Testing for Complementarities Among Countable Strategies*

Eugenio J. Miravete†

September 19, 2011

Abstract

I present a multivariate count data regression model that is suitable to test for the existence of strategic complementarities when firms make use of countable strategies. The estimator accommodates both over and underdispersion. It also allows for correlations of any sign among counts independently of the dispersion parameters. I apply the model to address whether the pricing strategies of competing duopolists in the early U.S. cellular telephone industry can be considered strategic complements or substitutes.

Keywords: Strategic Complementarity; Number of Tariff Options; Multivariate Count Data Models; Sarmanov Distributions.

JEL Codes: C16, C35, L11

---

* Thanks are due to Jason Abrevaya, Stephen Donald, and Jerry Hausman for their comments. A previous version was presented at the Texas Camp Econometrics XIV in New Braunfels.

†The University of Texas at Austin, Department of Economics, 1 University Station C3100, Austin, TX 78712; and CEPR, London, UK. Phone: 512-232-1718. Fax: 512-471-3510. E–mail: *miravete@eco.utexas.edu*; http://www.eugeniomiravete.com

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 counts 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), moment-based estimation methods, *e.g.*, Gourieroux, Monfort, and Trognon (1984), and more recently copulas, in particular in the field of financial econometrics, *e.g.*, see Bien, Nolte, and Pohlmeier (2011), Heinen and Rengifo (2007), or Nolte (2008). However, most models available are difficult to extend beyond the bivariate case, and they fail to consider any effect of unobserved heterogeneity other than the well known overdispersion of the distribution of counts.[2] Furthermore, correlation between counts is generally assumed to depend on the same parameter identifying overdispersion effects, therefore restricting it to be *necessarily positive* as the same source of unobserved heterogeneity must explain simultaneously overdispersion and correlation.

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

[1] Windmeijer and Santos-Silva (1997) successfully deal with an endogenous count and an endogenous continuous variable while Hausman, Leonard, and McFadden (1995) address the case of an endogenous count and an endogenous dichotomous variable.

[2] 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 by the data. 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.

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 a 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* both the overdispersion of the marginal distributions of counts and their correlation.[3] In this paper I make use of the Sarmanov family of distributions with double Poisson marginals to build a multivariate count data regression model that is flexible in the sense that it can accommodate both over and underdispersion independently of any correlation patterns among the counts.[4]

The Sarmanov count data regression model introduced in this paper has some remarkable features that overcome most difficulties in extending the existing single-dimensional count data 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

[3] 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 still *necessarily positive*.

[4] The double Poisson introduced by Efron (1986) is one of the few discrete univariate distributions that can accommodate both over and underdispersion. Winkelmann (1995) builds a similarly flexible unidimensional model but based on the continuous gamma distribution of latent 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/underdispersion of the distribution of events that take place within an arbitrarily defined time interval.

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.[5] Third, dispersion and correlation 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 be extended beyond the bivariate case. Finally, 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 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.

In this paper I study the pricing strategies of competing duopolists in the early U.S. cellular telephone industry in order to evaluate whether the number of tariff options offered by competing firms are strategic complements and the pricing practices of firms corresponds to a supermodular game, *e.g.*, Topkis (1998, §4) and Vives (1990), or if alternatively, firms use their pricing offerings to differentiate themselves in attracting customers, *e.g.*, Yang and Ye (2008).[6]

---

[5] 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 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 model presented in this paper can address both over and underdispersion *separately* from correlation.

[6] The need to estimate joint demands of countable products or services arise in many environments 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).

Back in the 1980s, cellular carriers implemented nonlinear tariffs by means of a menu of self-selecting tariff options. In the absence of strategic considerations, carriers offered the number of tariff plans that was optimal to screen a customer base with some degree of heterogeneity while compensating for the costs of design and commercialization. Alternatively, 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. A 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. A negative correlation arises if firms use the number of tariff plans as a device to segment markets and differentiate themselves from each other. However, results do not favor such interpretation.

The paper is organized as follows. Section 2 presents the bivariate count data regression model based on a bivariate and multivariate Sarmanov distributions with double Poisson marginals. This section also describes the properties of the Sarmanov family of distributions in reference to the proposed model with specific double Poisson marginal frequencies. Section 3 discusses the estimation of this model. Section 4 estimates the two bivariate specification of the double Poisson-Sarmanov model to study the determinants of the number of tariff options offered by competing cellular telephone carriers in the U.S. during the mid-1980s. Section 5 concludes.

