

Forecasting compositional risk allocations

Tim J. Boonen^a, Montserrat Guillen^b and Miguel Santolino^{b*}

^a University of Amsterdam

^b Riskcenter, University of Barcelona

October 8, 2018

Abstract

We analyse models for panel data that arise in risk allocation problems, when a given set of sources are the cause of an aggregate risk value. We focus on the modeling and forecasting of proportional contributions to risk over time. Compositional data methods are proposed and the time-series regression is flexible to incorporate external information from other variables. We guarantee that projected proportional contributions add up to 100%, and we introduce a method to generate confidence regions with the same restriction. An illustration is provided for risk capital allocations.

JEL classification: C02, G22, D81.

Keywords: Simplex, capital allocation, dynamic risk management.

1 Introduction

What managers have in mind when they analyse the aggregation of risks is the “diversification effect”. The different sources of risk compensate each other and the overall risk is lower than what would result from considering the sources separately.

The calculus of the contribution of each source of risk to the aggregate result corresponds exactly to the concept of *risk allocation*. A risk allocation assigns a percent value to each component, in such a way that the contribution of that source to the aggregated value is reflected by the percentage. Risk allocations in enterprise risk management distribute the contribution to the total risk by different sources. It also helps to assess risk diversification, because it allows to find the percent contribution and so to identify sources that produce all the risk. As

***Corresponding Author.** Department of Econometrics, University of Barcelona, Diagonal 690, 08034-Barcelona, Spain. Tel.:+34-93-4020484; fax: +34-93-4021983; e-mail: msantolino@ub.edu

such, risk allocations are expressed in percent of the total risk. Allocations can be used in the context of investments, business lines and other possible activities that have multiple risk components.

Predictions of risk allocations over time need to preserve the fact that they add up to one hundred per cent. This means that univariate time-series analysis of each historical time series does not guarantee that predictions will satisfy the restriction that the sum of risk allocations expressed in percent is 100%, unless a restriction is included in the process. This is precisely the problem that we study by implementing compositional data time-series analysis.

An advantage of compositional data is the reduction of the number of parameters in time-series modelling. If we only care about future compositions, we illustrate that it may be beneficial to transform the data to compositions. This reduces the dimension of the data by one, and this reduction may yield a better model performance, as we show in this paper. We introduce compositional models to actuarial science, and we differ from the literature on compositional data in two more ways. First, we compare compositional models with standard non-compositional models, and study the model performance. Second, we introduce a method to compute confidence intervals of the forecasts.

The dynamics of allocations are modeled using compositional time series regressions. Modeling multivariate time series of compositional data have gained some attention in last years (Kynclová et al., 2015; Snyder et al., 2017; Perez-Foguet et al., 2017). These previous studies focused on the expected projections. However, a major issue in risk management is the analysis of scenarios involving deviations from expectations. We propose a methodology to compute compositional confidence intervals of projections over time. Projected compositions over these confidence intervals provide valuable information to risk managers of allocations under less likely scenarios.

Applications of compositional techniques in risk-management are scarcely found in the literature. Glassman and Riddik (1996) analyse the performance of international portfolio models using log-ratio transformations. Belles-Sampera et al. (2016) show the connection between capital allocation principles and compositional data for a one-period horizon. These authors apply compositional techniques to compute the distance between a set of risk allocation solutions and to compute the distance between them, but they do not address forecasting problems. Verbelen et al. (2018) study compositional regression models in motor insurance pricing, but they do not study the time series and forecasting.

As an important application, we study risk capital allocations. Risk capital allocations problems emerge when an aggregate capital determined by risk measure needs to be allocated to the risk units within a firm. On a one-period horizon, risk capital allocations have been studied intensively in the literature, for instance, with game-theoretic and optimization-based approaches (see, among others, Tasche, 1999; Denault, 2001; Buch et al., 2011; Dhaene et al., 2012; Zaks and Tsanakas, 2014; Boonen et al., 2017).

In this paper, our focus differs from the conventional literature about risk capital allocation in two fundamental ways. First, we are interested in the dynamics of allocations over time. Second, we do not focus on the absolute value of risk capital allocations, but we are interested in the proportional contributions (*compositions*). Risk capital allocations fluctuate over time due to economic developments, but we abstract from this and model explicitly how the relative risk capital allocations to sources of risk evolve over time.

We claim that the methods presented here are useful to analyse the past behaviour of relative risk capital allocations and to predict future percent values directly together with confidence intervals. Forecasting of relative risk capital allocations has not been considered before by previous authors. The analysis is then flexible enough to analyse a panel of n different sources of risk and their corresponding contributions to the overall result along time, that is over T periods. Alternatively, one can forecast the underlying time-series in order to determine a future relative risk capital allocation. This alternative method requires nested simulations, which may be time-consuming to simulate (for efficient methods to solve nested simulations, see, among others, Bauer et al., 2012; Gan and Lin, 2015). An advantage of our method is that we do not need nested simulations. We also consider the possibility of introducing external information from other variables, such as market uncertainty, the external evolution of the economy or, other indicators. Our results facilitate decision-making processes in enterprise management that need to consider the dynamics of relative risk capital allocations. This is simply done, by means of the compositional panel data regression that incorporates other external variables.

The article is organized as follows. The vector space structure of compositional data is introduced in the next section. The regression model and the confidence regions for time-series compositional data are presented in Section 3. In Section 4, risk capital allocation models are described, and we provide a numerical illustration. Concluding remarks can be found in Section 5.

2 The clr and ilr operations

In this section, we review notation and operators for compositional data. Unless mentioned otherwise, we assume that vectors have length $n > 1$. The set of compositions is denoted by the following simplex:

$$\mathbb{S}^n = \left\{ x = (x_1, \dots, x_n) : x_i \geq 0, \sum_{i=1}^n x_i = 1 \right\}.$$

In this paper, we only consider compositions that have no coordinates equal to zero, i.e., we focus on the set:

$$\mathbb{S}_+^n = \left\{ x = (x_1, \dots, x_n) : x_i > 0, \sum_{i=1}^n x_i = 1 \right\}. \quad (2.1)$$

For econometric modelling, it is not convenient that the range of compositional data is constrained. Therefore, we transform compositional data to allow for the range \mathbb{R} per coordinate. In this way, we can construct a regular econometric model to forecast the compositions. The first isometric transformation we use is the centered log-ratio transformation (clr)

$$\text{clr}(x) = \left(\ln \frac{x_i}{g(x)} \right)_{i=1, \dots, n}, \text{ with } g(x) = \left(\prod_{i=1}^n x_i \right)^{1/n}, x \in \mathbb{S}_+^n, \quad (2.2)$$

where the log ratio of the vector is applied component-wise (Aitchison, 1986; Aitchison and Egozcue, 2005). The inverse clr transformation is straightforward: if $x^* = \text{clr}(x)$, then $x = C[\exp(x^*)]$, where the exponential is applied by components and where C is the closing function:

$$C(x) = \frac{x}{\sum_{i=1}^n x_i}.$$

By definition, the clr transformed components sum up to zero. This may be a source of problems when doing statistical analysis, as, e.g., the variance matrix of a clr-transformed composition is singular (Pawlowsky-Glahn et al., 2011; van den Boogaart and Tolosana-Delgado, 2013).

