Chilean Journal of Statistics
Vol. 9, No. 2, September 2018, 33-53

Compositional regression modeling under tilted normal errors: An application to a Brazilian Super League Volleyball data set

Francisco Louzada[1], Taciana K. O. Shimizu[1,*], Adriano K. Suzuki[1], Josmar Mazucheli[2], and Paulo H. Ferreira[3]

[1]Department of Applied Mathematics & Statistics, ICMC, University of São Paulo, São Carlos, SP, Brazil,
[2]Department of Statistics, State University of Marigá, Maringá, PR, Brazil,
[3]Department of Statistics, IME, Federal University of Bahia, Salvador, BA, Brazil.

(Received: April 2017 · Accepted in final form: March, 2018)

Abstract

Simplistically, compositional data are characterized by components representing proportions or fractions of a whole. In this study, we aimed to apply a compositional regression model using Additive Log-Ratio (ALR) transformation for the response variables and assuming asymmetric errors, more specifically, the tilted normal distribution. This distribution is an alternative to the skew normal distribution. The inferential procedure is based on the usual maximum likelihood estimation. A simulation study was performed to verify the asymptotic properties of the maximum likelihood estimates. Real data set on percentages of players' points in the Brazilian Super League 2014/2015 was used to illustrate the proposed methodology and also to compare our modeling with the skew normal and normal distributions.

Keywords: Asymmetry · Compositional data · Tilted normal distribution · Uncorrelated errors.

Mathematics Subject Classification: 62J05 · 62H99.

1. Introduction

Compositional data consist of vectors whose components are non-negative and represent proportions with unit-sum constraint. Such data have sample space called simplex $\mathbb{S}^D$, defined as $\mathbb{S}^D = \{(x_1, \ldots, x_D) : x_j > 0 \text{ for } j = 1, \ldots, D \text{ and } \sum_{j=1}^{D} x_j = 1\}$, where D is the number of variables (components).

Attempts of standard statistical methods applied for compositional data analysis, without considering the simplex sample space, result in inappropriate inference. The suitable methodology was developed with the contributions of Aitchison and Shen (1980) and Aitchison (1982, 1986), which is based on transformations of compositional data from restricted simplex sample space to well-defined real sample space $\mathbb{R}$. These transformations allow the use of standard statistical methods, with the possibility to back-transform to the original space (Filzmoser et al., 2009). Aitchison and Shen (1980) developed the logistic-

*Corresponding author. Email: taci_kisaki@yahoo.com.br

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

normal class of distributions by transforming the D-components vector $\mathbf{x}$ to a vector $\mathbf{y}$ in $\mathbb{R}^{D-1}$ considering the Additive Log-Ratio (ALR) function. More recently, some contributions on the theory and applications of compositional data have been developed, for instance, in Pawlowsky (2011, 2015); van den Boogaart and Tolosana-Delgado (2013).

The use of multivariate normal distribution for compositional data can be found in Hron et al. (2012) and Egozcue et al. (2011), among others. However, sometimes the assumption of symmetry is violated and it is necessary to assume appropriate distributions for the errors of the modeling. In this way, Azzalini (1985) developed the skew normal univariate distribution, which has normal distribution as a special case. Since then, several works have been developed to obtain more flexible classes. Some extensions, properties and applications of the skew normal distribution were presented by Genton (2004). The skew generalized normal distribution was proposed by Arellano et al. (2004).

In the context of the regression models, the skew normal distribution was introduced by Azzalini (2005). Guedes et al. (2014) applied the regression model with skew normal errors to data on the height of bedding plants, Arellano et al. (2005) proposed an alternative method to linear mixed models by assuming that both the random effects and the model errors follow a skew normal distribution. For the compositional data, Mateu-Figueras and Pawlowsky-Glahn (2007) introduced the skew normal distribution on the simplex, while Martins et al. (2014) presented an application of the bivariate skew normal model to compositional data.

