty}\frac{(\theta_m\mu_m)^{\theta_m y_m}\exp(-y_m)}{\Gamma(\theta_m y_m+1)}\right\}}{\prod_{m=1}^{n} Q(\mu_m, \theta_m)}\right].
\end{aligned}
\tag{A.1}
$$

For the gamma-Sarmanov model, multivariate extensions beyond three dimensions are easily obtained after tediously repeating the steps of Section 5.2. Suppose we want to estimate an n-variate, gamma-Sarmanov count data regression model. To derive the multivariate analogs of equations (32) and (50), we need a general expression for the multivariate survival function:

$$
\begin{aligned}
\text{Prob}(\tau_{q_1} \ge T, \ldots \tau_{q_n} \ge T) = 1 &- \sum_{k=1}^{n}\text{Prob}(\tau_{q_k} \le T) + \sum\sum_{1\le l_1\le l_2\le n}\text{Prob}(\tau_{q_{l_1}} \le T, \tau_{q_{l_2}} \le T) \\
&- \sum\sum\sum_{1\le l_1\le l_2\le l_3\le n}\text{Prob}(\tau_{q_{l_1}} \le T, \tau_{q_{l_2}} \le T, \tau_{q_{l_3}} \le T) \\
&+ \cdots + (-1)^n\text{Prob}(\tau_{q_{l_1}} \le T, \tau_{q_{l_2}} \le T, \tau_{q_{l_3}} \le T)\,.
\end{aligned}
\tag{A.2}
$$

Making use again of the properties of the gamma distribution, the multivariate density (45) leads to the following expression for the multivariate gamma-Sarmanov distribution that corresponds to (42) and (47) in the main text:

$$\begin{aligned}\mathrm{Prob}(\tau_{q_1} \leq T, \ldots, \tau_{q_n} \leq T) = \left(\prod_{k=1}^{n} G_k(\tau_{q_k}|\cdot)\right) \times \Bigg(1 + \sum\sum_{1\leq l_1 \leq l_2 \leq n} \rho_{l_1 l_2} \prod_{m=1}^{2} \frac{H_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})}{G_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})} + \\ \sum\sum\sum_{1\leq l_1 \leq l_2 \leq l_3 \leq n} \rho_{l_1 l_2 l_3} \prod_{m=1}^{3} \frac{H_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})}{G_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})} + \cdots + \rho_{12\ldots n} \prod_{m=1}^{n} \frac{H_m(q_m\theta_m, \alpha_m)}{G_m(q_m\theta_m, \alpha_m)}\Bigg) .\end{aligned} \tag{A.3}$$

Therefore, after computing $\mathrm{Prob}(y_1 = q_1, \ldots, y_n = q_n)$ from the multivariate survival equation (A.2) and substituting (A.3) into each term of this probability, the likelihood of observing any particular n-tuple $\{q_1, q_2, \ldots, q_n\}$ corresponding to equations (43) and (57) in the main text is:

$$\begin{aligned}&\mathrm{Prob}(y_1 = q_1, \ldots, y_n = q_n) = \\ &\left(\prod_{k=1}^{3} \left[(1 - L_{q_k}^{-1}) G_k(q_k\theta_k, \alpha_k)\right]\right) \times \Bigg(1 + \sum\sum_{1\leq l_1 \leq l_2 \leq n} \rho_{l_1 l_2} \prod_{m=1}^{2} \frac{\left[(1 - L_{q_{l_m}}^{-1}) H_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]}{\left[(1 - L_{q_{l_m}}^{-1}) G_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]} + \\ &\sum\sum\sum_{1\leq l_1 \leq l_2 \leq l_3 \leq n} \rho_{l_1 l_2 l_3} \prod_{m=1}^{3} \frac{\left[(1 - L_{q_{l_m}}^{-1}) H_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]}{\left[(1 - L_{q_{l_m}}^{-1}) G_{l_m}(q_{l_m}\theta_{l_m}, \alpha_{l_m})\right]} + \cdots + \rho_{12\ldots n} \prod_{m=1}^{n} \frac{\left[(1 - L_{q_m}^{-1}) H_m(q_m\theta_m, \alpha_m)\right]}{\left[(1 - L_{q_m}^{-1}) G_m(q_m\theta_m, \alpha_m)\right]}\Bigg) .\end{aligned} \tag{A.4}$$

B Econometric Implementation

Many of the expressions of the double Poisson-Sarmanov model include the sum of an infinite series that need to be approximated for the purpose of estimation. Equations (15)-(19) include the following element:

$$S_k(\mu_k, \theta_k) = \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} . \tag{B.1}$$

