Partially Schur-constant models

Anna Castañer[a], M. Mercè Claramunt[a], Claude Lefèvre[b], Stéphane Loisel[c,*]

[a]*Departament de Matemàtica Econòmica, Financera i Actuarial, Universitat de Barcelona, 690 Avinguda Diagonal, E-08034 Barcelona, Spain*
[b]*Département de Mathématique, Université Libre de Bruxelles, Campus de la Plaine C.P. 210, B-1050 Bruxelles, Belgium*
[c]*Univ Lyon, Université Lyon 1, ISFA, LSAF EA2429, 50, avenue Tony-Garnier, F-69007 Lyon, France*

Abstract

In this paper, we introduce a new multivariate dependence model that generalizes the standard Schur-constant model. The difference is that the random vector considered is partially exchangeable, instead of exchangeable, whence the term partially Schur-constant. Its advantage is to allow some heterogeneity of marginal distributions and a more flexible dependence structure, which broadens the scope of potential applications. We first show that the associated joint survival function is a monotonic multivariate function. Next, we derive two distributional representations that provide an intuitive understanding of the underlying dependence. Several other properties are obtained, including correlations within and between subvectors. As an illustration, we explain how such a model could be applied to risk management for insurance networks.

Keywords: Multivariate multiple monotonicity, Partially exchangeable vector, Risk management, Schur-constant model.
AMS 2010 *Subject Classifications:* 60G09, 91B30.

1. Introduction

Schur-constant models describe random lifetimes with a particular dependence. They are used in a variety of areas, including reliability, survival analysis, marketing, insurance and finance. Traditionally, the lifetimes are absolutely continuous random variables evaluated in $\mathbb{R}_+$. Properties of continuous Schur-constant models have been studied, e.g., in [3, 4, 7, 18, 20, 24].

Discrete versions of Schur-constant models have received little attention so far. However, they may be more appropriate in a variety of applications, particularly in reliability and lifetime studies; see, e.g., [19]. Recently, Castañer et al. [6] introduced Schur-constant models for vectors valued in $\mathbb{N}_0^n$, where $\mathbb{N}_0 = \{0, 1, \ldots\}$. Lefèvre et al. [15] showed for discrete vectors with different ranges that the associated partial sum process is a non-homogeneous Markov chain. Castañer and Claramunt [5] discussed the link with equilibrium distributions. Jones and Marchand [13] developed a generalized model based on a sum and share decomposition.

In the following, we will focus on the discrete framework for brevity and because the continuous case study is roughly similar. We recall that for any integer $n \geq 2$, a discrete random vector $\mathbf{X} = (X_1, \ldots, X_n)$ is Schur-constant on $\mathbb{N}_0^n$ if its joint survival function can be expressed as

$$\Pr(X_1 \geq x_1, \ldots, X_n \geq x_n) = S(x_1 + \cdots + x_n), \tag{1}$$

for all $x_1, \ldots, x_n \in \mathbb{N}_0$. Here, S is some admissible function $S(x) : \mathbb{N}_0 \to [0, 1]$, called generator of the model. From (1), such a random vector $\mathbf{X}$ is exchangeable [9]. Thus, the variables X_i have the same distribution and the covariances between any pair of variables (X_i, X_j) are equal. Clearly, this property is very restrictive, if not unrealistic, for a number of real-life situations. This is the case for example in the survival study of a given population.

---

*Corresponding author
Email addresses: `acastaner@ub.edu` (Anna Castañer), `mmclaramunt@ub.edu` (M. Mercè Claramunt), `clefevre@ulb.ac.be` (Claude Lefèvre), `stephane.loisel@univ-lyon1.fr` (Stéphane Loisel)

To overcome this difficulty, we follow de Finetti's idea and propose to construct a more general model incorporating a property of partial Schur constancy. Specifically, we partition the vector $\mathbf{X}$ into $m \geq 2$ groups, namely $\mathbf{X}_1 = (X_{1,1}, \ldots, X_{1,n_1})$ of size $n_1 \geq 1$, ..., $\mathbf{X}_m = (X_{m,1}, \ldots, X_{m,n_m})$ of size $n_m \geq 1$, with $n_1 + \cdots + n_m = n$. The model (1) is then generalized as follows.

Definition 1. *The discrete random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ is partially Schur-constant on $\mathbb{N}_0^n$ if its joint survival function is of the form*

$$\Pr(\mathbf{X}_1 \geq \mathbf{x}_1, \ldots, \mathbf{X}_m \geq \mathbf{x}_m) = S(|\mathbf{x}_1|, \ldots, |\mathbf{x}_m|), \tag{2}$$

for all $\mathbf{x}_j = (x_{j,1}, \ldots, x_{j,n_j}) \in \mathbb{N}_0^{n_j}$, where $|\mathbf{x}_j| = x_{j,1} + \cdots + x_{j,n_j}$ for all $j \in \{1, \ldots, m\}$. The function S is some admissible m-variates function $S(x_1, \ldots, x_m) : \mathbb{N}_0^m \rightarrow [0, 1]$, called a generator.

We note that the partially Schur-constant model (2) keeps the property of indifference relative to aging satisfied by the model (1). This means that any two vectors of residual lifetimes in the m groups, $(X_{1,k_1} - x_{1,k_1}, \ldots, X_{m,k_m} - x_{m,k_m})$ and $(X_{1,\ell_1} - x_{1,\ell_1}, \ldots, X_{m,\ell_m} - x_{m,\ell_m})$, have the same conditional distributions even if they have different ages. Indeed, we see from (2) that

$$\Pr(X_{1,k_1} - x_{1,k_1} \geq t_1, \ldots, X_{m,k_m} - x_{m,k_m} \geq t_m | \mathbf{X}_1 \geq \mathbf{x}_1, \ldots, \mathbf{X}_m \geq \mathbf{x}_m)$$
$$= S(|\mathbf{x}_1| + t_1, \ldots, |\mathbf{x}_m| + t_m) / S(|\mathbf{x}_1|, \ldots, |\mathbf{x}_m|)$$
$$= \Pr(X_{1,\ell_1} - x_{1,\ell_1} \geq t_1, \ldots, X_{m,\ell_m} - x_{m,\ell_m} \geq t_m | \mathbf{X}_1 \geq \mathbf{x}_1, \ldots, \mathbf{X}_m \geq \mathbf{x}_m).$$

By definition (2), the vector $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ is a particular case of partially exchangeable vectors [10]. Diaconis [11] and Aldous [1] present a complete analysis of the notion of partial exchangeability. Obviously, the model (2) constructed for a single group ($m = 1$) gives the exchangeable Schur-constant model (1). The advantage of (2) is that the previous symmetry in $\mathbf{X}$ is now partially broken because $\mathbf{X}$ is divided into m different symmetric groups. Thus, the marginal distributions in each group are equal but vary from one group to another. The covariances within each group are identical but the dependencies between groups are allowed to vary. As a result, the proposed model gains flexibility and greatly increases the scope of potential applications.

It should be mentioned that discrete Schur-constant models can be considered as analogs to the continuous models underlying Archimedean copulas [20]. In this context, the extension to partial Schur-constancy recalls, to some extent, Hierarchical Archimedean Copulas (HAC). This similarity is close but different, however, since the notion of hierarchy is not explicitly taken into account in modeling by (2). Besides, correlation inside groups is usually stronger than between members of different groups in HAC models, which is not always the case in partially Schur-constant models.

The paper is organized as follows. In Section 2, we show that the generator S is an m-dimensional function with a monotonicity degree of order $(n_1, \ldots, n_m)$. In Section 3, we derive different joint distributions and provide two partial Schur-constancy characterizations: the first by conditioning with respect to the sums of variables in the m groups and the second by using an m-dimensional doubly mixed multinomial distribution. In Section 4, we prove that if a model is partially Schur-constant for any order $(n_1, \ldots, n_m)$, all its submodels have an m-dimensional mixed geometric distribution. In Section 5, we determine the correlation coefficients within and between the m groups and illustrate with numerical examples. In Section 6, we use the second characterization of the model to discuss an application in the risk management of insurance networks.