1.1 The Tilted Normal Distribution

The tilted nornal TN distribution was proposed by Maiti and Dey (2012) and it is defined as follows. Let X be a normal random variable with mean μ and standard deviation σ. Following Marshall and Olkin (1997), the TN probability density function (pdf) can be written as

$$f(x|\mu,\sigma,\gamma) = \frac{\frac{\gamma}{\sigma}\phi(\frac{x-\mu}{\sigma})}{[1-(1-\gamma)\{1-\Phi(\frac{x-\mu}{\sigma})\}]^2},$$

where $-\infty < x, \mu < \infty$, $\gamma, \sigma > 0$, $\phi(\cdot)$ is the pdf of a standard normal distribution, $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution, and γ is the skewness parameter such that the TN density is skewed to the left if $\gamma > 1$ and to the right if $0 < \gamma < 1$; if $\gamma = 1$, this indicates a standard normal density function (Maiti and Dey, 2012). We will denote this extension by the notation $X \sim TN\,(\mu, \sigma, \gamma)$.

Figure 1 shows the behavior of the TN distribution for different values of γ. As pointed out in the Introduction section, it can accommodate both skewness to the left or right.

1.2 Aim and Organization

The aim of this paper is to perform compositional regression modeling considering the ALR transformation and assuming asymmetric and independent errors with tilted normal (TN) distribution proposed by Maiti and Dey (2012) as an alternative to normal (N) and skew normal (SN) distributions. The TN distribution is a unimodal density which is an alternative model to SN distribution for left skewed data. According to Gupta and Gupta (2008), the SN distribution is problematic when the sample size is not large enough for the estimation of the skewness parameter. Whereas the TN distribution has the N distribution as a special case and is capable of modeling both kind of skewed data (skewed to the left or right) (Maiti and Dey, 2012). This fact is relevant because such distribution becomes more flexible for modeling data, mainly when the data present skewness. Indeed, to the best

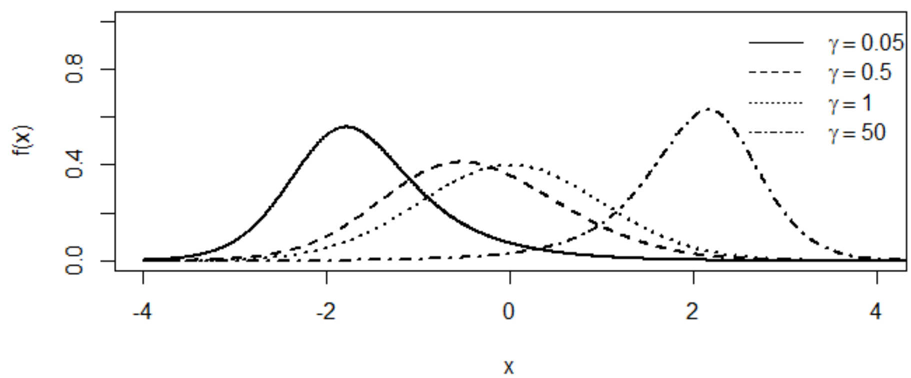

Figure 1. The pdf of TN for different values of γ ($\mu = 0$ and $\sigma = 1$).

of our knowledge this is the first paper considering compositional regression model with ALR transformation, assuming asymmetric and independent errors with tilted Normal distribution.

The paper is organized as follows. Section 2 introduces the formulation of the compositional regression model applied through the ALR transformation and assuming uncorrelated errors with TN distribution. Section 3 presents the inferential procedures. Section 4 provides the results of the applications to an artificial data set and to a real data set related to the Brazilian Men's Volleyball Super League 2014/2015. Section 5 ends the paper with some final remarks.

2. The New Compositional Regression Model

By definition, $\boldsymbol{x} = (x_1, \ldots, x_D)'$ is a compositional vector when x_j is a non-negative value and $\sum_{j=1}^{D} x_j = 1$, for $j = 1, \ldots, D$. The ALR transformation for the analysis of compositional data is given by

$$\text{ALR}: \quad \mathbb{S}^D \to \mathbb{R}^{D-1}, \quad y_{ij} = H\left(\frac{x_{ij}}{x_{iD}}\right) = \log\left(\frac{x_{ij}}{x_{iD}}\right),$$

where $H(\cdot)$ is the chosen transformation function that assures that the resulting vector has real components, x_{ij} represents the i-th observation for the j-th component, such that $x_{i1} > 0, \ldots, x_{iD} > 0$ and $\sum_{j=1}^{D} x_{ij} = 1$, for $i = 1, \ldots, n$.