Notice that for $\theta_k = 1$, the series $S_k(\mu_k, \theta_k)$ converges to:

$$S_k(\mu_k, 1) = \sum_{y_k=0}^{\infty} \frac{[\mu_k \exp(-1)]^{y_k}}{y!} = \exp\left(\frac{\mu_k}{e}\right) , \tag{B.2}$$

because of the well known Taylor expansion of the exponential function. Using Stirling's formula:

$$\lim_{y_k \to \infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\sqrt{2\pi}\sqrt{\theta_k y_k}(\theta_k y_k)^{\theta_k y_k} \exp(-\theta_k y_k)} = \lim_{y_k \to \infty} \frac{1}{\sqrt{2\pi\theta_k y_k}} \cdot \lim_{y_k \to \infty} \left(\frac{\mu_k}{y_k}\right)^{\theta_k y_k} \cdot \lim_{y_k \to \infty} [\exp(-y_k)]^{1-\theta_k} = 0\,, \tag{B.3}$$

so that the sum $S_k(\mu_k, \theta_k)$ converges for any value of θ_k. However, the length of the series needed to approximate sum $S_k(\mu_k, \theta_k)$ varies greatly with θ_k. We can rewrite equation (B.1) as:

$$S_k(\mu_k, \theta_k) = \sum_{y_k=0}^{\infty} \frac{\left[(\theta_k \mu_k)^{\theta_k} \exp(-1)\right]^{y_k}}{y_k!} \frac{y_k!}{(\theta_k y_k)!}\,, \tag{B.4}$$

to show that the number of elements of the sum in (B.1) needed to approximate $S_k(\mu_k, \theta_k)$ with any given precision decreases with θ_k, and in general, longer series are needed to approximate $S_k(\mu_k, \theta_k)$ when the distribution of y_k is overdispersed. The following expression can be used to determine the length of the series, *i.e.*, the value of y_k needed to increase estimate of $S_k(\mu_k, \theta_k)$ by at least δ:[17]

$$y_k^{\star}(\delta) \in \operatorname*{argmin}_{y_k} \frac{1}{\sqrt{2\pi\theta_k y_k}} \cdot \left(\frac{\mu_k}{y_k}\right)^{\theta_k y_k} \cdot [\exp(-y_k)]^{1-\theta_k} = \delta\,. \tag{B.5}$$

Contrary to the double Poisson-Sarmanov, the gamma-Sarmanov model does not build upon the distribution of the counts but rather on the multivariate distribution of arrival times associated to each type of event. The consequence is that the constraint (44), or its multivariate counterpart (46), becomes extremely tight if there are observations with zeroes in any dimension. Indeed if the sample includes just one observation with zeros in all dimensions, $\{q_1, q_2\} = \{0, 0\}$, those constraints become a singleton. This result follows from constraint (44) in the bivariate case, as q_k enters the squared root of the numerator. Thus, (44) becomes $0 \leq \rho_{12} \leq 0$.

There are at least two possibilities to circumvent this difficulty. One is to restrict the application of the gamma-Sarmanov model to situations where the endogenous counts exclude zeroes. This is the case of the present application since cellular carriers need to offer at least one tariff option if they are active. The second alternative is to re-scale the endogenous variable to $y_k = q_k + \varsigma$. Table 6 compares the estimates of the main text to alternative definitions of y_k for $\varsigma \in \{-0.5, -0.25, 0.25, 0.50\}$ in order to evaluate whether this re-scaling has any effect on the sign and size of the estimated parameters and correlation coefficients. Sign and significance of parameter estimates are robust but not their size. Absolute estimates values become larger as ς increases while the estimate of ρ becomes larger the closer ς is to zero. Thus, re-scaling allows us to obtain an estimate of the parameters, including a non-zero ρ but the double Poisson-Sarmanov outperforms the gamma-Sarmanov model when the endogenous counts include zeroes.

---

[17] The estimation uses the first 40 elements of the infinite series to approximate (B.1), leading to an approximation error $\delta = 7 \times 10^{-107}$ for y_1 and $\delta = 1 \times 10^{-77}$ for y_2 at the current estimated values of the parameters. Consequently, varying the number of elements from 30 to 50 has no practical consequences for the estimation of the double Poisson-Sarmanov model.