2 A Bivariate Double Poisson-Sarmanov Count Data Model

Consider a sample of size n of a K-variate process. Let $y_k = 0, 1, 2, \ldots$ be distributed as a double Poisson distribution with parameters μ_k and θ_k, conditional on a set of regressors $\mathbf{x_k}$ in a sample with $i = 1, 2, \ldots, n$ observations. After studying the properties of the double exponential family of distributions Efron (1986) shows that the probability frequency function of a double Poisson distribution is:

$$
\begin{aligned}
\tilde{f}_k(y_k|\mu_k,\theta_k) &= c(\mu_k,\theta_k) f_k(y_k|\mu_k,\theta_k)\,, && \text{(1a)}\\
f_k(y_k|\mu_k,\theta_k) &= \sqrt{\theta_k}\exp(-\theta_k\mu_k)\exp(-y_k)\frac{{y_k}^{y_k}}{y_k!}\left(\frac{e\mu_k}{y_k}\right)^{\theta_k y_k}\,, && \text{(1b)}\\
\frac{1}{c(\mu_k,\theta_k)} &= \sum_{y_k=0}^{\infty} f_k(y_k|\mu_k,\theta_k) \simeq 1+\frac{1-\theta_k}{12\theta_k\mu_k}\left(1+\frac{1}{\theta_k\mu_k}\right)\,, && \text{(1c)}
\end{aligned}
$$

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 $\mathbf{x_k}$, the expected count corresponding to observation i and its variance are:[7]

$$
\begin{aligned}
\mathrm{E}[y_k|\mathbf{x_k}] &\simeq \mu_k\,, && \text{(2a)}\\
\sigma_k^2=\mathrm{Var}[y_k|\mathbf{x_k}] &\simeq \frac{\mu_k}{\theta_k}\,. && \text{(2b)}
\end{aligned}
$$

Hence, the double Poisson includes the standard Poisson as a particular case when $\theta_k=1$ but it allows for overdispersion if $\theta_k<1$ 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:

$$
\mu_k=\exp\left(\mathbf{x_k'}\beta_\mathbf{k}\right)\,. \tag{3}
$$

[7] Efron (1986, Table 2) explores the discrepancy between $\tilde{f}_k(y_k|\mu_k,\theta_k)$ and $f_k(y_k|\mu_k,\theta_k)$ as a function of parameters θ_k and μ_k. For the application of Section 4 the approximate probability is just 1.17% larger than the exact probability frequency function at the estimated value of the parameters. Notice however that all relations below make use of the exact double Poisson density.

Using Stirling's formula $z! \simeq \sqrt{2\pi z} \cdot z^z \cdot \exp(-z)$ repeatedly for $z = y_k$ and $z = \theta_k y_k$, the frequency function (1b) is accurately approximated by:[8]

$$f_k(y_k|\mu_k, \theta_k) \simeq \theta_k \exp(-\theta_k \mu_k) \frac{(\theta_k \mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k + 1)} \,. \tag{4}$$

Sarmanov (1966) introduced a family of flexible bivariate distributions with given marginals. The double Poisson-Sarmanov count data regression models assumes that the marginal distributions are double Poisson. The bivariate probability frequency function takes the following form:

$$f_{12}(y_1, y_2) = f_1(y_1) f_2(y_2) \times [1 + \omega_{12} \psi_1(y_1) \psi_2(y_2)] \,, \tag{5}$$

where $f_k(y_k)$ for $k = 1, 2$ corresponds to the double Poisson marginal frequency (4). For the case of positive counts, the mixing functions $\psi_k(y_k)$ are the bounded and nonconstant functions:[9]

$$\sum_{y_k=0}^{\infty} \psi_k(y_k) f_k(y_k) dy_k = 0 \,. \tag{6}$$

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

$$\psi_k(y_k) = \exp(-y_k) - L_k(1), \qquad \forall y_k \geq 0 \,, \tag{7}$$

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

[8] The use of Stirling's formula is useful for practical purposes to ensure the stability of estimation for large counts (not an issue in the application of this paper). Moreover, using it twice for $z = y_k$ and $z = \theta_k y_k$ simplifies ${y_k}^{y_k}/y_k!$ in (1b) and helps showing the convergence of the infinite sums of the double Poisson-Sarmanov model in Section 3.2.

[9] In the literature on copulas mixing functions $\psi_k(y_k)$ are known as copula generators. See Fisher and Klein (2007, §2) and Nelsen (2006).

$$L_k(\zeta) = \sum_{y_k=0}^{\infty} \exp(-\zeta y_k) f_k(y_k) dy_k \,, \quad \text{at} \;\; \zeta = 1 \,. \tag{8}$$

Substituting the marginal frequency function (4) we obtain an approximation to the Laplace transform of the double Poisson distribution, which evaluated at $\zeta = 1$ becomes:

$$L_k(1|\mu_k, \theta_k) \simeq c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} \,. \tag{9}$$