The regression model assuming an ALR transformation for the response variables is given by

$$\boldsymbol{y}_i = \boldsymbol{z}_i'\boldsymbol{\beta} + \boldsymbol{\epsilon}_i,$$

where $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{id})$ is a vector $(1 \times d)$ of response variables with $d = D - 1$, $\boldsymbol{z}_i' = (z_{i1}, \ldots, z_{ip})$ is a vector $(1 \times p)$ of known independent variables where, usually, $z_{i1} = 1$ (intercept term), $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_d]$ is an unknown parameter matrix $(p \times d)$ with $\boldsymbol{\beta}_j = (\beta_{1j}, \ldots, \beta_{pj})'$, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{id})$ is a random errors vector $(1 \times d)$ whose elements $\epsilon_{ij} \sim TN(0, \sigma_j, \gamma_j)$, for $i = 1, \ldots, n$ and $j = 1, \ldots, d$. Such model represents a compositional regression model with the TN distribution for the error terms.

In this work, we approached independence among the errors of the regression model, being that the marginal functions for ϵ_{ij} are given by

$$f(\epsilon_{ij}|\sigma_j,\gamma_j) = \frac{\frac{\gamma_j}{\sigma_j}\phi(\frac{\epsilon_{ij}}{\sigma_j})}{[1-(1-\gamma_j)\{1-\Phi(\frac{\epsilon_{ij}}{\sigma_j})\}]^2}, \quad \text{for } j=1,\ldots,d.$$

3. Maximum Likelihood Estimation

The inferential procedure is based on the usual maximum likelihood estimators (MLEs). In our case, the MLEs are applied for the estimation of the parameters of the TN regression model considering compositional data. It is worth pointing out that it is assumed independence among the response variables $\boldsymbol{y}_i$.

Then, the logarithm of the likelihood function for $\boldsymbol{\theta} = (\boldsymbol{\beta}'_1,\ldots,\boldsymbol{\beta}'_d,\sigma_1,\ldots,\sigma_d,\gamma_1,\ldots,\gamma_d)'$ is given by

$$\begin{aligned} l(\boldsymbol{\theta}) = \sum_{j=1}^{d}\Bigg(& n\log\gamma_j - n\log\sigma_j + \sum_{i=1}^{n}\log\phi\left(\frac{y_{ij}-\boldsymbol{z}'_i\boldsymbol{\beta}_j}{\sigma_j}\right) \\ & - 2\sum_{i=1}^{n}\log\left[1-(1-\gamma_j)\left\{1-\Phi\left(\frac{y_{ij}-\boldsymbol{z}'_i\boldsymbol{\beta}_j}{\sigma_j}\right)\right\}\right]\Bigg). \end{aligned} \tag{1}$$

Following Migon et al. (2014), approximate $100(1-\alpha)\%$ confidence intervals for the parameters β_{mj}, σ_j and γ_j are given, respectively, by $\widehat{\beta}_{mj} \pm \xi_{\delta/2}\sqrt{Var(\widehat{\beta}_{mj})}$, $\widehat{\sigma}_j \pm \xi_{\delta/2}\sqrt{Var(\widehat{\sigma}_j)}$ and $\widehat{\gamma}_j \pm \xi_{\delta/2}\sqrt{Var(\widehat{\gamma}_j)}$, where the "hat" represents the MLE of the corresponding parameter, $\xi_{\delta/2}$ is the upper $\delta/2$ percentile of a standard normal distribution and $Var(\cdot)$ is the variance operator, for $m=1,\ldots,p$ and $j=1,\ldots,d$.

In this paper, we shall consider some model selection criteria, namely, the Akaike Information Criterion (AIC) proposed by Akaike (1974) and the Bayesian Information Criterion (BIC) proposed by Schwarz (1978). These criteria are defined by AIC $= -2\log(L)+2k$ and BIC $= -2\log(L)+k\log(n)$, where k is the number of estimated parameters, n is the sample size and L is the maximized value of the likelihood function.