Table 6: Gamma – Sarmanov Regression

Variables	$q_i - 0.50$		$q_i - 0.25$		$q_i + 0.25$		$q_i + 0.50$	
Incumbent								
CONSTANT	6.2553	(13.77)	9.6897	(14.95)	17.5328	(14.30)	22.8900	(12.89)
YEAR92	3.2848	(4.69)	3.7105	(4.47)	4.1992	(4.62)	4.5202	(4.43)
COMMUTING	–4.5129	(11.71)	–5.6022	(11.19)	–5.7804	(10.67)	–6.2680	(10.74)
POPULATION	–0.2318	(0.57)	–0.3348	(0.72)	–0.3040	(0.56)	–0.3249	(0.55)
EDUCATION	0.9550	(4.16)	2.8217	(8.56)	1.2495	(4.06)	1.1467	(3.11)
BUSINESS	0.0225	(0.05)	0.1622	(0.33)	0.0440	(0.08)	0.0563	(0.09)
GROWTH	0.3585	(1.31)	0.5008	(1.53)	0.4563	(1.26)	0.4936	(1.27)
INCOME	6.3171	(11.67)	6.8356	(10.32)	7.9562	(11.26)	8.5434	(12.07)
MULTIMARKET	0.1612	(1.01)	0.1845	(0.95)	0.1898	(1.11)	0.1991	(1.10)
REGULATED	0.3985	(0.79)	0.6236	(1.19)	0.5155	(0.85)	0.5583	(0.87)
AMERITECH	–1.1200	(2.20)	–2.0652	(3.65)	–1.4732	(2.34)	–1.6053	(2.26)
BELLATL	4.2097	(5.19)	4.2530	(4.97)	5.4100	(5.83)	5.7984	(5.43)
BELLSTH	–5.5745	(8.71)	–7.4564	(7.10)	–6.9358	(6.15)	–7.4081	(6.08)
CENTEL	–1.4865	(2.17)	–2.6263	(3.24)	–1.9823	(2.30)	–2.1662	(2.23)
CONTEL	–3.9163	(5.95)	–5.4874	(6.45)	–4.9587	(4.80)	–5.3267	(4.92)
GTE	–4.7547	(6.45)	–6.5454	(7.36)	–5.9863	(6.49)	–6.4134	(6.16)
NYNEX	3.8586	(6.46)	3.9321	(6.10)	4.9587	(6.25)	5.3103	(6.23)
PACTEL	–5.1267	(6.79)	–6.8463	(7.03)	–6.4992	(6.71)	–6.9933	(6.61)
SWBELL	–2.6569	(4.03)	–4.0878	(4.65)	–3.4077	(4.40)	–3.6738	(4.10)
USWEST	–0.2104	(0.38)	–1.2412	(1.78)	–0.3066	(0.38)	–0.3547	(0.42)
θ_1	4.0881	(15.96)	4.8538	(16.12)	5.1392	(15.11)	5.4930	(13.76)
Entrant								
CONSTANT	–37.0262	(9.49)	–37.9920	(13.00)	–40.3350	(14.56)	–40.8613	(13.50)
YEAR92	1.6869	(4.55)	1.7333	(4.07)	2.1986	(4.23)	2.3701	(3.88)
COMMUTING	3.9989	(8.76)	4.2782	(9.90)	5.1966	(11.58)	5.5911	(11.15)
POPULATION	0.1454	(0.52)	0.1357	(0.42)	0.1705	(0.40)	0.1765	(0.39)
EDUCATION	8.6876	(8.99)	9.0227	(12.07)	10.5322	(13.28)	11.2097	(12.96)
BUSINESS	–0.7294	(2.21)	–0.8009	(2.01)	–0.9178	(1.79)	–0.9756	(1.80)
GROWTH	–1.1227	(4.86)	–1.2127	(4.77)	–1.4088	(4.57)	–1.4990	(4.31)
INCOME	3.5168	(8.52)	4.1546	(10.67)	4.5070	(10.47)	4.8003	(9.68)
MULTIMARKET	0.2626	(2.05)	0.2807	(2.36)	0.3230	(2.62)	0.3416	(2.63)
REGULATED	1.6543	(4.69)	1.8414	(4.80)	2.1163	(4.63)	2.2683	(4.18)
AMERITECH	0.3271	(1.30)	0.1543	(0.55)	0.4482	(1.20)	0.4862	(1.17)
BELLATL	–0.2594	(0.71)	–0.5497	(1.44)	–0.3175	(0.67)	–0.3443	(0.62)
BELLSTH	–2.8575	(7.16)	–3.3553	(9.05)	–3.5567	(7.18)	–3.7884	(6.63)
CENTEL	2.8025	(10.23)	2.8240	(9.32)	3.5896	(8.31)	3.8329	(8.31)
CONTEL	–2.2878	(6.21)	–2.6606	(7.10)	–2.7112	(6.86)	–2.8522	(6.24)
GTE	–1.0714	(2.53)	–1.4739	(3.50)	–1.3747	(2.66)	–1.4788	(2.48)
MCCAW	1.8056	(4.49)	1.9977	(3.57)	2.3533	(4.59)	2.5311	(4.45)
NYNEX	2.0863	(3.42)	2.0475	(4.76)	2.6469	(5.13)	2.8226	(4.29)
PACTEL	–0.6136	(3.22)	–0.8522	(2.79)	–0.7577	(2.55)	–0.8086	(2.32)
SWBELL	–0.5418	(1.86)	–0.8090	(2.62)	–0.6476	(1.78)	–0.6919	(1.81)
USWEST	1.4736	(6.01)	1.4076	(4.80)	1.9981	(3.90)	2.1531	(3.82)
θ_2	2.4897	(11.87)	2.7692	(17.15)	3.1733	(16.74)	3.3932	(14.55)
ρ	0.0108	(1.83)	0.0105	(1.97)	0.0113	(2.46)	0.0109	(2.42)
$-\ln \mathcal{L}$	1,769.68		1,769.70		1,764.26		1,763.45	

Marginal effects evaluated at the sample mean of regressors. Endogenous variables are the number of tariff options of of each competing firm. The absolute value of a 2,000 replication, scaled, 10-step, bootsrapped t-statistics are reported between parentheses.