Note that $\text{clr}(x) \in \mathbb{R}^n$. We define the isometric log-ratio transformation, denoted by $\text{ilr}(x) \in \mathbb{R}^{n-1}$, as follows. There exist an isometric linear mapping between the simplex \mathbb{S}_+^n and the $(n-1)$ -dimensional Euclidean space \mathbb{R}^{n-1} . This mapping is called the isometric log-ratio transformation (ilr).

The isometry is constructed by representing the result in a basis of the $(n-1)$ -dimensional image space of the clr transformation. This is constructed by taking an orthonormal basis of \mathbb{R}^n including the vector $v_n = (1, \dots, 1)$, i.e., some $(n-1)$ linearly independent vectors $\{v_1^*, \dots, v_{n-1}^*\}$. In computational terms, one can arrange the vectors $\{v_j^*\}$ by columns in a $n \times (n-1)$ -element matrix, denoted by V (a *contrast* matrix), with the following properties:

- V is a quasi-orthonormal matrix, as

$$V^T \cdot V = 1_{n-1}, \text{ and } V \cdot V^T = 1_n - \frac{1}{n} 1_{n \times n}, \quad (2.3)$$

where 1_n is the $n \times n$ identity matrix, and $1_{n \times n}$ is a $n \times n$ matrix full of ones.

- The columns of V sum up to zero,

$$1 \cdot V = 0, \quad (2.4)$$

where 1 indicates a row vector of ones.

Thanks to these properties, we can find simple expressions to have an equivalence between coordinates ζ and compositions $x \in \mathbb{S}_+^n$:

$$\begin{aligned} \text{clr}(x) \cdot V &= \ln(x) \cdot V := \text{ilr}(x), \\ \text{ilr}^{-1}(\zeta) &= C[\exp(\zeta \cdot V^T)]. \end{aligned}$$

Through these expressions, we also defined explicitly the so-called isometric log-ratio transformation (ilr): this is nothing else than a transformation that provides the coordinates of any composition with respect to a given orthonormal basis. There are as many ilr as orthonormal basis can be defined, as well as matrices V satisfying (2.3) and (2.4).

Note that the operation ilr depends on V , and in the sequel of this paper we will fix V as follows. The sequential binary partition proposed by Egozcue et al. (2003) has been broadly used because it allows the interpretation in terms of grouped parts of the composition. To create the first coordinate, the parts of the composition are divided into two groups. Parts in the first group are coded $+1$ and placed in the numerator. Parts in the second group are coded -1 and placed in the denominator. The first coordinate describes the balance between these two groups. In the second and following steps, each group is in turn divided into two groups. The splitting process of a group stops when it only has one component, which is coded 0 in the subsequent steps.

Each coordinate describes the balance between two groups: $+1$ parts and -1 parts. For this reason, coordinates are also called balances. In step k , coordinate $\text{ilr}(x)_k$ (balance z_k) is created as follows. The r_k parts of the first group and s_k parts of the second group have elements $(x_{n_1}, x_{n_2}, \dots, x_{n_{r_k}})$ and $(x_{d_1}, x_{d_2}, \dots, x_{d_{s_k}})$, respectively, where r_k indicates the number of parts that takes value $+1$ and s_k the number of parts that takes value -1 in step k . The isometric log-ratio transformation $\text{ilr}(x)_k$ is given by

$$z_k = \text{ilr}(x)_k = \sqrt{\frac{r_k \cdot s_k}{r_k + s_k}} \ln \frac{(x_{i_1} \cdots x_{i_{r_k}})^{1/r_k}}{(x_{j_1} \cdots x_{j_{s_k}})^{1/s_k}}, \quad k = 1, \dots, n-1. \quad (2.5)$$

The number of steps to obtain a single component in each group is exactly $n-1$.

3 Method

We show the vector autoregressive model for compositional data in Section 3.1, our method to construct confidence intervals of our predictions in Section 3.2, and model selection in Section 3.3.

3.1 Regression model

Let x_t be a composition for all t . It holds that $x_t = \left(\frac{W_{1,t}}{W_t}, \dots, \frac{W_{n,t}}{W_t}\right)$, where $W_{i,t}$ is the amount allocated to the i -th source of risk at time t for $i = 1, \dots, n$, and W_t the total amount to allocate among the parts, $W_t = W_{1,t} + \dots + W_{n,t}$. The composition can be equivalently expressed as $x_t = C(\mathbf{W}_{\cdot,t})$ with $\mathbf{W}_{\cdot,t} = (W_{1,t}, \dots, W_{n,t})$. We assume that there

are no zeros in the compositional data vector. We build a regression model with compositional response¹:

$$\text{ilr}(x_t) = b + B \cdot \text{ilr}(x_{t-1}) + \varepsilon_t, \quad (3.1)$$

where ilr is the isometric log-ratio transformation, b is a vector of parameters in \mathbb{R}^{n-1} , B is an $(n-1) \times (n-1)$ matrix of parameters in the real space, and $\varepsilon_t \sim N_{n-1}(0, \Sigma_\varepsilon)$. This is a Vector Autoregressive (VAR) model, with one lag. Point predictions of compositions are invariant on the choice of the orthonormal basis V for the isometric log-ratio transformation (Snyder et al., 2017; Pawłowsky-Glahn et al., 2011).