4. Data Experiments

This section reports a simulation study for the compositional data and illustrates an application of the proposed methodology through ALR transformation. We considered the R software (R Core Team, 2013) with the package `maxLik` (Henningsen and Toomet, 2011) for the applications shown below. Since the R package `maxLik` is useful to introduce information on the first and second derivatives of the log-likelihood function (2), we report these quantities in the Appendix A.

4.1 Simulation Study

The simulation study was performed in order to verify the asymptotic properties of the maximum likelihood estimates. The study was based on 1,000 samples generated from the independent bivariate TN distribution with sample sizes $n = 30, 50, 70, \ldots, 250$. A dichotomized covariate z was generated through a Bernoulli distribution with probability of success 0.5, i.e. $z \sim Ber(0.5)$.

The fixed values considered for the parameters are the following ones:

i. $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 0.2$;
ii. $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 1.5$;
iii. $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2.0, \gamma_2 = 0.5$;
iv. $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2.0, \gamma_2 = 2.0$.

Figures 4 - 11 (see Appendix C) display the bias and mean square error (MSE) for the parameter estimates according to the four scenarios presented above. We can observe that the estimates are asymptotically unbiased for the parameters. Moreover, when the sample size increases, the MSE values decrease.

4.2 Real Data Application

In this section, we consider a real data set to illustrate an application of the proposed methodology, where the sample corresponds to 127 players extracted from the Brazilian Volleyball Confederation (Portuguese: Confederação Brasileira de Voleibol, CBV) (CBV, 2016). The data related to proportions of the volleyball players who participated in the Brazilian Men's Volleyball Super League 2014/2015 are available in Appendix B. The methodology of compositional data was applied to the points scored by the players during all the League, in which the considered components are: attack (x_{i1}), block (x_{i2}) and serve (x_{i3}). The covariate associated to the model is z, where $z = 1$ when the player belongs either to the first or second most efficient teams (i.e. the first and second teams with the best results) at the League 2014/2015, and $z = 0$, otherwise.

Figure 2 presents a ternary diagram for the three compositions: attack, block and serve. This type of graphic is similar to scatterplots, but representing a 3-part composition using a 2-dimensional plot (van den Boogaart and Tolosana-Delgado, 2013). We can observe that proportions of attack are higher than block and serve components. For both covariate levels, it is observed a concentration of values in direction to the attack component, but with more intensity to the players who belong to one of the two most efficient teams. On the other hand, there are some values closer to the block component, corresponding to the players who do not belong to the two most efficient teams. Some space values are tending to the center of the triangle, showing a little variability.

We applied the ALR transformation to the response variables y_1 and y_2. Therefore, we have $y_{i1} = \log(x_{i1}/x_{i3})$ and $y_{i2} = \log(x_{i2}/x_{i3})$, for $i = 1, \ldots, 127$. Table 1 shows the descriptive statistics for the data set.

Table 1. Summary statistics of ALR-transformed variables - points of the volleyball players.

Transformed	Descriptive Statistics							
Variable	Min.	1st. Qu.	Median	Mean	3rd Qu.	Max.	S.D.	Skewness
$y_1 = \log(x_1/x_3)$	-0.693	1.990	2.522	2.461	3.135	4.852	1.094	-0.698
$y_2 = \log(x_2/x_3)$	0.837	0.405	0.898	0.918	1.478	3.258	0.837	0.047

For sake of comparison, the model with TN, SN and N errors were fitted to the data set. Table 2 shows the estimation results of the fitted models. Note that the results present similar values for the parameters, i.e. the parameters $\beta_{01}, \beta_{12}, \sigma_1, \sigma_2$ were significant for all three models and, moreover, the skewness parameter γ_1 was significant for the models with SN and TN errors, indicating left skewness for the variable y_1. The results of the model comparison criteria, AIC and BIC, shown in Table 3, are suggesting that the fitted regression model assuming (uncorrelated) errors with TN distribution is the adequate choice, since it provides the lowest AIC and BIC values.