Thus, according to (7), the mixing function of the double Poisson-Sarmanov distribution is:

$$\psi_k(y_k|\mu_k, \theta_k) \simeq \exp(-y_k) - c(\mu_k, \theta_k)\theta_k \exp(-\theta_k \mu_k) \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} \,. \tag{10}$$

A common theme in the literature on copula functions is the maximum range of variation of correlation. For expression (5) to properly define a bivariate density function the value of ω_{12} needs to fulfill the following constraint:

$$\omega_{12} \in \mathbb{R} : \; 1 + \omega_{12}\psi_1(y_1)\psi_2(y_2) \geq 0 \qquad \forall y_1, y_2 \,. \tag{11}$$

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 not only a wider range for correlation coefficients, but also the possibility that correlation is negative, a rare feature of copula functions, *e.g.*, Joe (1997, 5.4). Indeed, Lee (1996) shows that restriction (11) holds when ω_{12} falls within the following bounds:

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

which needs to be fulfilled for *every* observation in the sample.

Next, it is useful to define the following mixing function weighted mean:

$$\begin{aligned}\nu_k(\mu_k,\theta_k) = &\sum_{y_k=0}^{\infty} y_k\psi_k(y_k)f_k(y_k)dy_k = -L'_k(1|\mu_k,\theta_k) - L_k(1|\mu_k,\theta_k)\mu_k \\ &\simeq c(\mu_k,\theta_k)\theta_k\exp(-\theta_k\mu_k)\sum_{y_k=0}^{\infty}\frac{(\theta_k\mu_k)^{\theta_k y_k}\exp(-y_k)}{\Gamma(\theta_k y_k+1)}(y_k-\mu_k)\ ,\end{aligned} \tag{13}$$

where $L'_k(1|\mu_k,\theta_k)$ denotes the value of the derivative of the Laplace transform (9) of the double Poisson frequency evaluated at $\zeta = 1$, and where $\Gamma(\theta_k y_k + 1)$ is the gamma function:

$$\Gamma(\theta_k y_k+1) = \int_0^{\infty}\eta^{\theta_k y_k}\exp(-\eta)d\eta\,. \tag{14}$$

Notice that integrating the product y_1y_2 with respect to (5) and making use of (13), the product moment can be written as:

$$E\left[y_1y_2\right] = \mu_1\mu_2 + \omega_{12}\nu_1\nu_2\,, \tag{15}$$

so that the correlation coefficient of a well defined double Poisson-Sarmanov distribution is:

$$\begin{aligned}\rho_{12} = \frac{\omega_{12}\nu_1\nu_2}{\sigma_1\sigma_2} &\simeq \omega_{12}\prod_{k=1}^{2}\left\{\frac{c(\mu_k,\theta_k)\theta_k\exp(-\theta_k\mu_k)}{\sqrt{\mu_k/\theta_k}}\sum_{y_k=0}^{\infty}\frac{(\theta_k\mu_i)^{\theta_k y_k}\exp(-y_k)}{\Gamma(\theta_k y_k+1)}(y_k-\mu_k)\right\} \\ &= \omega_{12}\prod_{k=1}^{2}Q(\mu_k,\theta_k)\,.\end{aligned} \tag{16}$$

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

Combining all these elements into (5) we obtain the probability of observing simultaneously a pair of counts $\{y_1, y_2\}$ generated by the double Poisson-Sarmanov distribution:

$$f_{12}(y_1,y_2) \simeq \left(\prod_{k=1}^{2}\left\{c(\mu_k,\theta_k)\theta_k\exp(-\theta_k\mu_k)\frac{(\theta_k\mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k+1)}\right\}\right)\times\left(1+\rho_{12}\prod_{m=1}^{2}\frac{S(\mu_m,\theta_m)}{Q(\mu_m,\theta_m)}\right)\ , \tag{17}$$

where:

$$S(\mu_m, \theta_m) = \exp(-y_m) - c(\mu_m, \theta_m)\theta_m \exp(-\theta_m \mu_m) \sum_{y_m=0}^{\infty} \frac{(\theta_m \mu_m)^{\theta_m y_m} \exp(-y_m)}{\Gamma(\theta_m y_m + 1)} \,. \tag{18}$$

In addition, for this double Poisson-Sarmanov to be coherent and properly define a bivariate probability frequency function, the constraint corresponding to the general case (12) needs to hold. Making use of (9) and (16), the constraint can be written in terms of ρ_{12} as follows:

$$\underline{\omega}_{12} \prod_{k=1}^{2} Q(\mu_k, \theta_k) \ \leq \rho_{12} \leq \ \overline{\omega}_{12} \prod_{k=1}^{2} Q(\mu_k, \theta_k)\,, \quad \forall\ i. \tag{19}$$