Secondly, we propose the following model:

$$\text{ilr}(x_t) = b + B \cdot \text{ilr}(x_{t-1}) + \gamma \cdot \log(W_{t-1}) + \varepsilon_t, \quad (3.2)$$

where W_{t-1} is a control variable and γ is the $(n-1) \times 1$ vector of coefficients defined in the real space. This is a vector auto-regression model with exogenous variables collecting the total amount to allocate among the parts at time $t-1$. This exogenous variable is selected to analyse if the size of the total amount to allocate (in log scale) at time $t-1$ has explanatory capacity on $\text{ilr}(x_t)$. This model may be straightforwardly extended to a higher number of lags and/or more exogenous variables.

Finally, we compare the regression model with compositional data with a naive model. The naive model is a VAR(1) model:

$$\mathbf{W}_{\cdot,t} = c + D \cdot \mathbf{W}_{\cdot,t-1} + \varepsilon_t, \quad (3.3)$$

where c is an $(n \times 1)$ vector and D an $(n \times n)$ matrix, both defined in the real space, and ε_t has an n -dimensional multivariate normal distribution with expectation 0 and covariance matrix Σ_ε . Note that in this model, the forecasts are not compositional, so to compare results the closing function would be applied to forecasts.

The naive regression approach is also considered with the exogenous variable:

$$\mathbf{W}_{\cdot,t} = c + D \cdot \mathbf{W}_{\cdot,t-1} + \gamma \cdot \log(W_{t-1}) + \varepsilon_t, \quad (3.4)$$

where D is again an $(n \times n)$ matrix and c and γ are $(n \times 1)$ vectors.

3.2 Confidence interval of predictions

We use bootstrapping methods to compute confidence intervals of predictions. We first simulate the predictive distribution for the next r periods for the model (3.1).

At time t , the algorithm design to the next period (the first unobserved period) is as follows:

¹A geometric interpretation on the simplex \mathbb{S}^n requires the introduction of special perturbation and powering operators for compositions (Aitchison and Egozcue, 2005; Kynclová et al., 2015).

1. We simulate S matrices of estimated coefficients \hat{B}^s , based on $N(\hat{B}, \text{var}(\hat{B}))$.
2. We estimate $\text{ilr}(\hat{x}_{t+1})^s = \text{ilr}(x_t) \cdot \hat{B}^s$ for $s = 1, \dots, S$.
3. We simulate S error matrices ε^s based on $N_{n-1}(0, \Sigma_\varepsilon)$.
4. The prediction error is incorporated $\text{ilr}(\tilde{x}_{t+1})^s = \text{ilr}(\hat{x}_{t+1})^s + \varepsilon^s$ for $s = 1, \dots, S$.

With this algorithm, S simulated forecasts of $\text{ilr}(\tilde{x}_{t+1})$ are obtained. Note that parameter uncertainty is included in step 2. Steps 3 and 4 include uncertainty associated to the process (prediction). This is a common way to simulate the predictive distribution by bootstrapping. The predictive distribution could be easily estimated, for instance, by means of non-parametric methods or fitting a parametric distribution. To obtain the predicted α -confidence region, we identify the α percent of the simulated forecasts with the highest squared Mahalanobis distance to the $E[\text{ilr}(\tilde{x}_{t+1})]$, where $E[\text{ilr}(\tilde{x}_{t+1})] = \text{ilr}(x_t) \cdot \hat{B}$. The squared Mahalanobis distance of the s -th simulated forecast is computed as, $\left(\text{ilr}(\tilde{x}_{t+1})^s - E[\text{ilr}(\tilde{x}_{t+1})]\right) \cdot \Sigma_\varepsilon^{-1} \cdot \left(\text{ilr}(\tilde{x}_{t+1})^s - E[\text{ilr}(\tilde{x}_{t+1})]\right)^T$. Finally, the inverse of the ilr transformation is considered to transform $\text{ilr}(\tilde{x}_{t+1})$ to a projected composition, \tilde{x}_{t+1} .

Similarly, the algorithm design to the remaining periods is as follows:

1. We simulate S matrices of estimated coefficients \hat{B}^s , based on $N(\hat{B}, \text{var}(\hat{B}))$.
2. We estimate $\text{ilr}(\hat{x}_{t+r})^s = \text{ilr}(\tilde{x}_{t+r-1})^s \cdot \hat{B}^s$ for $s = 1, \dots, S$ and $r > 1$.
3. We simulate S error matrices ε^s based on $N_{n-1}(\mathbf{0}, \Sigma_\varepsilon)$.
4. The prediction error is incorporated $\text{ilr}(\tilde{x}_{t+r})^s = \text{ilr}(\hat{x}_{t+r})^s + \varepsilon^s$ for $s = 1, \dots, S$.

At any projected period, $t+r$ with $r > 1$, the predicted α -confidence regions are again obtained identifying the α percent of the simulated forecasts with the highest squared Mahalanobis distance to the $E[\text{ilr}(\tilde{x}_{t+r})]$, where $E[\text{ilr}(\tilde{x}_{t+r})] = E[\text{ilr}(\tilde{x}_{t+r-1})] \cdot \hat{B}$.

The same algorithm may be adapted to simulate the predictive distribution for the next r periods for models (3.2),(3.3) and (3.4). In models with exogenous variables, when $r > 1$, we need to forecast W_{t+r-1} as well. We forecast W_{t+r-1} with an AR(1) model, i.e.,

$$W_{t+r} = \alpha + \beta W_{t+r-1} + \zeta_t,$$

where α and β are scalar parameters, and $\zeta_t \sim N(0, \sigma_\zeta^2)$.

3.3 Model selection

First, the models dealing with compositional data and those dealing with allocated amounts can be compared. Both models are Vector Autoregressive (VAR) models, with or without an exogenous variable.

For such models, we can include more lags. For model selection, we consider the following two well-known information criteria, Akaike’s information criterion (AIC, Akaike 1973) and the Bayesian information criterion (BIC, Schwarz 1978):

$$\begin{aligned} \text{AIC} &= -2\ell + 2K, \\ \text{BIC} &= -2\ell + \ln(N)K, \end{aligned}$$

where ℓ is the log-likelihood with the estimated parameters, K the number of parameters, and N the number of observations. The model with the smallest AIC/BIC is selected. Yang (2005) explains the differences between the AIC and BIC and van den Boogaart and Tolosana-Delgado (2013) argue that the use of the AIC and BIC levels for model selection with compositional data is suitable. The AIC and BIC are invariant to the choice of the contrast matrix V for the ilr transformation (Kynclová et al., 2015).