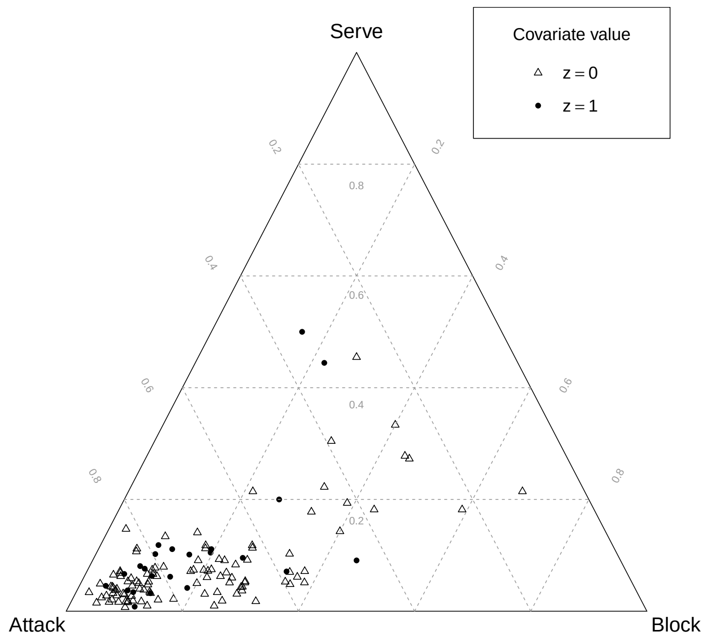

Figure 2. Ternary diagram for the components: attack, block and serve.

Following the results of Table 3, it was performed an outlier detection procedure for the TN model. According to Filzmoser et al. (2009), the Mahalanobis distance is one of the multivariate outlier detection procedures based on the estimation of the covariance structure. It can be defined for $(D-1)$-dimensional real space R^{D-1} as $\mathrm{MD}(\boldsymbol{x}_i) = [(\boldsymbol{x}_i - \boldsymbol{T})' C^{-1} (\boldsymbol{x}_i - \boldsymbol{T})]^{1/2}$, for $i = 1, \ldots, n$, where $\boldsymbol{T}$ and C are multivariate location and covariance estimators, respectively. In the case of a multivariate normal distribution, the best choices for $\boldsymbol{T}$ and C are the multivariate arithmetic mean and the sample covariance matrix, respectively. Thus, the squared Mahalanobis distances follow approximately a Chi-square distribution with $(D-1)$ degrees of freedom, i.e. $[\mathrm{MD}(\boldsymbol{x}_i)]^2 \sim \chi^2_{D-1}$, for $i = 1, \ldots, n$. Following Rousseeuw and van Zomeren (1990), data points with squared Mahalanobis distance higher than the cut-off value, like the 97.5% quantile of χ^2_{D-1}, are considered as potential outliers. This approach is the generalization of the outlier detection procedure proposed by Rousseeuw and Leroy (2003) and by Filzmoser et al. (2009), to compositional data. Thus, we note, through Figure 3, that #111 and #103 are potential influence points. The next step was to verify the impact of these observations on the estimates of the parameters. First, we refitted the model after removal of the #111 and #103 points individually, and also after both observations were removed from the set "A", that is, the original data set. The #111 point refers to a player who scored more on the block (67,9%) than attack (10.7%) and he does not belong to one of the two most efficient teams, i.e. an atypical player compared to the others who scored more on attack. And the #103 point refers to a player who scored more on the attack (82.4%), block (2.9%) and serve (14.7%).

We present the relative changes (in percentage) of each parameter estimate, defined by $RC_{\boldsymbol{\theta}_j} = \left[(\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_{j(i)})/\hat{\boldsymbol{\theta}}_j\right] \times 100$, where $\hat{\boldsymbol{\theta}}_{j(i)}$ is the MLE of $\boldsymbol{\theta}_j$ without the i-th observation, the parameter estimates and the corresponding p-values in Table 4. One can observe from this table that the estimates of parameters are sensitive under exclusion of the outstanding observations. Note that the significance of the parameter estimates does not alter after discarding the observations #111 and #103 (at a level of 5%). The considerable variations are present for the parameters β_{01}, β_{02}, γ_1 and γ_2 when the observations #111 and #103 are removed from the data. Thus, we considered the model fitted without both observations.