2.1 Multivariate Extension

Multivariate extensions of this model are of clear interest for practical purposes. Lee (1996, §8) suggests a generalization of the joint density function of a multivariate Sarmanov distribution that accounts for higher order correlation among counts:

$$f_{1,2,\ldots,K}(y_1, \ldots, y_K) = \left[\prod_{k=1}^{K} f_k(y_k)\right] \times \left[1 + R_{\psi_1,\ldots,\psi_K,\mathbf{\Omega_K}}(y_1, \ldots, y_K)\right] \,, \tag{20}$$

where correlations must fulfill the following condition:

$$\begin{aligned} 1 + R_{\psi_1,\ldots,\psi_K,\mathbf{\Omega_K}}(y_1, \ldots, y_K) &= 1 + \sum\sum_{1\leq l_1 \leq l_2 \leq K} \omega_{l_1 l_2} \prod_{m=1}^{2} \psi_{l_m}(y_{l_m}) \\ + \sum\sum\sum_{1\leq l_1 \leq l_2 \leq l_3 \leq K} \omega_{l_1 l_2 l_3} \prod_{m=1}^{3} \psi_{l_m}(y_{l_m}) &+ \cdots + \omega_{12\ldots K} \prod_{m=1}^{K} \psi_m(y_m) \geq 0, \quad \forall y_1, \ldots, y_K \,. \end{aligned} \tag{21}$$

Extending the double Poisson-Sarmanov to more than two dimensions reduces to repeating the bivariate analysis of this section and substituting the probability frequency function (4) and mixing function (10) into (20) and (21):

$$
f_{1,2,\dots,K}(y_1,\dots,y_K) \simeq \left[\prod_{k=1}^{K}\left\{c(\mu_k,\theta_k)\theta_k \exp(-\theta_k\mu_k)\frac{(\theta_k\mu_k)^{\theta_k y_k}}{\Gamma(\theta_k y_k+1)}\right\}\right] \times
$$

$$
\left[1 + \sum\sum_{1\le l_1\le l_2\le K} \rho_{l_1 l_2} \prod_{m=1}^{2} \frac{S(\mu_{l_m},\theta_{l_m})}{Q(\mu_{l_m},\theta_{l_m})} + \sum\sum\sum_{1\le l_1\le l_2\le l_3\le K} \rho_{l_1 l_2 l_3} \prod_{m=1}^{3} \frac{S(\mu_{l_m},\theta_{l_m})}{Q(\mu_{l_m},\theta_{l_m})} \right. \tag{22}
$$

$$
\left. + \cdots + \rho_{12\dots K} \prod_{m=1}^{K} \frac{S(\mu_{l_m},\theta_{l_m})}{Q(\mu_m,\theta_m)}\right],
$$

where, similarly to equation (16), higher order correlation coefficients are given by:

$$
\rho_{12\dots K} = \omega_{12\dots K} \prod_{k=1}^{K}\left(\frac{\nu_k}{\sigma_k}\right), \tag{23}
$$

which in turn makes use of the product moment:

$$
E\left[y_1 y_2 \dots y_K\right] = \prod_{k=1}^{K}\mu_k + \omega_{12\dots K}\prod_{k=1}^{K}\nu_k\,. \tag{24}
$$

3 Estimation

Despite the apparently cumbersome notation, estimation of the proposed model is relatively straightforward. Take for instance the scalar realizations y_{1i} and y_{2i} of two count random variables given two vectors of regressors $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$, parameter vectors $\gamma_{\mathbf{1}}$ and $\gamma_{\mathbf{2}}$, as well as parameter scalar ω_{12}. Estimation by maximum likelihood maximizes the probability of jointly observing $\{y_{11},y_{21}\},\{y_{12},y_{22}\},\dots,\{y_{1n},y_{2n}\}$ in an n-size sample. Using the general bivariate Sarmanov distribution (5), the log-likelihood function can be written as:

$$
\mathcal{L}\left(\gamma_{\mathbf{1}},\gamma_{\mathbf{2}},\omega_{12}\right) = \sum_{i=1}^{n}\sum_{k=1}^{2}\ln f_k\left(y_{ki}|\mathbf{x_{ki}},\gamma_{\mathbf{ki}}\right) + \sum_{i=1}^{n}\ln\left[1+\omega_{12}\prod_{k=1}^{2}\psi_k\left(y_{ki}|\mathbf{x_{ki}},\gamma_{\mathbf{ki}}\right)\right]. \tag{25}
$$