Next, we test our forecast accuracy via backtesting. Assume that our data-set contains T periods. For a jump-off period $k \in \{T - h, \dots, T - 1\}$, with $h > 1$ we study the forecast performance using the mean Aitchison distance of prediction errors (MADPE):

$$\text{MADPE}(k, m) = \frac{1}{T - k} \sum_{t=k+1}^T d_a(x_t, \hat{x}_t^{k,m}), \quad (3.5)$$

where $\hat{x}_t^{k,m}$ is the at time k -forecasted composition using model m for the period t . And, d_a corresponds to the Aitchison distance and it holds that $d_a(x_t, \hat{x}_t^{k,m}) = d(\text{ilr}(x_t), \text{ilr}(\hat{x}_t^{k,m}))$, where d is the Euclidean distance (Aitchison, 1986; Egozcue et al., 2003). The forecast accuracy is best for the model with the lowest MADPE.

4 Illustration in risk capital allocations

In this section, we provide the main econometric analysis of this paper based on the data as described in Section 4.2. The two main models that we study are (3.1) and (3.2), which we will compare with the naive models in (3.3) and (3.4). The Value-at-Risk risk measure is considered for the application. The chosen allocation method is the one based on the Euler allocation rule. Results were obtained in the programming language R, with packages *vars* (Pfaff and Stigler, 2013), *robCompositions* (Templ et al., 2017), and *compositions* (van den Boogaart et al., 2015).

4.1 Euler allocation rule

A capital allocation problem arises when an amount associated to the whole has to be distributed among its parts. Many previous studies deal with the analysis of capital allocation problems in risk management where an economic amount is fixed to cover the risk (see, among

others, Furman and Zitikis, 2008; Tsanakas, 2004; Kalkbrenner, 2005; Dhaene et al., 2012; Tsanakas, 2009; Urbina and Guillen, 2014; Zaks and Tsanakas, 2014). In this section, we briefly discuss the well-known Euler allocation rule.

We assume that the amount to share among parts is associated with the positive random variable Y_t that indicates the (unknown) daily value of the aggregate investment in month t . The positive random variable in risk management can also indicate loss severity. Let us assume that the aggregate investment consists of n different assets, where $Y_{i,t}$ is the positive random variable indicating the daily investment value of the i -th asset in the month t , i.e. $Y_t = \sum_{i=1}^n Y_{i,t}$.

The value to allocate among the n parts is computed by means of a risk measure, ρ . A risk measure ρ maps random variables into real numbers. We assume that ρ is positive homogeneous, i.e., it holds that $\rho(cX) = c\rho(X)$ for all $c > 0$ and all random variables X . We want to allocate a strictly positive amount to the n assets that are \mathcal{F}_t -measurable. To take into account dependency between assets, the Euler allocation rule can be used. Following the Euler allocation rule (Tasche, 1999; Denault, 2001), in the month t the risk proportion allocated to asset i is given by

$$\frac{\frac{\partial}{\partial \lambda_i} \rho(\sum_{j=1}^n \lambda_j Y_{j,t}) \Big|_{\lambda=(1,\dots,1)}}{\rho(Y_t)}, \quad (4.1)$$

where $\rho(Y_t)$ is the value that the risk measure assigns in the month t to the daily investment value of the portfolio and λ_i indicates the number of units of the i -th asset, $i = 1, \dots, n$. The expression

$$\frac{\partial}{\partial \lambda_i} \rho(\sum_{j=1}^n \lambda_j Y_{j,t}) \Big|_{\lambda=(1,\dots,1)}$$

can be interpreted as the partial contribution of $Y_{i,t}$ to the value that the risk measure assigns to the aggregate random variable Y_t . Since the risk measure is positively homogeneous, it holds that (Tasche, 1999):

$$\rho(Y_t) = \sum_{i=1}^n \frac{\partial}{\partial \lambda_i} \rho(\sum_{i=1}^n \lambda_i Y_{i,t}) \Big|_{\lambda=(1,\dots,1)}.$$

Therefore, the Euler allocation rule guarantees full risk allocation, i.e. the sum of risk proportions obtained by expression (4.1) is equal to one. Moreover, Tasche (1999) characterizes the Euler allocation rule as the unique one satisfying a suitability property, which is a property derived from performance management.

Some popular examples of positively homogeneous risk measures are the Value-at-Risk (VaR) and the conditional Value-at-Risk (Buch et al., 2011; Cossette et al., 2012; Asimit et al., 2011). Let us consider the Value-at-Risk with confidence level $\alpha \in (0, 1)$. The α -Value-at-Risk is defined as

$$\rho(Y_t) = \text{VaR}_\alpha(Y_t) = \sup \{y \mid F_{Y_t}(y) \leq 1 - \alpha\}.$$

If the cumulative distribution function F_{Y_t} is continuous and strictly increasing then $\text{VaR}_\alpha(Y_t) = F_{Y_t}^{-1}(1 - \alpha)$, where $F_{Y_t}^{-1}$ is the quantile distribution function. The risk capital proportion allocated to the i -th asset defined in (4.1) can be expressed as:

$$\frac{E[Y_{i,t} | Y_t = \text{VaR}_\alpha(Y_t)]}{\text{VaR}_\alpha(Y_t)}. \quad (4.2)$$

The risk capital proportion defined in (4.2) can be interpreted as the i -th component of a composition at time t . In this framework, the composition at time t is $x_t = (x_{1,t}, \dots, x_{n,t})$, where $x_{i,t}$ is computed as in (4.2) for $i = 1, \dots, n$. If we assume that $(Y_{1,t}, \dots, Y_{n,t})$ is a multivariate normally distributed random vector, $(Y_{1,t}, \dots, Y_{n,t}) \sim N_n(\mu_t, \Sigma_t)$ where μ_t is the vector of means and Σ_t the covariance matrix at time t , then it holds that:

$$E[Y_{i,t} | Y_t = \text{VaR}_\alpha(Y_t)] = E[Y_{i,t}] + \frac{\text{cov}(Y_{i,t}, Y_t)}{\sqrt{\text{var}(Y_t)}} \cdot \text{VaR}_\alpha(Z) \quad (4.3)$$

where $\text{cov}(Y_{i,t}, Y_t)$ is the covariance between $Y_{i,t}$ and Y_t , $\text{var}(Y_t)$ the variance of Y_t and Z is a standard normal random variable. The Euler allocation rule defined in (4.3) can result in null or negative capitals, although in practice it barely occurs. In the application of our compositional time series model, we assume that the values of the compositional Euler allocation rule in (4.1) are strictly positive, i.e., the compositions are in \mathbb{S}_+^n .