2. Monotonicity of the generator

The function $S(x_1, \ldots, x_m)$, for $(x_1, \ldots, x_m) \in \mathbb{N}_0^m$, must satisfy certain conditions to generate a partially Schur-constant model (2). Let us begin with straightforward observations. The generator S is an m-dimensional survival function such that

$$S(x_1, \ldots, x_m) = \Pr(X_{1,i_1} \geq x_1, \ldots, X_{m,i_m} \geq x_m), \tag{3}$$

for all $i_1 \in \{1, \ldots, n_1\}, \ldots, i_m \in \{1, \ldots, n_m\}$. Each vector $\mathbf{X}_j$ is a Schur-constant model with

$$\Pr(\mathbf{X}_j \geq \mathbf{x}_j) = S(0, \ldots, 0, |\mathbf{x}_j|, 0, \ldots, 0),$$

i.e., its generator is $S_j(x) = S(0, \ldots, 0, x, 0, \ldots, 0)$, $x \in \mathbb{N}_0$. More generally, any subvector of $\mathbf{X}$ is partially Schur-constant with a generator obtained similarly from S.

For the Schur-constant model (1), Castañer et al. [6] proved that the generator S is a univariate survival function that is necessarily n-monotone. Recall that a function S is n-monotone for some integer $n \geq 1$ if, for all $x \in \mathbb{N}_0$, it satisfies the conditions

$$\forall_{k \in \{1, \ldots, n\}} \quad (-1)^k \Delta^k S(x) \geq 0, \tag{4}$$

where Δ is the forward difference operator, viz. $\Delta f(x) = f(x+1) - f(x)$, and Δ^k is its kth iterated.

We will generalize that result to a partially Schur-constant model by using the property of multivariate monotonicity. For $\mathbf{k} = (k_1, \ldots, k_m) \in \mathbb{N}_0^m$ and a function $S : \mathbb{N}_0^m \to \mathbb{R}$, we set $\Delta^{\mathbf{k}} S(\mathbf{x}) = \Delta_1^{k_1} \cdots \Delta_m^{k_m} S(\mathbf{x})$, where the operator $\Delta_j^{k_j}$ applies to the argument x_j. As usual, $\mathbf{0}$ ($\mathbf{1}$) denote vectors with all 0 (1)s, and $\mathbf{k} \leq \mathbf{l}$ means $k_j \leq \ell_j$ for all j while $\mathbf{k} \lneqq \mathbf{l}$ is the same with $k_j < \ell_j$ for at least one j.

Definition 2. *A function $S : \mathbb{N}_0^m \to \mathbb{R}$ is $\mathbf{n}$-monotone, $\mathbf{n} \in \mathbb{N}^m$ if, for all $\mathbf{x} \in \mathbb{N}_0^m$, it satisfies the conditions*

$$\forall_{\mathbf{0} \lneqq \mathbf{k} \leq \mathbf{n}} \quad (-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} S(\mathbf{x}) \geq 0. \tag{5}$$

This property generalizes, of course, the previous monotonicity (4). In the univariate case, the continuous multiple monotonicity is a classical notion with many applications; see, e.g., [2, 8, 17, 25]. A discrete version like (4) was used by, e.g., Lefèvre and Loisel [14] and Mai et al. [16] (with the references therein) for certain characterization problems. In the multivariate case, Ressel [21] studied the continuous multiple monotonicity and its properties. Recently, Shenkman [22] dealt with parameterization problems of multivariate distributions using a concept of discrete monotonicity such as (5).

Proposition 1. *A function $S : \mathbb{N}_0^m \to [0, 1]$ is the generator of a partially Schur-constant model if and only if $S(\mathbf{0}) = 1$, $S(\mathbf{x}) = 0$ when $x_j = \infty$ for some j and S is $\mathbf{n}$-monotone on $\mathbb{N}_0^m$.*

Proof. A function $\bar{F}(\mathbf{x}): \mathbb{N}_0^n \to [0, 1]$ is the survival function of a $\mathbb{N}_0^n$-valued vector $\mathbf{x}$ if and only if $\bar{F}(\mathbf{0}) = 1$, $\bar{F}(\mathbf{x}) = 0$ when $x_j = \infty$ for some j and

$$\forall_{\ell \in \{1, \ldots, n\}} \quad (-1)^\ell \Delta_{i_1} \cdots \Delta_{i_\ell} \bar{F}(x_{i_1}, \ldots, x_{i_\ell}) \geq 0. \tag{6}$$

This is shown, e.g., in Lemma 2.1 of [6]. Consider now a partial Schur-constant vector. The first two conditions give $S(\mathbf{0}) = 1$ and $S(\mathbf{x}) = 0$ when $x_j = \infty$ for some j. Let us examine the condition (6). From (2), we can write that

$$\bar{F}(x_{i_1}, \ldots, x_{i_\ell}) = S(|\mathbf{x}_1|, \ldots, |\mathbf{x}_m|),$$

where $|\mathbf{x}_j|$ is defined as before. The ℓ involved variables, $X_{i_1}, \ldots, X_{i_\ell}$, are distributed among the m groups. Let k_j be the number of variables in group $j \in \{1, \ldots, m\}$; of course, $\mathbf{0} \lneqq (k_1, \ldots, k_m) \leq \mathbf{n}$. We then see that

$$\Delta_{i_1} \cdots \Delta_{i_l} \bar{F}(x_{i_1}, \ldots, x_{i_l}) = \Delta_1^{k_1} \cdots \Delta_m^{k_m} S(|\mathbf{x}_1|, \ldots, |\mathbf{x}_m|). \tag{7}$$

Moreover, we notice that $\ell = k_1 + \cdots + k_m$. So, inserting (7) in (6) yields (5), i.e., S is $\mathbf{n}$-monotone. $\square$

In other words, an m-dimensional survival function S is a generator for the model (2) provided that it is an $\mathbf{n}$-monotone function. Let p be the probability mass function (pmf) associated to S, i.e., using (3),

$$p(x_1, \ldots, x_m) = \Pr(X_{1,i_1} = x_1, \ldots, X_{m,i_m} = x_m).$$

Corollary 1. *A function $p : \mathbb{N}_0^m \to [0, 1]$ is the pmf of a partially Schur-constant generator if and only if the $p(\mathbf{x})$s are of sum 1 and p is $(\mathbf{n} - \mathbf{1})$-monotone on $\mathbb{N}_0^m$.*

Proof. Clearly, $p(\mathbf{x}) = (-1)^m \Delta^{\mathbf{1}} S(\mathbf{x})$. Applying $\Delta^{\mathbf{k}}$ to both sides yields $\Delta^{\mathbf{k}} p(\mathbf{x}) = (-1)^m \Delta^{\mathbf{k}+\mathbf{1}} S(\mathbf{x})$. Since $m = |\mathbf{1}|$, we then obtain

$$(-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} p(\mathbf{x}) = (-1)^{|\mathbf{k}+\mathbf{1}|} \Delta^{\mathbf{k}+\mathbf{1}} S(\mathbf{x}).$$

By Proposition 1, the right-hand side is non-negative for $\mathbf{0} \lneqq \mathbf{k} + \mathbf{1} \leq \mathbf{n}$. Thus, the left-hand side is non-negative for $\mathbf{0} \lneqq \mathbf{k} \leq \mathbf{n} - \mathbf{1}$, i.e., p is $(\mathbf{n} - \mathbf{1})$-monotone. $\square$

3

3. Joint distributions and characterizations

We start by deriving simple formulas for different joint distributions on subvectors of $\mathbf{X}$ and related sums. For clarity, we will use the notation

$$\binom{\mathbf{k} + \mathbf{l}}{\mathbf{l}} = \binom{k_1 + \ell_1}{\ell_1} \cdots \binom{k_m + \ell_m}{\ell_m}.$$

Inside the vector $\mathbf{X}$, consider any subvector $\mathbf{X_k} = (\mathbf{X}_{1,k_1}, \ldots, \mathbf{X}_{m,k_m})$, where for each group, $\mathbf{X}_{j,k_j} = (X_{j,1}, \ldots, X_{j,k_j})$ with $k_j \in \{1, \ldots, n_j\}$. To $\mathbf{X_k}$, we associate the vector of total sums $\mathbf{T_k} = (T_{1,k_1}, \ldots, T_{m,k_m})$ where for each group, $T_{j,k_j} = X_{j,1} + \cdots + X_{j,k_j}$. This vector plays a key role.

To obtain its distribution (10) below, we define the vector of partial sums $\mathbf{T_{(k)}} = (\mathbf{T}_{1,(k_1)}, \ldots, \mathbf{T}_{m,(k_m)})$ where $\mathbf{T}_{j,(k_j)} = (T_{j,1}, \ldots, T_{j,k_j})$. More generally, we also introduce the subvectors $\mathbf{T_{(l,k)}} = (\mathbf{T}_{1,(l_1,k_1)}, \ldots, \mathbf{T}_{m,(l_m,k_m)})$ where $\mathbf{T}_{j,(l_j,k_j)} = (T_{j,\ell_j}, \ldots, T_{j,k_j})$ with $\ell_j \leq k_j$.