Notice that ω_{12} only enters the term between brackets. The estimation thus proceeds iteratively, alternatively fixing the value of ω_{12} or $\gamma_{\mathbf{1}}$ and $\gamma_{\mathbf{2}}$ until we achieve convergence. Initial values $\hat{\gamma}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\gamma}_{\mathbf{2}}^{(\mathbf{0})}$ 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 (25) over the interval defined by the constraint (12) while holding the estimated $\hat{\gamma}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\gamma}_{\mathbf{2}}^{(\mathbf{0})}$ constant. With this new value of $\hat{\omega}_{12}^{(0)}$, new estimates $\hat{\gamma}_{\mathbf{1}}^{(\mathbf{1})}$ and $\gamma_{\mathbf{2}}^{(\mathbf{1})}$ are obtained by maximizing (25) while holding ω_{12} constant at the estimated value $\hat{\omega}_{12}^{(0)}$. The process is repeated until convergence is achieved.[10]

Estimating a trivariate or multivariate model is slightly more convoluted because in principle equation (21) 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 $\{y_1, y_2, \ldots, y_K\}$ are jointly distributed according to a K-variate Sarmanov distribution, then any subset of $\{y_1, y_2, \ldots, y_K\}$ 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 an n-size sample is:

$$\begin{aligned}\mathcal{L}\left(\gamma_{\mathbf{1}}, \gamma_{\mathbf{2}}, \gamma_{\mathbf{3}}, \omega_{12}, \omega_{13}, \omega_{23}, \omega_{123}\right) &= \sum_{i=1}^{n}\sum_{k=1}^{3} \ln f_k\left(y_{ki}|\mathbf{x_{ki}}, \gamma_{\mathbf{ki}}\right) \\ &+ \sum_{i=1}^{n} \ln\left[1 + \sum\sum_{1\le l_1 \le l_2 \le 3} \omega_{l_1 l_2} \prod_{m=1}^{2} \psi_{l_m}\left(y_{l_m i}|\mathbf{x_{l_m i}}, \gamma_{\mathbf{l_m i}}\right) + \omega_{123} \prod_{k=1}^{3} \psi_k\left(y_{ki}|\mathbf{x_{ki}}, \gamma_{\mathbf{ki}}\right)\right].\end{aligned} \tag{26}$$

Under the assumption of independence, single dimensional count data regressions produce initial estimates for $\gamma_{\mathbf{1}}$, $\gamma_{\mathbf{2}}$, and $\gamma_{\mathbf{3}}$. Conditioning on $\hat{\gamma}_{\mathbf{1}}^{(\mathbf{0})}$ and $\hat{\gamma}_{\mathbf{2}}^{(\mathbf{0})}$ in (25) 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 (25) on $\{\hat{\gamma}_{\mathbf{1}}^{(\mathbf{0})}, \hat{\gamma}_{\mathbf{3}}^{(\mathbf{0})}\}$ and

[10] The GAUSS code used for the estimation of the bivariate model of Section 4 is available upon request.

$\{\hat{\gamma}_{\mathbf{2}}^{(\mathbf{0})}, \hat{\gamma}_{\mathbf{3}}^{(\mathbf{0})}\}$, respectively. Then, maximizing (26) produces an estimate of ω_{123} while holding all the other parameters constant. Once $\hat{\omega}_{123}^{(0)}$ has been obtained, new estimates $\gamma_{\mathbf{1}}^{(\mathbf{1})}$, $\gamma_{\mathbf{2}}^{(\mathbf{1})}$, and $\gamma_{\mathbf{3}}^{(\mathbf{1})}$ are estimated by maximizing (26) 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.

3.1 Inference

We first need to evaluate whether we can consistently estimate parameters that may lie on the boundary generically defined by condition (11). Notice that the Sarmanov model only imposes a constraint on the correlation coefficient that must be fulfilled by *every* observation in the sample, a condition that varies with the regressors considered in the estimation. All other parameters, although they define the range of variation of the correlation coefficient, remain unrestricted. Furthermore, the range defined by (12) 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.

Thus, in order to obtain consistent inference for these parameter estimates, we need to address the possibility that estimated parameters may lie on the boundary defined by the constraint (19). 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 (11) and (21) for the bivariate and multivariate case, respectively. Rather than computing common bootstrap standard errors Andrews (1999, §6.4) suggests the use of a rescaled bootstrap method in which bootstrap samples of size $b < n$ are employed.[11] Andrews

[11] 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).

(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. To speed up the process of obtaining robust inference I follow Andrews (2002) and compute a 10-step version of the rescaled bootstrap.