4.2 Data

We assume that $n = 3$, and that an investor selects the S&P 500 index, the DAX index, and the CAC 40 index, respectively. We use these indexes from January 1st, 2000, to December 31st, 2016, and monitor for every working day their daily value. We consider a buy-and-hold portfolio with an initial equal investment to each index on January 1st, 2000. The original aggregate investment in the portfolio was 100,000 dollars. The daily value of the investment in each index is considered a source of risk, and the aggregated daily value of the portfolio has an overall risk, which is displayed in Figure 1.

The time series is divided in monthly periods. We fit every month (20 working days) a multivariate normal distribution to the three risk components of the past month. Then, we apply this approximation so that the aggregate risk is normally distributed, and calculate the 95%-Value-at-Risk of this approximated aggregate risk. In this manner, we obtain 210 risk measure values associated to the aggregate daily portfolio at each month, $W_t = \text{VaR}_{95\%}(Y_t)$ for $t = 1, \dots, 210$. All values are positive.

We assume that the risk capital allocation is determined by the Euler allocation rule, as defined in (4.3). In this way, we obtain each component of the vector $\mathbf{W}_{\cdot,t} = (W_{1,t}, W_{2,t}, W_{3,t})$. We construct the compositional risk capital allocation for each month according to (4.2). Our

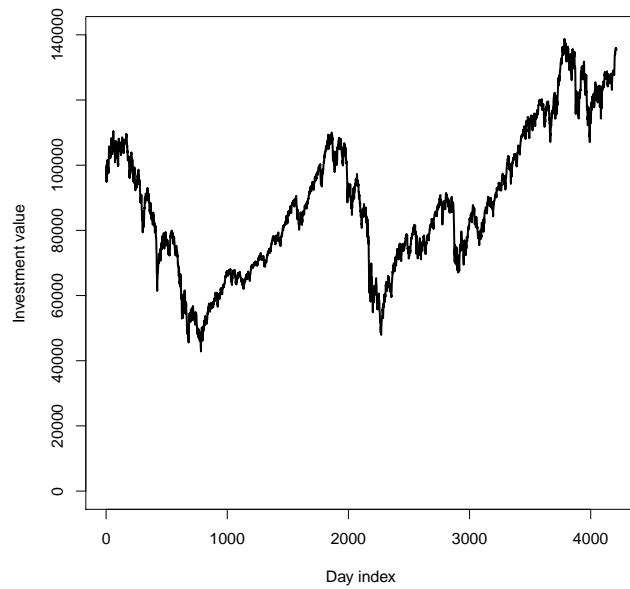


Figure 1: Evolution of the aggregate daily value of a portfolio that invests 100,000 units in the S&P 500 index, the DAX index, and the CAC 40. The investment is equally split in the three components on January 1st, 2000 and the investment goes on until December 31st, 2016.

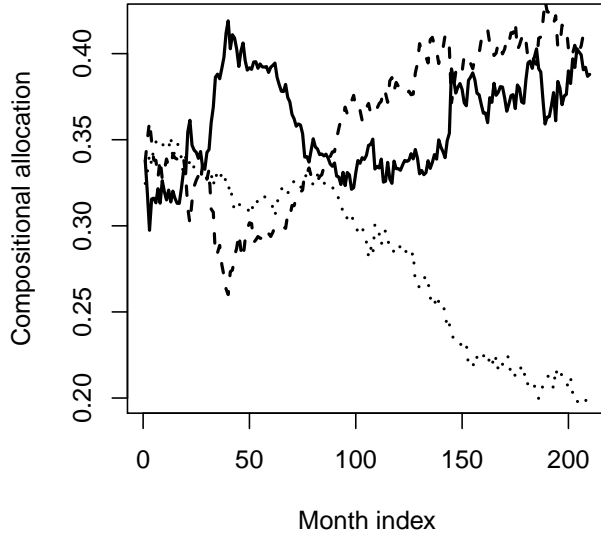


Figure 2: Evolution of the compositional data that capture the relative contribution of each index to the overall risk. Here, the solid line is the relative risk due to the S&P 500 index, the dashed line is the relative risk due to the DAX index, and the dotted line is the relative risk caused by the CAC 40.

data consists of 210 compositions. We verify that the compositional data is positive, so that $(x_{1,t}, x_{2,t}, x_{3,t}) \in \mathbb{S}_+^3$ for all $t = 1, \dots, 210$.

We display the compositional data in Figure 2. The dynamics of the relative risk capital allocations over time are shown. We observe the decreasing evolution of the third index risk contribution, which is the CAC 40 index. The classical form to display three-part compositions is by means of ternary diagrams. We provide a ternary diagram of the compositions in the simplex \mathbb{S}^3 in Figure 3.

We now express the compositional data with respect to an orthonormal basis of the underlying vector space with Euclidean structure. We use the sequential binary partition shown in (2.5). Table 1 shows the choice of the sequential binary partition for this example and Table 2 shows the associated contrast matrix of this sequential binary partition.

This particular choice of sequential binary partition leads to the following balances,

$$z_1 = \sqrt{\frac{2}{3}} \ln \frac{x_1}{\sqrt{x_2 \cdot x_3}}, \quad z_2 = \frac{1}{\sqrt{2}} \ln \frac{x_2}{x_3}. \quad (4.4)$$

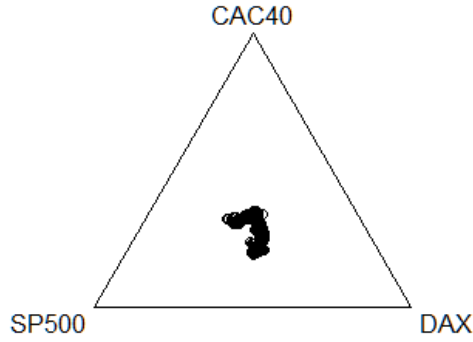


Figure 3: A ternary diagram of the compositional data in our numerical illustration.

Order	x_1	x_2	x_3	r	s
z_1	+1	-1	-1	1	2
z_2	0	+1	-1	1	1

Table 1: A 3-dimensional sequential binary partition.

x_1	x_2	x_3
$\sqrt{\frac{2}{3}}$	$-\frac{1}{2}\sqrt{\frac{2}{3}}$	$-\frac{1}{2}\sqrt{\frac{2}{3}}$
0	$\sqrt{\frac{1}{2}}$	$-\sqrt{\frac{1}{2}}$

Table 2: The contrast matrix V^T associated with the sequential binary partition in Table 1.