Table 2. Maximum likelihood estimates, standard error estimates and 95% confidence intervals for the parameters of the independent N, SN and TN distributions.

Model	Parameter	Estimate	Standard Error	95% Confidence Interval
N	β_{01}	2.522	0.106	(2.314, 2.730)
	β_{02}	0.992	0.080	(0.834, 1.149)
	β_{11}	-0.341	0.249	(-0.829, 0.148)
	β_{12}	-0.407	0.189	(-0.777, -0.037)
	σ_1	1.082	0.068	(0.949, 1.215)
	σ_2	0.819	0.051	(0.718, 0.919)
SN	β_{01}	3.636	0.158	(3.327, 3.945)
	β_{02}	0.517	0.421	(-0.307, 1.341)
	β_{11}	-0.355	0.232	(-0.810, 0.099)
	β_{12}	-0.407	0.188	(-0.776, -0.038)
	σ_1	1.551	0.146	(1.265, 1.837)
	σ_2	0.947	0.219	(0.518, 1.375)
	γ_1	-2.179	0.563	(-3.283, -1.075)
	γ_2	0.808	0.866	(-0.890, 2.506)
TN	β_{01}	0.719	0.257	(0.215, 1.222)
	β_{02}	1.226	0.393	(0.456, 1.996)
	β_{11}	-0.390	0.214	(-0.809, 0.028)
	β_{12}	-0.405	0.186	(-0.770, -0.040)
	σ_1	1.218	0.093	(1.036, 1.400)
	σ_2	0.819	0.055	(0.713, 0.927)
	γ_1	16.745	3.773	(9.351; 24.139)
	γ_2	0.604	0.493	(-0.363, 1.572)

Table 3. Comparison of the models with N, SN and TN errors.

Model	Log-likelihood	AIC	BIC
N	-345.086	702.172	719.237
SN	-340.391	696.782	719.536
TN	-338.053	692.106	714.860

A detailed study of the fundamentals on the efficient players is essential in order to the hole team becomes highly skilled, consequently, getting success in competitions. Based on the estimated parameters of the model fitted without observations #111 and #103, we can obtain the original proportions of the attack, block and serve for the players. The results are presented in Table 5, which shows that the estimated proportions were different for the three fitted models. For sake of comparison, for the TN distribution, the proportions of the block presented higher values than under the other distributions, i.e. the TN distribution pointed out that the most significant component was the block. The proportions of the serve presented little difference in relation to the associated covariate (i.e. the players who scored higher on serve belong to the first or second most efficient teams at the League).

According to this result, the serve points are important to the team become efficient at the League.

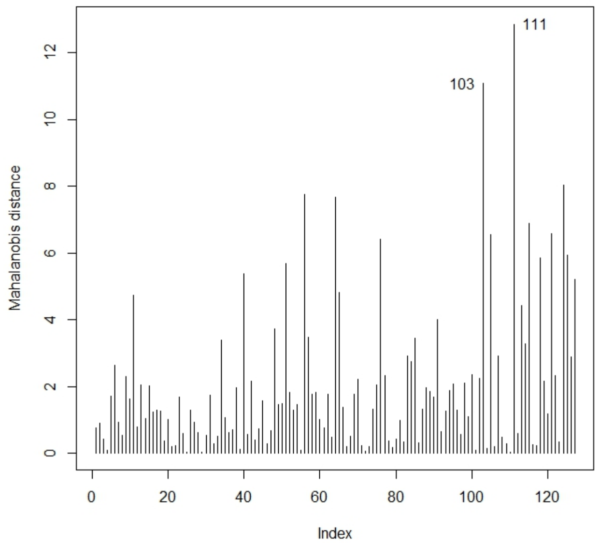

Figure 3. Mahalanobis distance for the model with TN errors.

Table 4. Estimates, relative changes [RC in %] and corresponding p-values (in parenthesis) for the parameters of the model with TN errors.