3.2 Econometric Implementation

Equation (18) includes the following sum:

$$S_k(\mu_k, \theta_k) = \sum_{y_k=0}^{\infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\Gamma(\theta_k y_k + 1)} \,. \tag{27}$$

We thus need to decide how many terms of the infinite sums in equations (9)-(17) to account for in the estimation. Notice that for $\theta_k = 1$, the series $S_k(\mu_k, \theta_k)$ converges to:

$$S_k(\mu_k, 1) = \sum_{y_k=0}^{\infty} \frac{[\mu_k \exp(-1)]^{y_k}}{y!} = \exp\left(\frac{\mu_k}{e}\right) , \tag{28}$$

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

$$\lim_{y_k \to \infty} \frac{(\theta_k \mu_k)^{\theta_k y_k} \exp(-y_k)}{\sqrt{2\pi}\sqrt{\theta_k y_k}(\theta_k y_k)^{\theta_k y_k} \exp(-\theta_k y_k)} = \lim_{y_k \to \infty} \frac{1}{\sqrt{2\pi\theta_k y_k}} \cdot \lim_{y_k \to \infty} \left(\frac{\mu_k}{y_k}\right)^{\theta_k y_k} \cdot \lim_{y_k \to \infty} [\exp(-y_k)]^{1-\theta_k} = 0 \,, \tag{29}$$

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 $S_k(\mu_k, \theta_k)$ varies greatly with θ_k. We can rewrite equation (27) as:

$$S_k(\mu_k, \theta_k) = \sum_{y_k=0}^{\infty} \frac{\left[(\theta_k \mu_k)^{\theta_k} \exp(-1)\right]^{y_k}}{y_k!} \frac{y_k!}{(\theta_k y_k)!} \,, \tag{30}$$

so that the number of elements of the sum in (27) needed to approximate $S_k(\mu_k, \theta_k)$ decreases with θ_k for any given precision level. Thus, longer series are needed to approximate $S_k(\mu_k, \theta_k)$ the more overdispersed the distribution of y_k is.

4 Number of Tariff Options in Duopoly Competition

At the beginning of the 1980s, technology was a barrier for competition in cellular telephony, essentially because of the large size of bandwidth needed for transmission and the scarce radio spectrum available. To solve this problem the Federal Communications Commission (*FCC*) divided the U.S. into 305 non-overlapping markets corresponding to the Standard Metropolitan Statistical Areas (*SMSA*s). In 1981, the FCC set aside 50 MHz of spectrum in the 800 MHz band for cellular services. One of the two cellular channel blocks in each market —the B block or *wireline* license— was awarded to a local incumbent carrier, while the A block —the *nonwireline* license— was awarded by comparative hearing to an entrant carrier other than a local *wireline* incumbent. After awarding the first thirty *SMSA* licenses by means of this expensive and time consuming approach, rules were adopted in 1984 and 1986 to award the remaining *nonwireline* licenses through lotteries. Depending on the market, there were between 6 and 579 contenders for a single *nonwireline* license. The administrative decision to award the second license to one out of hundreds of applicants was customarily contested in court in a process that took several years. As the licenses were finally awarded, entrant firms had six month to be fully operative, something that was facilitated by the *FCC* requirement hat the incumbent had to share its installed base of antennae with the entrant in this early stage of the market in order to promote competition.[12]

This section studies whether there were strategic consideration in designing the pricing strategies of local duopolists in local cellular telephone markets in the early U.S. cellular telephone

[12] For an institutional and historical account of the poorly designed awarding process of licenses in the early U.S. cellular telephone industry see Hausman (2002), Parker and Röller (1997), or Murray (2002).

Table 1: 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.

industry. Tariffs in the early U.S. cellular industry were quite simple. A tariff option was normally a three-part tariff consisting of an allowance of "free" minutes per month, a fixed monthly fee, and a fixed rate per minute. Tariff options normally distinguished between peak (comprising on average about 13 hours a day at that time) and off-peak marginal rates.[13] I will thus focus on the total number of tariffs offered by the competing firms as strategic choice variable to sign up new customers with heteregoneous calling needs. Data available contain a complete description of the tariff options offered by any of the two firms present in the 100 largest markets of the U.S. between 1984 and

[13] Other value added services such as detailed billing, call waiting, no-answer transfer, call forwarding, three way calling, busy transfer, call restriction, and voice mail were priced independently and rarely bundled together with particular tariff options.

1988. 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, Marciano (2000) combined information for the same carriers from *Cellular Directions, Inc.*, the *Cellular Telephone Industry Association*, and direct interviews with managers.

Table 1 presents the market and firm specific characteristics used in the estimation.[14] 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.[15] 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 compete against each other.[16] The latter is a dummy variable that indicates whether new tariffs need to be approved by the regulator.[17] In order to control for potential firm effects, I also include firm dummies to identify the largest shareholder of each cellular carrier (available from the *FCC*). Only those carriers with

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

[15] 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.

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

[17] Shew (1994) claims 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 2: 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.

at least 4% of licenses in this sample are identified.[18] Lastly, YEAR92 identifies those observations from 1992, when arguably, the cellular market had matured.