Proposition 2. *Let $\mathbf{1} \leq \mathbf{k} \leq \mathbf{n}$. For $\mathbf{x_k} \in \mathbb{N}_0^{|\mathbf{k}|}$,*

$$\Pr(\mathbf{X_k} = \mathbf{x_k}) = (-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} S(|\mathbf{x}_{1,k_1}|, \ldots, |\mathbf{x}_{m,k_m}|), \tag{8}$$

or equivalently, for $\mathbf{t_{(k)}} \in \mathbb{N}_0^{|\mathbf{k}|}$ where $t_{j,1} \leq \cdots \leq t_{j,k_j}$ for all $j \in \{1, \ldots, m\}$,

$$\Pr(\mathbf{T_{(k)}} = \mathbf{t_{(k)}}) = (-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} S(t_{1,k_1}, \ldots, t_{m,k_m}). \tag{9}$$

For $\mathbf{t_k} \in \mathbb{N}_0^m$,

$$\Pr(\mathbf{T_k} = \mathbf{t_k}) = (-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} S(t_{1,k_1}, \ldots, t_{m,k_m}) \binom{\mathbf{t_k} + \mathbf{k} - \mathbf{1}}{\mathbf{k} - \mathbf{1}}, \tag{10}$$

which gives, for $\mathbf{t_{(k-1)}} \in \mathbb{N}_0^{|\mathbf{k-1}|}$ with $\mathbf{k} \geq \mathbf{2}$,

$$\Pr(\mathbf{T_{(k-1)}} = \mathbf{t_{(k-1)}} \mid \mathbf{T_k} = \mathbf{t_k}) = 1 \Big/ \binom{\mathbf{t_k} + \mathbf{k} - \mathbf{1}}{\mathbf{k} - \mathbf{1}}. \tag{11}$$

In general, for $\mathbf{t_{(l,k)}} \in \mathbb{N}_0^{|\mathbf{k-l}|}$ where $t_{j,\ell_j} \leq \cdots \leq t_{j,k_j}$ for all $j \in \{1, \ldots, m\}$,

$$\Pr(\mathbf{T_{(l,k)}} = \mathbf{t_{(l,k)}}) = (-1)^{|\mathbf{k}|} \Delta^{\mathbf{k}} S(t_{1,k_1}, \ldots, t_{m,k_m}) \binom{\mathbf{t_l} + \mathbf{l} - \mathbf{1}}{\mathbf{l} - \mathbf{1}}. \tag{12}$$

Proof. From (2), the subvector $\mathbf{X_k}$ is Schur-constant with

$$\Pr(\mathbf{X_k} \geq \mathbf{x_k}) = S(|\mathbf{x}_{1,k_1}|, \ldots, |\mathbf{x}_{m,k_m}|). \tag{13}$$

Note that, for $j \in \{1, \ldots, m\}$,

$$\Pr(\mathbf{X}_{j,k_j} = \mathbf{x}_{j,k_j}) = (-1)^{k_j} \Delta_1 \cdots \Delta_{k_j} \Pr(\mathbf{X}_{j,k_j} \geq \mathbf{x}_{j,k_j}). \tag{14}$$

From (13) and (14), we then deduce formula (8). This is evidently equivalent to (9). For $\mathbf{T_k}$, we observe that

$$\Pr(\mathbf{T_k} = \mathbf{t_k}) = \Pr(\mathbf{T_{(k)}} = \mathbf{t_{(k)}}) N_{\mathbf{k}},$$

where $N_{\mathbf{k}}$ counts the different ways to put t_{j,k_j} balls in k_j urns for $j \in \{1, \ldots, m\}$. Thus, we get $N_{\mathbf{k}} = \binom{\mathbf{t_k} + \mathbf{k} - \mathbf{1}}{\mathbf{k} - \mathbf{1}}$. Substituting this and using (9), we then obtain (10). Formula (11) is immediate from (9) and (10). Formula (12) can be derived by extending the argument followed for (10). □

We now obtain a general representation for a partially Schur-constant generator S. For that, we will compute the conditional distribution of the subvector $\mathbf{X_1} = (\mathbf{X}_{1,1}, \ldots, \mathbf{X}_{m,1})$ given the total sums vector $\mathbf{T_n} = (T_{1,n_1}, \ldots, T_{m,n_m})$.

Proposition 3. [First characterization] *The model* $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ *is partially Schur-constant if its generator* $S(x_1, \ldots, x_m)$, *with* $(x_1 \ldots, x_m) \in \mathbb{N}_0^m$, *can be represented as*

$$S(x_1, \ldots, x_m) = \mathrm{E}\left\{ \prod_{j=1}^{m} \binom{Z_j - x_j + n_j - 1}{n_j - 1} \Big/ \binom{Z_j + n_j - 1}{n_j - 1} \right\}, \tag{15}$$

where $(Z_1, \ldots, Z_m)$ *is a random vector which is distributed as* $\mathbf{T_n} = (T_{1,n_1}, \ldots, T_{m,n_m})$, *i.e., with a joint pmf given by* (10) *for* $\mathbf{k = n}$.

Proof. Let $\mathbf{1} \leq \mathbf{k} \leq \mathbf{n}$. For $\mathbf{x_k} \in \mathbb{N}_0^{|\mathbf{k}|}$ and $\mathbf{t_n} \in \mathbb{N}_0^m$, we have, using (10),

$$\Pr(\mathbf{X_k} \geq \mathbf{x_k}|\mathbf{T_n} = \mathbf{t_n}) = \frac{\Pr(\mathbf{X_k} \geq \mathbf{x_k}, \mathbf{T_n} = \mathbf{t_n})}{(-1)^{|\mathbf{n}|} \Delta^{\mathbf{n}} S(t_{1,n_1}, \ldots, t_{m,n_m}) \binom{\mathbf{t_n + n - 1}}{\mathbf{n - 1}}}. \tag{16}$$

For the numerator, we easily see from (9) that

$$\Pr(\mathbf{X_k} \geq \mathbf{x_k}, \mathbf{T_n} = \mathbf{t_n}) = (-1)^{|\mathbf{n}|} \Delta^{\mathbf{n}} S(t_{1,n_1}, \ldots, t_{m,n_m}) M_{\mathbf{k,n}},$$

where $M_{\mathbf{k,n}}$ counts the different ways to put t_{j,n_j} balls in n_j urns with at least $x_{j,1}$ balls in urn 1, ..., x_{j,k_j} balls in urn k_j, for $j \in \{1, \ldots, m\}$. Thus, we get $M_{\mathbf{k,n}} = \binom{\mathbf{t_n - |x_k| + n - 1}}{\mathbf{n - 1}}$. Inserting this in (16) yields

$$\Pr(\mathbf{X_k} \geq \mathbf{x_k}|\mathbf{T_n} = \mathbf{t_n}) = \binom{\mathbf{t_n - |x_k| + n - 1}}{\mathbf{n - 1}} \Big/ \binom{\mathbf{t_n + n - 1}}{\mathbf{n - 1}}. \tag{17}$$

When $\mathbf{k = 1}$ in (17), we then obtain

$$\Pr(X_{1,1} \geq x_{1,1}, \ldots, X_{m,1} \geq x_{m,1}|\mathbf{T_n} = \mathbf{t_n}) = \prod_{j=1}^{m} \binom{t_{j,n_j} - x_{j,1} + n_j - 1}{n_j - 1} \Big/ \binom{t_{j,n_j} + n_j - 1}{n_j - 1},$$

and the announced formula (15) directly follows. $\square$

Using this result, we are able to derive a general distributional representation for a partially Schur-constant vector $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$. This illuminating representation is very useful for modeling purposes, as will be illustrated in Section 6.