With this transformation, balance z_1 contains relative information (ratio) of the part x_1 to the parts x_2 and x_3 , and balance z_2 only relative information of the part x_2 to the part x_3 . Note that x_1 , x_2 and x_3 indicate the proportion of the 95%-VaR risk portfolio value associated to the investment in the S&P 500 index, the DAX index, and in the CAC 40 index, respectively.

4.3 Model selection

Now, we compare the four regression models. In this example, the explanatory variable in the model specification (3.2) and (3.4) is the risk measure value in the previous month, $W_{t-1} = \text{VaR}_{95\%}(Y_{t-1})$. First, we compare the models in (3.1) and (3.2). We find that adding more lags leads to a higher AIC and BIC, and so the formulation in (3.1) and (3.2) with only one lag is preferred. We analyse the AIC and BIC levels for the models in (3.1) and (3.2) with one lag, and find that the model (3.1) is preferred (Model (3.1): AIC= -2,040.6, BIC=-2,020.5; Model (3.2): AIC= -2,038.1, BIC=-2,011.4). For Model (3.3) and Model (3.4), we find the same model selection conclusion, the model without exogenous variable is preferred (Model (3.3): AIC= 10,364.5, BIC=10,404.6; Model (3.4): AIC= 10,366.4, BIC=10,416.5).

As the dimensions of the dependent variables are not equal for models based on compositions and models based on absolute values, the AIC/BIC criteria is not used in the selection of models in (3.1) and (3.2) with models in (3.3) and (3.4). To compare these two model design strategies, we test their forecast accuracy. First, however, we show in Table 3 the number of (unknown) parameters to be estimated that are involved in each model specification.

	Model (3.1)	Model (3.2)	Model (3.3)	Model (3.4)
Regression	6	8	12	15
Covariance	3	3	6	6
Total	9	11	18	21

Table 3: Number of (unknown) parameters in Models (3.1), (3.2), (3.3) and (3.4).

In Table 3, we show that Model (3.1) is a sparse model. We believe that the advantage of having less parameters than other models makes this choice particularly suitable.

Next, we test our forecast accuracy via backtesting. Recall that our data-set contains 210 periods. For a jump-off period $k \in \{160, \dots, 209\}$, we study the forecast performance using MADPE defined in (3.5). For models in (3.3) and (3.4), compositions are obtained with the closing function, as $x_t = C(\mathbf{W}_t)$ and $\hat{x}_t^{k,m} = C(\hat{\mathbf{W}}_t^{k,m})$, $m \in \{3,4\}$. The forecast performance of four models is shown in figure 4.

The plot in Figure 4 shows that Model (3.1) has a relative small error compared to the other models and therefore, we believe this model

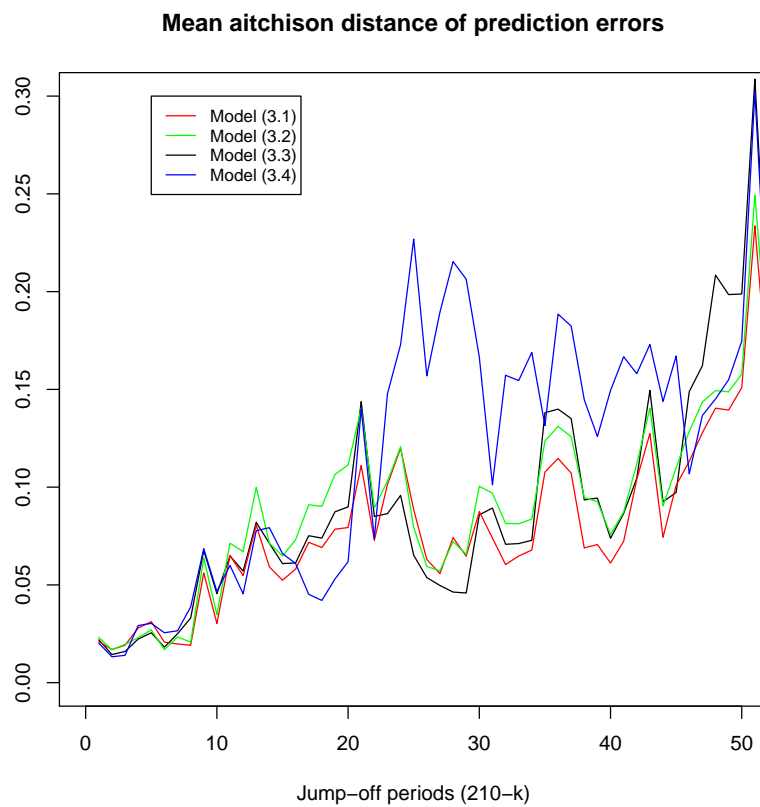


Figure 4: The mean Aitchison distance of prediction errors (MADPE) for jump-off periods 160,...,209.

provides a suitable combination of sparsity and precision.

4.4 Causality test

A frequent task in the context of multivariate time series analysis is testing for causality. For that reason, Granger causality was introduced, which represents a statistical concept that is based on prediction. This is a pairwise time series analysis that indicates if past values of a time series contain information to predict future values of another time series. In other words, we are interested in testing whether one variable could help to improve predictions of the variable of interest. This is done by including the variable with several lags, and test whether it improves statistically the predictions of the variable of interest. We check Granger causality of our ilr-transformed data and the aggregate risk level W_t in Table 4. We only find statistical evidence that $\text{ilr}(x_t)_1$ Granger causes W_t . So, effects of W_t are partially caused by changes in the composition $\text{ilr}(x_t)$, but not the other way around. Therefore, we have no evidence that W_t has an explanatory effect on $\text{ilr}(x_t)$. That is, information of the proportion of the 95%-VaR risk portfolio value associated to the investment in the S&P 500 index in relation to the investment in the DAX and CAC 40 indices is informative to predict the size of the 95%-VaR risk portfolio. The inverse does not hold².