Table 2 presents the marginal distribution of the total number of tariffs offered by incumbent and entrant carriers. Notice that incumbents only 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. This is consistent with the prediction of theoretical models of nonlinear pricing competition as markets mature and most potential customers have already signed up for one of the two carriers, *e.g.*, Armstrong and Vickers (2001) and Rochet and Stole (2002). 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. These features, and in particular the low number of telephone options by two competing firms, make the Sarmanov model of this paper suitable to be used in the estimation of the determinants of the number of plans offered by different cellular carriers.

[18] 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).

Table 3: Correlation Among Number of Tariff Options

	1984–1988							*1992*						
Plans	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.

Table 3 presents the bivariate frequency of each combination of the number of tariff options offered by incumbent and entrant carriers. Simple unconditional association measures indicate that the number of tariff options appear to be strategic complements when we measure the association between these strategies regardless 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.

4.1 Results

Table 4 presents the results of the estimation of the bivariate double-Poisson Sarmanov count data regressions model. Estimates capture the fact that the distribution of the number of tariffs are positively correlated and underdispersed. Table 4 also reports the estimates of the corresponding,

Table 4: Double Poisson – Sarmanov Regression

Variables	*Independent Regressions*				*Sarmanov Regression*					
	Incumbent		*Entrant*		*Incumbent*			*Entrant*		
CONSTANT	2.3986	(0.47)	–12.5831	(1.79)	2.3967	[25.87]	{52.93}	–12.6123	[27.21]	{61.25}
YEAR92	0.7212	(6.87)	0.6151	(4.18)	0.7115	[2.69]	{4.06}	0.6216	[2.98]	{3.68}
COMMUTING	–1.1548	(1.92)	1.5559	(2.06)	–1.1798	[11.88]	{14.25}	1.5674	[13.16]	{17.57}
POPULATION	–0.0489	(0.40)	0.0743	(0.59)	–0.0475	[0.41]	{0.54}	0.0652	[0.48]	{0.62}
EDUCATION	0.1227	(0.06)	2.8608	(0.97)	0.1284	[1.66]	{1.91}	2.8691	[20.19]	{36.52}
BUSINESS	0.0333	(0.26)	–0.2681	(2.01)	0.0237	[0.19]	{0.24}	–0.2694	[1.85]	{2.12}
GROWTH	0.0891	(1.51)	–0.4534	(6.76)	0.0874	[0.98]	{1.55}	–0.4500	[4.51]	{5.99}
INCOME	1.4633	(2.32)	1.3248	(1.64)	1.4941	[14.12]	{18.69}	1.3347	[13.57]	{18.36}
MULTIMARKET	0.0409	(1.80)	0.1082	(4.06)	0.0394	[0.80]	{1.60}	0.1037	[1.71]	{3.62}
REGULATED	0.0928	(0.87)	0.6520	(4.66)	0.0815	[0.54]	{0.76}	0.6342	[4.31]	{5.95}
AMERITECH	–0.2183	(0.87)	0.3169	(0.63)	–0.2299	[1.76]	{2.25}	0.2406	[1.80]	{2.07}
BELLATL	1.0770	(4.97)	0.1317	(0.31)	1.0957	[5.78]	{7.00}	0.0848	[0.54]	{0.62}
BELLSTH	–1.2825	(6.09)	–0.9200	(2.07)	–1.2362	[5.91]	{7.13}	–0.9414	[5.77]	{6.67}
CENTEL	–0.2719	(1.26)	1.3981	(2.94)	–0.2827	[1.58]	{2.02}	1.3036	[8.43]	{10.79}
CONTEL	–0.8500	(3.70)	–0.7116	(1.42)	–0.8524	[3.71]	{4.63}	–0.7340	[5.68]	{6.51}
GTE	–1.1022	(6.38)	–0.1997	(0.52)	–1.0929	[5.81]	{7.55}	–0.2429	[1.41]	{1.47}
MCCAW			0.8311	(2.76)				0.8508	[4.67]	{5.62}
NYNEX	0.9543	(5.30)	0.9591	(2.37)	0.9632	[5.53]	{6.32}	0.8880	[4.76]	{4.74}
PACTEL	–1.2295	(4.07)	–0.0734	(0.11)	–1.1967	[6.29]	{6.71}	–0.0748	[0.76]	{0.91}
SWBELL	–0.5886	(2.53)	0.0341	(0.06)	–0.5839	[3.49]	{4.55}	–0.0037	[0.03]	{0.04}
USWEST	–0.0150	(0.08)	0.7996	(1.69)	–0.0048	[0.03]	{0.03}	0.7491	[4.27]	{5.60}
θ	3.6324	(16.37)	2.3895	(15.24)	3.6585	[15.60]	{17.79}	2.4113	[13.82]	{18.33}
ρ						0.0396		[3.46]	{5.17}	
$–\ln \mathcal{L}$	830.17		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. Absolute value t-statistics computed using a 2,000 replication, 10-step bootstrapping. Heteroskedastic-consistent t-statistics are reported between parentheses. For the Sarmanov model t-statistics reported between brackets make use of rescaled bootstrapping while those between curly brackets make use of standard bootstrapping, respectively.