Proposition 4. [Second characterization] *The model* $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ *is partially Schur-constant if it has an m-dimensional doubly mixed multinomial* ($\mathcal{MM}_m$) *distribution of the form*

$$(\mathbf{X}_1, \ldots, \mathbf{X}_m) =_d \mathcal{MM}_m[(Z_j; U_{j,1}, \ldots, U_{j,n_j}), j \in \{1, \ldots, m\}]. \tag{18}$$

Here, the random vector $(Z_1, \ldots, Z_m)$ *gives the numbers of experiments made for the m types; it is distributed as* $\mathbf{T_n} = (T_{1,n_1}, \ldots, T_{m,n_m})$. *The random vectors* $(U_{j,1}, \ldots, U_{j,n_j})$, *with* $j \in \{1, \ldots, m\}$, *give the randomized cell probabilities for the m types; they are independent between them and of the vector* $(Z_1, \ldots, Z_m)$, *and they have a continuous Schur-constant survival function given, for all* $u_{j,1}, \ldots, u_{j,n_j} \in (0, 1)$, *by*

$$\Pr(U_{j,1} > u_{j,1}, \ldots, U_{j,n_j} > u_{j,n_j}) = [1 - (u_{j,1} + \cdots + u_{j,n_j})]_+^{n_j - 1}, \tag{19}$$

i.e., $(U_{j,1}, \ldots, U_{j,n_j})$ *is uniformly distributed over the* n_j*-dimensional unit simplex.*

Proof. We want to show that under the representation (18), (19), the pmf of the vector $\mathbf{X_k}$, with $\mathbf{1} \leq \mathbf{k} \leq \mathbf{n}$, is identical to the pmf provided by the generator (15). Of course, it suffices to prove it conditionally on the vector $(Z_1, \ldots, Z_m)$, which is distributed as $\mathbf{T_n}$ by assumption. Thus, let us determine $\Pr(\mathbf{X_k} = \mathbf{x_k}|\mathbf{T_n} = \mathbf{t_n})$, with $\mathbf{1} \leq \mathbf{k} \leq \mathbf{n - 1}$.

First, for the characterization through (15), the survival function of $\mathbf{X_k}$ given $\mathbf{T_n}$ is given by (17). To use formula (14), we notice that, for $j \in \{1, \ldots, m\}$,

$$\Delta_j^{k_j}\binom{t_{j,n_j} - x_{j,k_j} + n_j - 1}{n_j - 1} = (-1)^{k_j}\binom{t_{j,n_j} - x_{j,k_j} + n_j - 1 - k_j}{n_j - 1 - k_j}.$$

We then deduce from (17) that

$$\Pr(\mathbf{X_k} = \mathbf{x_k}|\mathbf{T_n} = \mathbf{t_n}) = \binom{\mathbf{t_n} - |\mathbf{x_k}| + \mathbf{n} - \mathbf{1} - \mathbf{k}}{\mathbf{n} - \mathbf{1} - \mathbf{k}} \Big/ \binom{\mathbf{t_n} + \mathbf{n} - \mathbf{1}}{\mathbf{n} - \mathbf{1}}. \tag{20}$$

Now, for the characterization (18), (19) stated above, we follow the method used to prove Proposition 3.2 in [6]. The calculations are close and omitted; they bring us back to the desired formula (20). $\quad\square$

4. Mixed geometric model as a limit

It is easily seen that inside each vector $\mathbf{X}_j$, the variables $(X_{j,1}, \ldots, X_{j,n_j})$ are independent if and only if they are geometrically distributed with some parameter q_j. Thus, a partially Schur-constant model with independent variables is of survival function

$$S(|\mathbf{x}_1|, \ldots, |\mathbf{x}_m|) = q_1^{|\mathbf{x}_1|} \times \cdots \times q_m^{|\mathbf{x}_m|}. \tag{21}$$

Interestingly, we are going to show that if a m-dimensional vector is partially Schur-constant for any order, all models involved have a m-dimensional mixed geometric distribution. Thus, we assume below that the condition (2) is satisfied for all $(n_1, \ldots, n_m) \in \mathbb{N}_0^m$. For each group j, we define as before the subvectors $\mathbf{X}_{j,k_j} = (X_{j,1}, \ldots, X_{j,k_j})$ and the associated sums $T_{j,k_j} = X_{j,1} + \cdots + X_{j,k_j}$, for all $k_j \geq 1$.

Proposition 5. *A vector $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ is partially Schur-constant for any order $(n_1, \ldots, n_m)$ if, and only if, for any $k_1, \ldots, k_m \geq 1$, the vector $(\mathbf{X}_{1,k_1}, \ldots, \mathbf{X}_{m,k_m})$ has a m-dimensional mixed geometric distribution such that, for all $\mathbf{x}_{j,k_j} \in \mathbb{N}_0^{k_j}$,*

$$\Pr\left(\mathbf{X}_{1,k_1} \geq \mathbf{x}_{1,k_1}, \ldots, \mathbf{X}_{m,k_m} \geq \mathbf{x}_{m,k_m}\right) = \mathrm{E}\left\{\prod_{j=1}^{m}\left(\frac{\Theta_j}{\Theta_j + 1}\right)^{|\mathbf{x}_{j,k_j}|}\right\}, \tag{22}$$

where

$$(\Theta_1, \ldots, \Theta_m) = \lim_{n_1, \ldots, n_m \to \infty} (T_{1,n_1}/n_1, \ldots, T_{m,n_m}/n_m) \quad a.s. \tag{23}$$

Proof. First, the Strong Law of Large Numbers for a partially exchangeable sequence $X_{1,n_1}, \ldots, X_{m,n_m}$ with finite means gives us the convergence result (23); see, e.g., [10]. Now, to obtain the representation (22), we insert the expression (15) of S into formula (13). This yields, for all $k_j \in \{1, \ldots, n_j\}$,

$$\Pr(\mathbf{X}_{1,k_1} \geq \mathbf{x}_{1,k_1}, \ldots, \mathbf{X}_{m,k_m} \geq \mathbf{x}_{m,k_m}) = \mathrm{E}\left\{\prod_{j=1}^{m}\binom{T_{j,n_j} - |\mathbf{x}_{j,k_j}| + n_j - 1}{n_j - 1}\Big/\binom{T_{j,n_j} + n_j - 1}{n_j - 1}\right\}$$

$$= \mathrm{E}\left[\prod_{j=1}^{m}\left\{\prod_{\ell=1}^{n_j-1}\left(1 - \frac{|\mathbf{x}_{j,k_j}|}{T_{j,n_j} + \ell}\right)\mathbf{1}(T_{j,n_j} \geq |\mathbf{x}_{j,k_j}|)\right\}\right], \tag{24}$$

where $\mathbf{1}(\cdot)$ is the indicator function. As all $n_j \to \infty$, we use Lebesgue's Dominated Convergence Theorem and rewrite the right-hand side of (24) as

$$\mathrm{E}\left[\prod_{j=1}^{m}\lim_{n_j \to \infty}\left\{\prod_{\ell=1}^{n_j-1}\left(1 - \frac{|\mathbf{x}_{j,k_j}|}{n_j\Theta_{j,n_j} + \ell}\right)\mathbf{1}(n_j\Theta_{j,n_j} \geq |\mathbf{x}_{j,k_j}|)\right\}\right], \tag{25}$$

6

after denoting $\Theta_{j,n_j} = T_{j,n_j}/n_j$. Thus, formula (22) will follow directly from (25) if each term $[\cdots]$ in (25) converges to $\{\Theta_j/(\Theta_j + 1)\}^{|\mathbf{x}_{j,k_j}|}$ as $n_j \to \infty$. This can be proved in the same way as in [6] using the limit result (23). $\square$

In other words, multivariate mixed geometric distributions can be considered as a limit of the partially Schur-constant models. As a referee pointed out, multivariate geometric distributions provide a wide and interesting class of discrete models. A thorough analysis of these distributions is done in [16] and extended in [22], with particular focus on multivariate extensions of the property of lack of memory. It should be noted that the exchangeable subclasses considered in these works are characterized by properties of multiple monotonicity of the parameters concerned. Let us also mention a different approach by Sreehari and Vasudeva [23] based on conditional distributions.

5. Correlations with illustrations

By partial exchangeability, within each vector $\mathbf{X}_j$, the variables $X_{j,1}, \ldots, X_{j,n_j}$ have the same mean, variance and Pearson correlation coefficients; these are denoted by μ_j, σ_j^2 and ρ_j. Moreover, between any two vectors $\mathbf{X}_j$ and $\mathbf{X}_k$, $j \neq k$, all the pairs of variables have the same correlation coefficient $\rho_{j,k}$, say.

Our goal here is to determine these parameters using the representation (18), (19). We denote by μ_{Z_j} and $\sigma_{Z_j}^2$ the mean and variance of Z_j, and by ρ_{Z_j,Z_k} the correlation coefficient between Z_j and Z_k, $j \neq k$. First, the vector $\mathbf{X}_j$ being Schur-constant, we know from [6] that

$$\mu_j = \mu_{Z_j}/n_j, \tag{26}$$

$$\sigma_j^2 = \frac{2\sigma_{Z_j}^2}{n_j(n_j + 1)} + \frac{\mu_{Z_j}^2(n_j - 1)}{n_j^2(n_j + 1)} + \frac{\mu_{Z_j}(n_j - 1)}{n_j(n_j + 1)}, \tag{27}$$

$$\rho_j = (\sigma_j^2 - \mu_j^2 - \mu_j)/2\sigma_j^2 \tag{28}$$

$$= \frac{n_j \sigma_{Z_j}^2 - \mu_{Z_j}^2 - n_j \mu_{Z_j}}{2n_j \sigma_{Z_j}^2 + (n_j - 1)\mu_{Z_j}^2 + n_j(n_j - 1)\mu_{Z_j}}. \tag{29}$$

Let us now examine the cross-correlation coefficient $\rho_{j,k}$. We need the coefficients of variation of $X_{j,1}, X_{k,1}$, denoted by v_j, v_k, and those of Z_j, Z_k, denoted by v_{Z_j}, v_{Z_k}.