Null hypothesis	p -value
$\text{ilr}(x_t)_1$ does not Granger cause $\text{ilr}(x_t)_2$	0.534
$\text{ilr}(x_t)_2$ does not Granger cause $\text{ilr}(x_t)_1$	0.201
$\text{ilr}(x_t)_1$ does not Granger cause W_t	0.007*
$\text{ilr}(x_t)_2$ does not Granger cause W_t	0.098
W_t does not Granger cause $\text{ilr}(x_t)_1$	0.746
W_t does not Granger cause $\text{ilr}(x_t)_2$	0.335

Table 4: Tests for Granger causality of the compositional data, and the variable W_t .

4.5 Coefficient estimates

We did not find indicators to prefer Model (3.2) above Model (3.1). We proceed our analysis with focusing only on Model 3.1. Table 5 shows the parameter estimates for the model in (3.1). The coefficient associated with the lag is statistically significant at the 0.05 significance level. So, information of the value of the balance at time $t - 1$ has explanatory capacity on the estimation of its value at time t .

²The same conclusions are obtained when $\log(W_t)$ is considered rather than W_t .

However, the coefficient estimate of the balance z_1 is not statistically significant at the 0.05 significance level when the dependent variable is the balance z_2 (and vice versa). According to this result, the relative information (ratio) of the part x_1 to the parts x_2 and x_3 , at time $(t - 1)$ would be not relevant to estimate the relative information of the part x_2 to the part x_3 at time t . That is, the proportion of the 95%-VaR risk portfolio value associated to the investment in the S&P 500 index in relation to the investment in the DAX index and in the CAC 40 index would be not informative to estimate in the next month the proportion of the 95%-VaR risk portfolio value associated to the investment in the DAX index in relation to the investment in the CAC 40 index.

Parameter	Model (3.1)	
b_1	0.003	(0.003)
b_2	0.001	(0.002)
$B_{1,1}$	0.965	(0.018)*
$B_{2,1}$	0.012	(0.009)
$B_{1,2}$	0.008	(0.012)
$B_{2,2}$	0.999	(0.006)*

Table 5: Parameter estimates of Model (3.1) and their standard errors (in brackets). * refers to significance at the 0.05 significance level.

4.6 Confidence interval of predictions

We study the confidence intervals of the forecasted relative risk capital allocations for Model 3.1. We apply the simulation approach of Section 3.2. In Figure 5, we display a three-dimensional scatter-plot of compositional simulations for the ten projected periods³. Moreover, the 5% simulated forecasts with the highest Mahalanobis distance to the point forecast in the simplex are shown in Figure 6.

The scatter-plots show that the projections present an elliptical shape. These simulations can be used to select worst-case scenarios, when using the boundaries of the resulting region. For instance, one can select four cases in the boundary of the projection and take them as extreme situations. Note that compositional simulations seem to be in the shape of ellipsoids with axes increasing with time projected periods. The center of the ellipsoids are the point compositional forecasts. We see that all simulations are compositional, as they are elements of \mathbb{S}^3 .

Some descriptive statistics can be computed for simulated forecasts. In Table 6 we compare at each projection period the correlation between simulated compositional forecasts expressed in ilr coordinates (all simulated balances) and the correlation between the 5%-of simu-

³Alternatively, simulations could be represented in a two-dimensional contour ternary diagram.

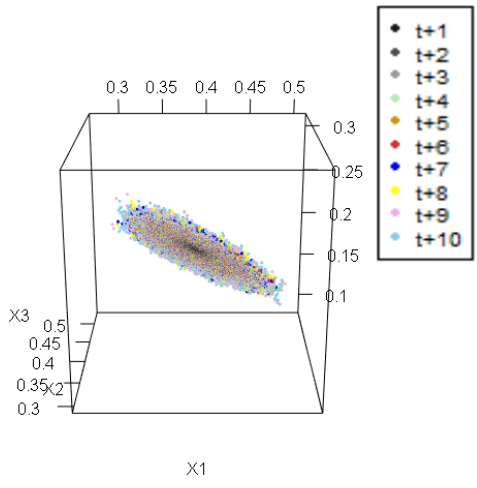


Figure 5: A scatter-plot of the simulated compositional forecasts for ten projected periods. Here, $t = 210$ is the last sample period.

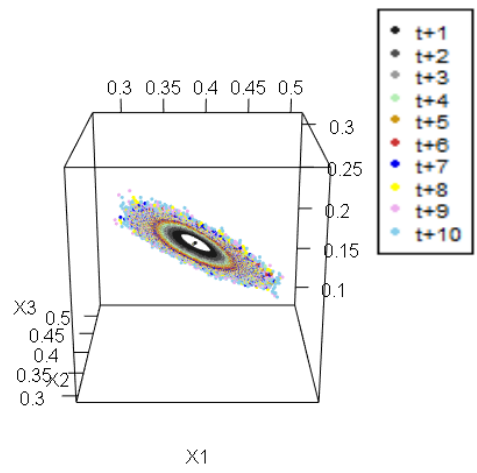


Figure 6: A scatter-plot of the 5% simulated compositional forecasts with the highest Mahalanobis distance for 10-projected periods. Here, $t = 210$ is the last sample period.

lated balances with the associated highest Mahalanobis distance (simulated balances in the 95% confidence region).

Forecasting period	Correlation (all simulations)	Correlation (5% simulations)
$t + 1$	-0.200	-0.423
$t + 2$	-0.199	-0.410
$t + 3$	-0.187	-0.390
$t + 4$	-0.180	-0.387
$t + 5$	-0.170	-0.374
$t + 6$	-0.162	-0.357
$t + 7$	-0.153	-0.325
$t + 8$	-0.144	-0.296
$t + 9$	-0.136	-0.277
$t + 10$	-0.125	-0.250

Table 6: Correlation between simulated balances for all simulations and for the 5% simulations with the highest Euclidean distance. Here, $t = 210$ is the last sample period.

The estimated correlation is weaker as the forecasting period is farther in the future. This result is expected. Note that balance z_1 is based on the ratio of x_1 with x_2 and x_3 and balance z_2 on the ratio of x_2 with x_3 . The projected value of z_1 provides less information on the projected value of z_2 as the projection is further in the future, and inverse.

In fact, a weak estimated correlation is observed in all periods when all simulations are considered. The proportion of the 95%-VaR risk portfolio value associated to the investment in the S&P 500 index in relation to the investment in the DAX index and in the CAC 40 index would be weakly correlated with the simulated prediction of the proportion of the 95%-VaR risk portfolio value associated to the investment in the DAX index in relation to the investment in the CAC 40 index. However, this conclusion varies when simulations in the confidence region are considered. In this case, a stronger correlation is observed in all periods (more than the double value). That is, the simulated projections of balances are more correlated in extreme scenarios.