restricted, independent, count data regression models. Table 4 shows that marginal effects are very similar. However the estimation of the correlation coefficient ρ is significant and the specification with independent count regression is rejected in favor of the bivariate double Poisson-Sarmanov model (likelihood ratio test of 12.12, 0.001 p-value). Since in addition of the sample considered, the effective range of the correlation coefficient is partially determined by the regressors included in the exponential mean function (3).

The estimate of correlation between the number of tariff options offered by competing firms in Table 4 is small but positive and significant.[19] Such result supports the view that the number of tariff options offered by competing cellular carriers are strategic complements.

Continuing with the effects of firm and market characteristics on the pricing decisions of cellular carriers, estimates show that ownership fixed effects are generally significant and indicate that cellphone carriers offer more tariff options in those markets where they act as new entrants relative to those markets where they are the incumbents. Many market characteristics 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.

[19] Notice that I have computed t-statistics based on two bootstrapping methods to evaluate whether the inference is sensible to dealing with the existence of bounds in the estimation given by the constraint (16). There are very few differences but in interpreting the results, I focus on the inference obtained making use of rescaled boostrapping, *i.e.*, t-statistics between brackets. This criteria is more conservative regarding the significance of parameters and also the appropriate one to deal with the existence of constraints involving the parameters estimate. Every estimation of the scaled bootstrap is run on a sample of size 1/3 of the full sample and where the same observation may be present multiple times. I repeated the analysis without any regressors. While correlation becomes not significant the simpler specification is rejected in favor of that of Table 4.

4.2 Interpreting the Sign of the Correlation Coefficient

The observed co-movements in the total number of plans offered by competing firms may respond not only to strategic interactions between firms who want to match competitors' practices (captured by ϑ_k above) but also by their attempt to extract as much surplus as possible from consumers' willingness to pay for cellular service (the effect of ϵ above) in a competitive environment.

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 in 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 to how tariffs are commonly implemented, *i.e.*, through a menu of tariff options.

The use of few tariff options to screen consumer might be due to the existence of 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. If this was 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 between counts 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 the 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. This is the scenario supported by results of Table 4. On the contrary, if correlation among the number of tariff options offered is negative, firms might be attempting 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. Results do not support this view, which is intuitive given the very early stage of development of the cellular industry during the 1980s.

5 Concluding Remarks

The Sarmanov count data model 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.

Results from analyzing the pricing strategies of cellular carriers in the U.S. during the mid-1980s indicates that the number of tariff options offered by these firms can be considered strategic complements and, as postulated by Wilson (1993, §8.3), that nonlinear pricing competition is mostly driven by how heterogeneous the valuations of consumers are as well as by an attempt to match the competitor pricing practices to avoid loosing early subscribers to the cellular service.

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.

Bien, K., I. Nolte, and W. Pohlmeier (2011): "An Inflated Multivariate Integer Count Hurdle Model: An Application to Bid and Ask Quote Dynamics." *Journal of Empirical Finance*, 26, 669–707.

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.

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

Fisher, M. and I. Klein (2007): "Constructing Generalized FGM Copulas by Means of Certain Univariate Distributions." *Metrika*, 65, 243–260.

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. (2002): "Mobile Telephone." In M. Cave, S. K. Majumdar, and I. Vogelsang (eds.): *Handbook of Telecommunications Economics,* Vol. I. Amsterdam, The Netherlands: North-Holland.

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.

Heinen, A. and E. Rengifo (2007): "Multivariate Autoregressive Modeling of Time Series Count Data Using Copulas." *Journal of Empirical Finance*, 14, 564–583.

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

Joe, H. (1997): *Multivariate Models and Dependence Concepts.* New York, NY: Chapman & Hall/CRC.

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.

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

Murray, J. B., Jr (2002): *Wireless Nation. The Frenzied Launch of the Cellular Revolution in America.* Cambridge, MA: Perseus Publishing.

Nelsen, R. B. (2006): *An Introduction to Copulas*, 2nd edn. New York, NY: Springer-Verlag.

Nolte, I. (2008): "Modeling a Multivariate Transaction Process." *Journal of Financial Econometrics*, 6, 143–170.

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.

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.

Topkis, D. M. (1998): *Supermodularity and Complementarity.* Princeton, NJ: Princeton University Press.

Vives, X. (1990): "Nash Equilibrium with Strategic Complementarities." *Journal of Mathematical Economics*, 19, 305–321.

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.