Proposition 6.

$$\rho_{j,k} = \rho_{Z_j,Z_k} \frac{v_{Z_j} v_{Z_k}}{v_j v_k}. \tag{30}$$

Proof. By definition, $\rho_{j,k}$ is the correlation between $X_{j,1}$ and $X_{k,1}$. From Proposition 4, these variables have a mixed binomial distribution, $\mathcal{MB}_j$ and $\mathcal{MB}_k$, of the form

$$X_{j,1} =_d \mathcal{MB}_j(Z_j; U_{j,1}) \quad \text{and} \quad X_{k,1} =_d \mathcal{MB}_k(Z_k; U_{k,1}),$$

where $U_{j,1}$ and $U_{k,1}$ are independent variables with survival functions $\Pr(U_{j,1} \geq u_{j,1}) = (1 - u_{j,1})^{n_j - 1}$ and $\Pr(U_{k,1} \geq u_{k,1}) = (1 - u_{k,1})^{n_k - 1}$, independently of the vector (Z_j, Z_k), and given fixed values for the random parameters, the (non-mixed) distributions B_j and B_k are independent binomials. Using a conditional argument and these independence properties, we then have

$$\begin{aligned}
\mathrm{E}(X_{j,1}X_{k,1}) &= \mathrm{E}\{\mathcal{MB}_j(Z_j; U_{j,1})\, \mathcal{MB}_k(Z_k; U_{k,1})\} \\
&= \mathrm{E}[\mathrm{E}\{B_j(Z_j; U_{j,1})\, B_k(Z_k; U_{k,1})|Z_j, Z_k, U_{j,1}, U_{k,1}\}] \\
&= \mathrm{E}[\mathrm{E}\{B_j(Z_j; U_{j,1})|Z_j, U_{j,1}\}\, \mathrm{E}\{B_k(Z_k; U_{k,1})|Z_k, U_{k,1}\}] \\
&= \mathrm{E}(Z_j U_{j,1} Z_k U_{k,1}) = \mathrm{E}(Z_j Z_k)\, \mathrm{E}(U_{j,1})\, \mathrm{E}(U_{k,1}). \tag{31}
\end{aligned}$$

Since $\mathrm{E}(U_{j,1}) = 1/n_j = \mu_j/\mu_{Z_j}$ — see (26) — we can rewrite (31) as $\mathrm{E}(X_{j,1}X_{k,1})/\mu_j\mu_k = \mathrm{E}(Z_jZ_k)/\mu_{Z_j}\mu_{Z_k}$, which is equivalent to (30) when expressed in terms of coefficients of variation. $\square$

7

Note that $\rho_{j,k}$ and ρ_{Z_j,Z_k} have the same sign. Using (26) and (27), we directly check that $v_j \geq v_{Z_j}$ for each j. Thus, (30) implies that $\rho_{j,k} \leq \rho_{Z_j,Z_k}$, which is an expected inequality.

Of course, $\rho_{j,k}$ can also be evaluated using the generator S. Indeed, from (8), we find that when $j < k$,

$$
\begin{aligned}
\mathrm{E}(X_{j,1}X_{k,1}) &= \sum_{x_{j,1}=0}^{\infty}\sum_{x_{k,1}=0}^{\infty} x_{j,1}x_{k,1}\Pr(X_{j,1}=x_{j,1}, X_{k,1}=x_{k,1}) \\
&= \sum_{x_{j,1}=0}^{\infty}\sum_{x_{k,1}=0}^{\infty} x_{j,1}x_{k,1}\Delta_j\Delta_k S(0,\ldots,x_{j,1},\ldots,x_{k,1},\ldots,0) \\
&= \sum_{x_{j,1}=0}^{\infty}\sum_{x_{k,1}=0}^{\infty} S(0,\ldots,x_{j,1}+1,\ldots,x_{k,1}+1,\ldots,0),
\end{aligned}
$$

the last equality following from a direct calculation.

As an illustration, we consider below several partially Schur-constant models with $m = 2$ groups. These examples also allow us to highlight the construction of the model and its flexibility. Remember that by virtue of (15), the generator has the form

$$
S(x_1, x_2) = \mathrm{E}\left\{ \frac{\binom{Z_1-x_1+n_1-1}{n_1-1}\binom{Z_2-x_2+n_2-1}{n_1-1}}{\binom{Z_1+n_1-1}{n_1-1}\binom{Z_2+n_2-1}{n_2-1}} \right\}, \tag{32}
$$

for some random vector (Z_1, Z_2) valued in $\mathbb{N}_0^2$.

Example 1. Suppose that (Z_1, Z_2) has a bivariate Bernoulli distribution with pmf given by

Z_1 / Z_2	0	1
0	0.60	0.15
1	0.15	0.10

So, Z_1 and Z_2 are two Bernoulli variables of parameter $p = 0.25$ such that $\Pr(Z_1 = Z_2 = 1) = 0.1$.

Take, for instance, $n_1 = n_2 = 2$. From (32), we get

$$
\begin{aligned}
S(x_1, x_2) &= 0.6(1-x_1)(1-x_2) + 0.15(2-x_1)(1-x_2)/2 + 0.15(1-x_1)(2-x_2)/2 \\
&\quad + 0.1(2-x_1)(2-x_2)/4 = 1 - 0.875x_1 - 0.875x_2 + 0.775x_1x_2,
\end{aligned} \tag{33}
$$

when $x_1, x_2 \in \{0, 1\}$, and null otherwise. From (2) and (33), the vector $(\mathbf{X}_1, \mathbf{X}_2)$ is of survival function

$$
\begin{aligned}
\Pr\{(X_{1,1} \geq x_{1,1}, X_{1,2} \geq x_{1,2}), (X_{2,1} \geq x_{2,1}, X_{2,2} \geq x_{2,2})\} &= S(x_{1,1}+x_{1,2}, x_{2,1}+x_{2,2}) \\
&= 1 - 0.875(x_{1,1}+x_{1,2}) - 0.875(x_{2,1}+x_{2,2}) + 0.775(x_{1,1}+x_{1,2})(x_{2,1}+x_{2,2}),
\end{aligned} \tag{34}
$$

when $x_{1,1} + x_{1,2}, x_{2,1} + x_{2,2} \in \{0, 1\}$. From (8) and (34), its pmf is

$$
\begin{aligned}
\Pr\{(X_{1,1} = x_{1,1}, X_{1,2} = x_{1,2}), (X_{2,1} = x_{2,1}, X_{2,2} = x_{2,2})\} &= \Delta_1^2\Delta_2^2 S(x_{1,1}+x_{1,2}, x_{2,1}+x_{2,2}) \\
&= 4S(x_{1,1}+x_{1,2}+1, x_{2,1}+x_{2,2}+1) - 2S(x_{1,1}+x_{1,2}+1, x_{2,1}+x_{2,2}) \\
&\quad - 2S(x_{1,1}+x_{1,2}, x_{2,1}+x_{2,2}+1) + S(x_{1,1}+x_{1,2}, x_{2,1}+x_{2,2}),
\end{aligned} \tag{35}
$$

if again $x_{1,1} + x_{1,2}, x_{2,1} + x_{2,2} \in \{0, 1\}$. Using (32) and (35), we obtain

$$
\begin{aligned}
\Pr\{(X_{11} = 0, X_{12} = 0), (X_{21} = 0, X_{22} = 0)\} &= 0.6, \\
\Pr\{(X_{11} = 1, X_{12} = 0), (X_{21} = 0, X_{22} = 0)\} &= 0.075, \\
\Pr\{(X_{11} = 0, X_{12} = 0), (X_{21} = 1, X_{22} = 0)\} &= 0.075, \\
\Pr\{(X_{11} = 1, X_{12} = 0), (X_{21} = 1, X_{22} = 0)\} &= 0.025.
\end{aligned}
$$

From (26), (27), (28), we directly compute the correlations intra-group as $\rho_1 = -0.142857 = \rho_2$. We also see that $\rho_{Z_1,Z_2} = [\Pr(Z_1 = Z_2 = 1) - p^2]/p(1-p) = 0.2$ and $v_{Z_j}v_{Z_k}/v_j v_k = (1-p)/(2-p) = 3/7$, so that (30) gives $\rho_{1,2} = 0.085714$.