5 Conclusion

We have presented a method that is useful to analyse panels of data that represent proportions of risk allocated to a give number of risk sources. In this article we present an application in risk capital allocation. For risk capital allocation problem, the firm needs to allocate proportions of the firm's risk capital to the different risk units within the firm. Of our interest is to *predict* the proportion of risk for each

unit in the next periods, and to analyse which factors influence the vector of proportions. There are many other examples where this is applicable.

The accuracy of the compositional risk allocation forecasts shown in the applications is heavily reliant on the quality of the data that the time series regression model is fitted on. For instance, in our application of risk capital allocation, we assume normality of the portfolio values within a month. A misspecification can introduce prediction errors.

Although it is less frequent, in some contexts the risk allocations can result in negative or null values (Panjer and Jing, 2001). Compositional data analysis cannot be applied in contexts involving negative values. However, compositional techniques can be still valid in case of components with zero values. The presence of zero components represents an obstacle in the application of log-ratio analysis. To solve it, alternative replacement strategies are usually applied to deal with zeros. This would then also solve the problem of having a source of risk that suddenly disappears from the scene due to a default.

Acknowledgements

The authors thank insightful comments from the editor and the anonymous reviewers that contributed to improve the manuscript. M.G. and M.S. acknowledge the support of the Spanish Ministry for grant ECO2016-76203-C2-2-P and M.S for grant ECO2015-66314-R. The authors also thank Roel Verbelen for constructive comments and suggestions.

References

- Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman and Hall.
- Aitchison, J. and J. Egozcue (2005). Compositional data analysis: Where are we and where should we be heading? *Mathematical Geology* 37(7), 829–850.
- Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*, pp. 267–281. Second International Symposium on Information Theory. Akademiai Kiado, Budapest.
- Asimit, A. V., E. Furman, Q. Tang, and R. Vernic (2011). Asymptotics for risk capital allocations based on Conditional Tail Expectation. *Insurance: Mathematics and Economics* 49(3), 310–324.
- Bauer, D., A. Reuss, and D. Singer (2012). On the calculation of the solvency capital requirement based on nested simulations. *ASTIN Bulletin* 42(2), 453–499.
- Belles-Sampera, J., M. Guillen, and M. Santolino (2016). Compositional methods applied to capital allocation problems. *Journal of Risk* 19(1), 1–15.
- Boonen, T. J., A. Tsanakas, and M. V. Wüthrich (2017). Capital allocation for portfolios with non-linear risk aggregation. *Insurance: Mathematics and Economics* 72, 95–106.
- Buch, A., G. Dorfleitner, and M. Wimmer (2011). Risk capital allocation for RORAC optimization. *Journal of Banking and Finance* 35(11), 3001–3009.
- Cossette, H., M. Mailhot, and E. Marceau (2012). TVaR-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance: Mathematics and Economics* 50, 247–256.
- Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk* 4, 1–34.
- Dhaene, J., A. Tsanakas, E. A. Valdez, and S. Vanduffel (2012). Optimal capital allocation principles. *Journal of Risk and Insurance* 79(1), 1–28.
- Egozcue, J. J., V. Pawlowsky-Glahn, G. Mateu-Figueras, and C. Barcelo-Vidal (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology* 35(3), 279–300.
- Furman, E. and R. Zitikis (2008). Weighted risk capital allocations. *Insurance Mathematics and Economics* 43(2), 263–269.

- Gan, G. and X. Lin (2015). Valuation of large variable annuity portfolios under nested simulation: A functional data approach. *Insurance: Mathematics and Economics* 62, 138–150.
- Glassman, D. and L. Riddik (1996). Why empirical international portfolio models fail: evidence that model misspecification creates home asset bias. *Journal of International Money and Finance* 15(2), 275–312.
- Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance* 15(3), 425–437.
- Kynclová, P., P. Filzmoser, and K. Hron (2015). Modeling compositional time series with vector autoregressive models. *Journal of Forecasting* 34(4), 303–314. for.2336.
- Panjer, H. and J. Jing (2001). Solvency and Capital allocation. Technical report, Institute of Insurance and Pension Research Research Report 01-14, University of Waterloo.
- Pawlowsky-Glahn, V., J. Egozcue, and R. Tolosana-Delgado (2011). *Lecture Notes on Compositional Data Analysis*. University of Girona.
- Perez-Foguet, A., R. Gine-Garriga, and M. Ortego (2017). Compositional data for global monitoring: The case of drinking water and sanitation. *Science of The Total Environment* 590(Supplement C), 554 – 565.
- Pfaff, B. and M. Stigler (2013). *vars: VAR Modelling*. R package version 1.5–2.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* 6(2), 461–464.
- Snyder, R. D., J. K. Ord, A. B. Koehler, K. R. McLaren, and A. N. Beaumont (2017). Forecasting compositional time series: A state space approach. *International Journal of Forecasting* 33(2), 502 – 512.
- Tasche, D. (1999). Risk contributions and performance measurement. Preprint, TU Munich.
- Templ, M., K. Hron, and P. Filzmoser (2017). *robComposition: Robust Estimation for Compositional Data*. R package version 2.0.6.
- Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics* 35(2), 223–243.
- Tsanakas, A. (2009). To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics* 44(2), 268–277.

- Urbina, J. and M. Guillen (2014). An application of capital allocation principles to operational risk and the cost of fraud. *Expert Systems with Applications* 41(16), 7023–7031.
- van den Boogaart, K. G., R. Tolosana, and M. Bren (2015). *compositions: Compositional Data Analysis*. R package version 1.40–1.
- van den Boogaart, K. G. and R. Tolosana-Delgado (2013). *Analyzing Compositional Data with R*. Springer-Verlag Berlin Heidelberg.
- Verbelen, R., K. Antonio, and G. Claeskens (2018). Unravelling the predictive power of telematics data in car insurance pricing. Forthcoming in the *Journal of the Royal Statistical Society: Series C*.
- Yang, Y. (2005). Can the strengths of aic and bic be shared? a conflict between model identification and regression estimation. *Biometrika* 92(4), 937–950.
- Zaks, Y. and A. Tsanakas (2014). Optimal capital allocation in a hierarchical corporate structure. *Insurance: Mathematics and Economics* 56, 48–55.