Example 2. Suppose that (Z_1, Z_2) has a bivariate Poisson distribution given by the common shock model

$$Z_1 = N_1 + M, \quad Z_2 = N_2 + M, \tag{36}$$

where N_1, N_2, M are three independent Poisson variables of parameters λ_1, λ_2, $\lambda > 0$, respectively. Thus, Z_1, Z_2, are two Poisson variables of parameter $\lambda_j + \lambda$ and their correlation is $\rho_{Z_1,Z_2} = \lambda/\{(\lambda_1 + \lambda)(\lambda_2 + \lambda)\}^{1/2}$. Of course, ρ_{Z_1,Z_2} is positive and increases with λ.

The correlation coefficients are easily calculated. From (26), (27), we have $\mu_j = (\lambda_j + \lambda)/n_j$ and $\sigma_j^2 = \mu_j + \mu_j^2(n_j - 1)/(n_j + 1)$. So, (29) gives $\rho_j = -1/\{n_j - 1 + n_j^2(n_j + 1)/(\lambda_j + \lambda)\}$, which is negative and decreases with λ, as expected. Moreover, as $v_{Z_j} = 1/(\lambda_j + \lambda)^{1/2}$ and $v_j = \{n_j/(\lambda_j + \lambda) + (n_j - 1)/(n_j + 1)\}^{1/2}$, we obtain from (30) that

$$\rho_{1,2} = \frac{\lambda}{\prod_{j=1}^{2}(\lambda_j + \lambda)\{n_j/(\lambda_j + \lambda) + (n_j - 1)/(n_j + 1)\}^{1/2}},$$

which is positive and not necessarily monotone in λ.

Take, for instance, $n_1 = 2$, $n_2 = 3$, $\lambda_1 = 0.05$ and $\lambda_2 = 0.3$. Table 1 gives ρ_{Z_1,Z_2}, ρ_1, ρ_2 and $\rho_{1,2}$ for different values of the shock parameter λ. The correlation $\rho_{1,2}$ first increases and then decreases with λ; the maximum is at $\lambda = 0.9990$. For a recent extension of model (36), see [12].

Table 1: Correlation coefficients when $n_1 = 2$, $n_2 = 3$ and (Z_1, Z_2) has a bivariate Poisson distribution of parameters $(\lambda_1 = 0.05, \lambda_2 = 0.3, \lambda)$, for different values of λ.

λ	ρ_{Z_1,Z_2}	ρ_1	ρ_2	$\rho_{1,2}$
0.1	0.4082	−0.0244	−0.0312	0.1594
0.2	0.5657	−0.0400	−0.0385	0.2174
0.8	0.8273	−0.1241	−0.0775	0.2906
0.9	0.8429	−0.1367	−0.0833	0.2919
1.5	0.8980	−0.2053	−0.1154	0.2866
2	0.9211	−0.2547	−0.1385	0.2760
3	0.9456	−0.3370	−0.1774	0.2525

Example 3. Suppose that (Z_1, Z_2) has a bivariate negative binomial distribution given by the common shock model (36), where N_1, N_2, M are three independent negative binomial variables of parameters (m_1, p_1), (m_2, p_2), (m_0, p_0), respectively. Thus, the pmf of M is given, for all $i \in \mathbb{N}_0$, by

$$\Pr(M = i) = \frac{\Gamma(m_0 + i)}{i!\Gamma(m_0)} p_0^i (1 - p_0)^{m_0},$$

where $m_0 > 0$ and $p_0 \in (0, 1)$. We reparametrize it in a standard way by setting $p_0 = h_0/(1 + h_0)$ (which yields h_0) and $m_0 = g_0/h_0$ (which yields g_0). This gives $E(M) = m_0 p_0/(1 - p_0) = g_0$, and $\sigma_M^2 = m_0 p_0/(1 - p_0)^2 = g_0(1 + 1/h_0)$ where $1/h_0$ corresponds to a clustering parameter ($1/h_0 = 0$ in the Poisson case). A similar operation is made on the parameters of N_1, N_2. We then obtain

$$\rho_{Z_1,Z_2} = \frac{g_0(1 + 1/h_0)}{[\{g_0(1 + 1/h_0) + g_1(1 + 1/h_1)\}\{g_0(1 + 1/h_0) + g_2(1 + 1/h_2)\}]^{1/2}},$$

which is positive and increasing with $1/h_0$.

Here too, the correlation coefficients can be determined explicitly. The formulas are omitted for brevity. Take, for instance, $n_1 = 2$, $n_2 = 3$, $(g_1 = 4, h_1 = 6)$, $(g_2 = 4, h_2 = 1)$ and $(g_0 = 5, h_0)$. Table 2 gives ρ_{Z_1,Z_2}, ρ_1, ρ_2 and $\rho_{1,2}$ for different values of the non-aggregation parameter h_0. We observe that all the correlations decrease with h_0, i.e., when the clustering strength becomes weaker.

9

Table 2: Correlation coefficients when $n_1 = 2$, $n_2 = 3$ and (Z_1, Z_2) has a bivariate negative binomial distribution of parameters $[(g_1 = 4, h_1 = 6), (g_2 = 4, h_2 = 2), (g_0 = 5, h_0)]$, for different values of h_0.

h_0	ρ_{Z_1,Z_2}	ρ_1	ρ_2	$\rho_{1,2}$
0.1	0.8971	0.0602	0.1364	0.4254
0.2	0.8266	−0.1248	0.0135	0.3199
0.8	0.6427	−0.4129	−0.1516	0.1678
0.9	0.6281	−0.4288	−0.1599	0.1598
1.5	0.5719	−0.4834	−0.1879	0.1326
2	0.5461	−0.5056	−0.1990	0.1216
3	0.5171	−0.5289	−0.2105	0.1102

Example 4. Suppose that (Z_1, Z_2) is a vector that has marginally two binomial distributions of parameters $(m_1 = 2, p_1 \in (0, 1))$, $(m_2 = 3, p_2 \in (0, 1))$, respectively. Their joint distribution must fulfill certain conditions to satisfy this property. Four examples are presented: two when $p_1 = p_2 = 0.8$ and two when $p_1 = p_2 = 0.3$. The pmf of (Z_1, Z_2) is provided in the right hand corner of each graph in Figure 1, as well as the correlation ρ_{Z_1,Z_2}.

We want to illustrate how the group sizes n_1 and n_2 can influence the correlation coefficient $\rho_{1,2}$. From (30), a simple calculation yields

$$\rho_{1,2} = \rho_{Z_1,Z_2} \frac{(1 - p_1)(1 - p_2)(n_1 + 1)(n_2 + 1)}{\{n_1(n_1 + 1 - 2p_1) + m_1 p_1(n_1 - 1)\}\{n_2(n_2 + 1 - 2p_2) + m_2 p_2(n_2 - 1)\}}.$$

From Figure 1, we see that if $\rho_{Z_1,Z_2} > 0$, then $\rho_{1,2}$ (> 0) decreases with n_1 and n_2, while if $\rho_{Z_1,Z_2} < 0$, then $\rho_{1,2}$ (< 0) increases with n_1 and n_2. This is consistent with intuition.
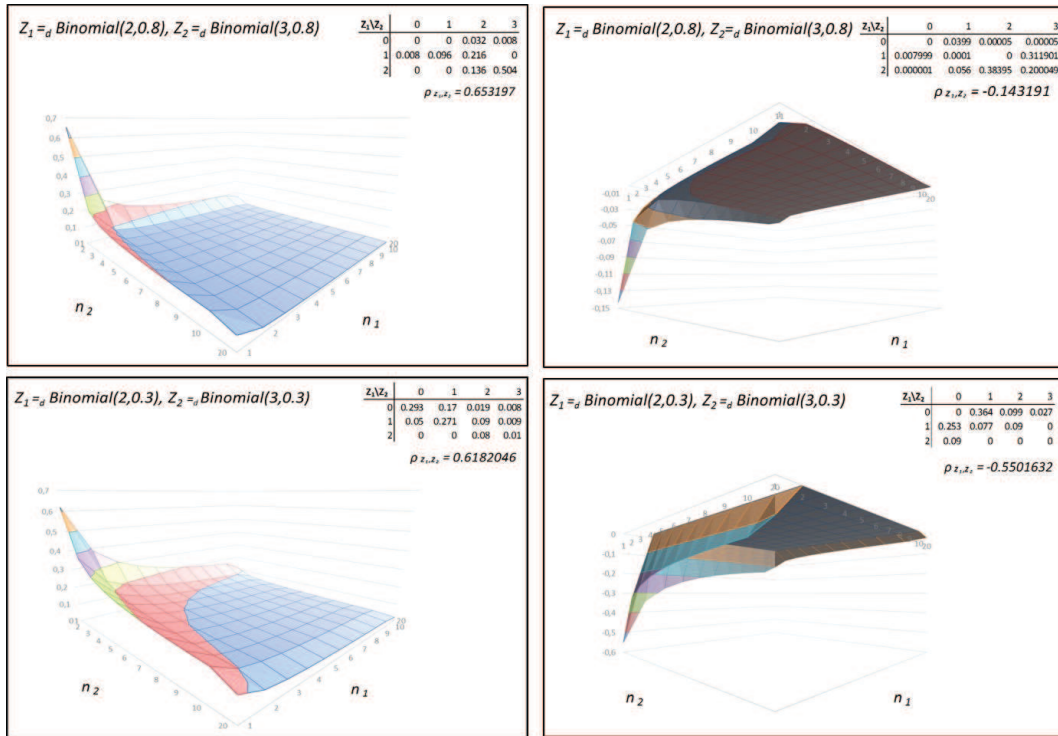

Figure 1: Correlations $\rho_{1,2}$ when (Z_1, Z_2) has four different joint pmf in which marginals are binomials of parameters $(m_1 = 2, p_1 \in (0, 1))$ and $(m_2 = 3, p_2 \in (0, 1))$ with $p_1 = p_2 = 0.8$ or $p_1 = p_2 = 0.3$, as a function of the group sizes n_1 and n_2.

6. An application to risk management

In this section, we illustrate how partially Schur-constant models can be involved in a macro-prudential risk management for insurance-reinsurance networks. When an extreme event occurs, it may cause some insurers and reinsurers to be bankrupt. The shock may propagate through the network by contagion, as some reinsurers may be unable to cover what they have to pay to the insurers who had transferred risks to them. Very often, regulators model these shocks by considering the (random) total insured loss, and then affect the losses to each insurer in proportion to the market shares. This means that they typically use a model for the sum of the variables representing the individual losses, and then assign that sum in deterministic proportions. The partially Schur-constant model improves this approach by considering random exchangeable proportions. Such a model can be continuous or discrete in nature. If it is continuous, a discretization may also be necessary for practical reasons. For example, regulators and central banks sometimes develop certain software to avoid spyware, and some of these tools must be used in a discrete environment. This is why we present here an illustration in the discrete frame. Of course, the current model could be developed in a continuous framework.

6.1. Basic model

Our regulator is concerned with systemic risk in an insurance-reinsurance network created to cover potentially important events (like nuclear terrorism or mega earthquakes). The regulator has partial information about the exposures to risk in the network. Firstly, there are $n_1 = 3$ large (type 1) insurers with a market share of 25% each (labeled $(1,1), (1,2), (1,3)$), and $n_2 = 5$ small (type 2) insurers with a market share of 5% each (labeled $(2,1), \ldots, (2,5)$). Customers of the large insurers are more concentrated in big cities, while small insurers have more customers in the countryside.

In addition, two reinsurers, A and B, can intervene according to certain rules. On the one hand, each insurer uses non-proportional reinsurance for each event. More precisely, the insurer (i, j) uses a treaty $\ell_{i,j}^A$ xs $r_{i,j}^A$ with the reinsurer A, and a treaty $\ell_{i,j}^B$ xs $r_{i,j}^B$ with the reinsurer B. The notation ℓ xs r means that the reinsurer pays the amount of the claim that exceeds a retention r, but with a total liability of at most ℓ. In other words, if the bulk cost is x, the reinsurer pays $\min(x, \ell + r) - \min(x, r)$. On the other hand, the reinsurer B retrocedes part of its risk to the reinsurer A with a treaty ℓ_R xs r_R. Figure 2 shows a scheme of the network.
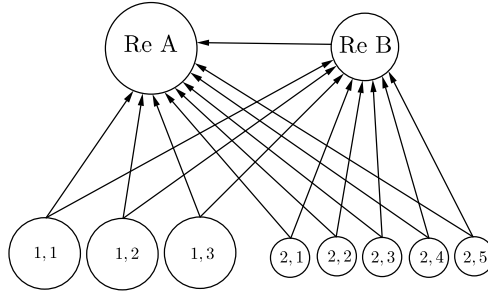

Figure 2: Scheme of the insurance-reinsurance network.

As mentioned earlier, the regulator often has a model that gives the distribution of the random total loss Z for the insurance industry if a catastrophic event occurs. It also has a joint model for the total insurance losses (Z_1, Z_2) of both types of insurers (with $Z = Z_1 + Z_2$). This model is given by the common shock model

$$Z_1 = 3N_1 + 3M, \quad Z_2 = N_2 + M,$$

where N_1, N_2 and M are iid random variables with a Zipf distribution of parameter $\alpha = 2.1$. The Zipf (or discrete Pareto) distribution of parameter $\alpha > 0$ has a pmf given by $p(k) = (1 + k)^{-(1+\alpha)}/\zeta(1 + \alpha)$, $k \in \mathbb{N}_0$, where $\zeta(1 + \alpha)$ is the Riemann zeta function. It has a heavy tail distribution and a finite variance when $\alpha > 2$. The variable M takes into account a factor that influences the two zones, while N_1 and N_2 are representative of a factor specific to each zone.

A partially Schur-constant vector provides a multivariate model for the claim amounts of the two groups of insurers. From (18), these claim amounts are then defined by

$$[(X_{1,1}, X_{1,2}, X_{1,3}), (X_{2,1}, \ldots, X_{2,5})] =_d \mathcal{MM}_2[(Z_1; U_{1,1}, U_{1,2}, U_{1,3}), (Z_2; U_{2,1}, \ldots, U_{2,5})],$$

with the constraints indicated through (18), (19).

A partially Schur-constant (PSC) model like this is a natural enhancement of the classical model (with fixed proportions) used by regulators. Within each zone, each insurer plays the same role, so that a claim may be attributed to each insurer according to a partially exchangeable scheme. It would be possible to use different exchangeable distributions on the unit simplex. However, since there is considerable uncertainty about where the disaster occurs and its impact, the regulator will generally tend to use, for each group, a uniform distribution over the unit simplex. This is precisely the framework described by the PSC model above.

Now, let us specify the recovery rules if one of the reinsurers is in default. Since the reinsurer A is at the top of the risk pyramid, we first settle its possible default cases. If A has to pay x to its counterparties and has only $\xi < x$ left, A pays each counterparty a propotion ξ/x of the corresponding liability. We then examine whether the reinsurer B is in default (after receiving full or partial recoveries of A). If B is in default, the same proportional method is used to settle a payment to his counterparties. Finally, insurers and reinsurers have capital buffers. For insurers, the objective is to avoid a recapitalization with a probability of 0.93, neglecting the risk of a reinsurer's default. For reinsurers, the goal is to avoid ruin with a probability of 0.996.

Table 3 shows the parameters of the reinsurance contracts and the capital buffers of the insurers and reinsurers.

Table 3: Parameters of the reinsurance contracts and capital buffers.

Insurer (i, j)	$r_{i,j}^A$	$\ell_{i,j}^A$	$r_{i,j}^B$	$\ell_{i,j}^B$	Capital buffers
$(1, 1), (1, 2), (1, 3)$	17	67	5	4	3
$(2, 1), \ldots, (2, 5)$	4	15	1	1	1
Reinsurer	r_R	ℓ_R			
A	—	—			9
B	11	6			9

6.2. Comparison with related models

For comparative purposes, we consider three other models of similar general structure, but with different forms of dependence:

(i) the conditionally independent (CIND) model in which the losses (Z_1, Z_2) are defined as before, but for each type $j \in \{1, 2\}$ of insurers, the losses for any insurer i are obtained by mixed binomial samplings $\mathcal{MB}(Z_j, U_{j,i})$, where the parameters $U_{j,i}$ are distributed as in the PSC model but are now assumed to be independent. Note that the aggregate losses in each group j will generally be different from Z_j;

(ii) the independent (IND) model in which the losses Z_1 and Z_2 have the same marginal distributions as before but are assumed to be independent, and the losses in both types of insurers are defined as for the CIND model;

(iii) the market share proportional (MSP) model in which losses are proportional to the global market shares, i.e., $(X_{1,1}, X_{1,2}, X_{1,3}) = (Z/4, Z/4, Z/4)$ and $(X_{2,1}, \ldots, X_{2,5}) = (Z/20, \ldots, Z/20)$. This model with fixed proportions corresponds to a current choice, pragmatic, made by many regulators.

To illustrate the effect of the dependencies, we calculate various probabilities of huge losses, insolvency or default for the basic model PSC and the other models CIND, IND and MSP. This is done by a standard Monte Carlo method (through 10 processes with 1,000,000 simulations each).

First, we use the classic value-at-risk concept at level p (denoted p-VaR) to determine the following joint or conditional probabilities of a huge loss for different sets of insurers (regardless of capital buffers):

S_1: the probabilities of huge losses for all the insurers, via the 0.99-VaR (of the PSC model);

12

S_2: the probabilities of huge losses for the large insurers $(1, 1)$ and $(1, 2)$ and the small insurers $(2, 1)$ and $(2, 2)$, via the 0.99-VaR (of the PSC model);

S_3: the conditional probabilities that the payment of the large insurer $(1, 1)$ is greater than its 0.99-VaR given that the payment of the small insurer $(2, 1)$ is greater than its 0.99-VaR (of each model);

S_4: the conditional probabilities that the payment of the large insurer $(1, 1)$ is greater than its 0.99-VaR, given that the payment of the large insurer $(1, 2)$ is greater than its 0.99-VaR (of each model).

Table 4: Joint and conditional probabilities of huge losses.

	PSC	CIND	IND	MSP
S_1	0.0001	0.0002	0.0000	0.0084
S_2	0.0007	0.0010	0.00001	0.0084
S_3	0.2470	0.2674	0.0583	1.0000
S_4	0.3495	0.5113	0.0181	1.0000

The results are given in Table 4. As expected, $S_1 \leq S_2$. For all probabilities, the worst situation is provided by the MSP model which corresponds to a comonotonic case. The second worst case is the CIND model in which there is high correlation inside each zone and between zones. The third case is the PSC model in which there are moderate correlations inside each zone (because of a mixture of positive and negative dependencies) and a high correlation between the zones. The best case is the IND model in which there is no correlation inside each zone and between the zones. Comparing the PSC and IND models, we thus see that, for all probabilities, the positive dependence in the PSC model due to the common factor M has a stronger effect than the negative dependence existing in each zone because the total sum is prescribed.

Then, we determine the following conditional probabilities of insolvency of insurers or default of reinsurers (taking into account capital buffers):

P_1: the probability of contagion from the reinsurers to the insurers, i.e., the conditional probability that at least one insurer becomes insolvent given that at least one reinsurer defaults;

P_2: the conditional probability that the small insurer $(2, 1)$ is insolvent given that the large insurer $(1, 1)$ is insolvent;

P_3: the conditional probability that the small insurers $(2, 1)$ and $(2, 2)$ are insolvent given that the large insurer $(1, 1)$ is insolvent;

P_4: the conditional probability that the reinsurer B defaults given that the reinsurer A defaults.

Table 5: Conditional probabilities of insolvency or default.

	PSC	CIND	IND	MSP
P_1	0.3228	0.3691	0.0130	0.4810
P_2	0.0729	0.0778	0.0276	0.2264
P_3	0.0247	0.0333	0.0009	0.2264
P_4	0.4932	0.6435	0.0060	1.0000

Of course, $P_2 \geq P_3$. Comparing the probabilities given in Table 5, we see that the four models follow the order indicated by Table 4, which is not surprising intuitively. However, a striking point is that these probabilities are of a very different magnitude. The IND model produces very low probabilities and seems to strongly underestimate the risk. On the contrary, the MSP model used by most regulators produces very high probabilities and seems to overestimate, in particular, the contagion probabilities P_1 which constitute the main quantities of interest for the regulators concerned by the systemic risk. As several other sources of risk are not taken into account by the regulator

but should be in the future, it is prudent not to rely on a scheme that is too conservative. Table 5 therefore emphasizes the interest of developing a more precise intermediate approach for the decomposition of the sum of the losses into individual random variables (i.e., losses for each insurer). In the absence of information on the allocation of losses, the use per group of a uniform distribution on the simplex, i.e., of a partially Schur-constant model for the individual losses, appears as a pragmatic way to take into account the randomness of the proportions of losses attributed to each insurer while maintaining the total sum per group.

References

[1] D.J. Aldous, Partial exchangeability, In: S. Kotz, N. Balakrishnan, C. Read, B. Vidakovic, N.L. Johnson (Eds.), Encyclopedia of Statistical Sciences, Wiley, New York, 2006, pp. 5946–5948.

[2] F. Balabdaoui, J.A. Wellner, Estimation of a k-monotone density: Limit distribution theory and the spline connection, Ann. Statist. 35 (2007) 2536–2564.

[3] R.E. Barlow, M.B. Mendel, Similarity as a probabilistic characteristic of aging, In: C.E. Clarotti, R.E. Barlow, F. Spizzichino, (Eds.), Reliability and Decision Making, Chapman & Hall, London, 1993, pp. 233–244.

[4] L. Caramellino, F. Spizzichino, Dependence and aging properties of lifetimes with Schur-constant survival function, Probab. Engrg. Inform. Sci. 8 (1994) 103–111.

[5] A. Castañer, M.M. Claramunt, Equilibrium distributions and discrete Schur-constant models. Methodol. Comput. Appl. Probab. (2018). https://dx.doi.org/10.1007/s11009-018-9632-5.

[6] A. Castañer, M.M. Claramunt, C. Lefèvre, S. Loisel, Discrete Schur-constant models, J. Multivariate Anal. 140 (2015) 343–362.

[7] Y. Chi, J. Yang, Y. Qi, Decomposition of a Schur-constant model and its applications, Insur. Math. Econom. 44 (2009) 398–408.

[8] C. Constantinescu, E. Hashorva, L. Ji, Archimedean copulas in finite and infinite dimensions, with application to ruin problems, Insur. Math. Econom. 49 (2011) 487–495.

[9] B. de Finetti, Funzione caratteristica di un fenomeno aleatorio, Mem. Reale Accad. Naz. Lincei 4 (1930) 86–133.

[10] B. de Finetti, Sur la condition d'équivalence partielle, Actualités Scientifiques et Industrielles 739, Herman, Paris, 1938.

[11] P. Diaconis, Recent progress on de Finetti's notions of exchangeabiity, In: J.M. Bernardo, M.H. DeGroot, D.V. Lindley, A.F.M. Smith (Eds.), Bayesian Statistics 3, Oxford University Press, Oxford, 1988, pp. 111–125.

[12] C. Genest, M. Mesfioui, J. Schulz, A new bivariate Poisson common shock model covering all possible degrees of dependence, Statist. Probab. Lett. 140 (2018) 202–209.

[13] M.C. Jones, É. Marchand, Multivariate discrete distributions via sums and shares, J. Multivariate Anal. 171 (2019) 83–93.

[14] C. Lefèvre, S. Loisel, On multiply monotone distributions, continuous or discrete, with applications, J. Appl. Probab. 50 (2013) 827–847.

[15] C. Lefèvre, S. Loisel, S. Utev, Markov property in discrete Schur-constant models, Methodol. Comput. Appl. Probab. 20 (2018) 1003–1012.

[16] J.-F. Mai, M. Scherer, N. Shenkman, Multivariate geometric distributions, (logarithmically) monotone sequences, and infinitely divisible laws, J. Multivariate Anal. 115 (2013) 457–480.

[17] A.J. McNeil, J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions, Ann. Statist. 37 (2009) 3059–3097.

[18] N.U. Nair, P.G. Sankaran, Modelling lifetimes with bivariate Schur-constant equilibrium distributions from renewal theory, Metron 72 (2014) 331–349.

[19] N.U. Nair, P.G. Sankaran, N. Balakrishnan, Reliability Modelling and Analysis in Discrete Time, Academic, San Diego, 2018.

[20] R.B. Nelsen, Some properties of Schur-constant survival models and their copulas, Braz. J. Probab. Stat. 19 (2005) 179–190.

[21] P. Ressel, Higher order monotonic functions of several variables, Positivity 18 (2014) 257–285.

[22] N. Shenkman, A natural parametrization of multivariate distributions with limited memory, J. Multivariate Anal. 155 (2017) 234–251.

[23] M. Sreehari, R. Vasudeva, Characterizations of multivariate geometric distributions in terms of conditional distributions, Metrika 75 (2012) 271–286.

[24] B.Q. Ta, C.P. Van, Some properties of bivariate Schur-constant distributions, Stat. Prob. Lett. 124 (2017) 69–76.

[25] R.E. Williamson, Multiply monotone functions and their Laplace transforms, Duke Math. J. 23 (1956) 189–207.