Parameter	A (Original data)	A-{#111}	A-{#103}	A-{#111,#103}
β_{01}	0.719 [-] (0.005)	0.892 [-24] (0.000)	0.702 [2] (0.018)	0.867 [-21] (0.000)
β_{02}	1.226 [-] (0.001)	1.233 [-1] (0.002)	1.417 [-16] (0.000)	1.420 [-16] (0.000)
β_{11}	-0.390 [-] (0.068)	-0.402 [-3] (0.057)	-0.400 [-3] (0.061)	-0.412 [-6] (0.050)
β_{12}	-0.405 [-] (0.030)	-0.402 [1] (0.031)	-0.431 [-6] (0.015)	-0.429 [-6] (0.016)
σ_1	1.218 [-] (0.000)	1.174 [4] (0.000)	1.221 [0] (0.000)	1.178 [3] (0.000)
σ_2	0.820 [-] (0.000)	0.823 [0] (0.000)	0.798 [3] (0.000)	0.801 [2] (0.000)
γ_1	16.745 [-] (0.000)	14.288 [15] (0.021)	17.408 [-4] (0.001)	15.048 [10] (0.000)
γ_2	0.604 [-] (0.221)	0.594 [2] (0.221)	0.407 [33] (0.190)	0.404 [33] (0.188)

We can also observe that the values of the attack proportions were very different for the three fitted models. The higher values of these proportions were observed for the SN distribution.

Thus, the results shown in Table 5 pointed out that the choice of the adequate errors distribution, in the case of regression modeling, can make the difference for coaching decisions, for instance, in which fundamentals the players are more efficient to become an efficient team at the League. In our case, if we compare the principal fundamentals of points scoring: attack, block and serve, there were differences among the fitted models. The model that provided the best fit was considering the TN distribution for errors, and it demonstrated that the points of block correspond to more than half ($> 50\%$) to a volleyball player be efficient, following by the fundamentals of attack and serve, respectively.

Table 5. Estimated original proportions of the attack, block and serve for the players, according to the errors distribution.

Component	Covariate	N	SN	TN
Attack	$z = 0$	0.771	0.934	0.317
	$z = 1$	0.760	0.926	0.299
Block	$z = 0$	0.167	0.041	0.550
	$z = 1$	0.154	0.039	0.511
Serve	$z = 0$	0.062	0.025	0.133
	$z = 1$	0.086	0.035	0.190

5. Concluding Remarks

In this paper, we have introduced a new compositional regression model considering ALR transformation and assuming independent errors with TN distribution. Inference is approached via maximum likelihood estimation. We have illustrated the proposed methodology considering a simulated data set and also a real data set on the percentages of players' points in the Brazilian Super League 2014/2015, in which it was considered a (independent) multivariate data structure. Thus, the compositional regression model with TN errors proved to be better than the SN and N distributions. Overall, the TN distribution showed to be a good alternative when the data are asymmetric, taking into account the addition of a parameter to the model. Moreover, the application of the adequate distribution for the errors in a regression modeling is important in order to provide suitable conclusions, such as the ones extracted from Table 5.

Our modeling assumes independence among the error components. However, future research needs to address this issue by considering dependence among the error components as pointed out by a referee. For instance, we could go in the direction of obtaining a multivariate extension of tilted normal distribution via from copulas (see Joe, 2014; Nelsen, 2006).

Acknowledgements

The research is partially supported by the Brazilian organizations: CNPq and FAPESP.

References

Aitchison, J., 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society. Series B 44, 139-177.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, London.

Aitchison, J., and Shen, S.M., 1980. Logistic-normal distributions: Some properties and uses. Biometrika 67, 261-272.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716-723.

Arellano-Valle, R.B., Bolfarine, H., and Lachos, V.H., 2005. Skew-normal linear mixed models. Journal of Data Science 3, 415-438.

Arellano-Valle, R.B., Gómez, H.W., and Quintana, F.A., 2004. A new class of skew normal distributions. Communications in Statistics - Theory and Methods 33, 1465-1480.

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12, 171-178.

Azzalini, A., 2005. The skew-normal distribution and related multivariate families. Scandinavian Journal of Statistics 32, 159-188.

Brazilian Volleyball Confederation (CBV). Data set of men's volleyball super league. Avaliable at: http://www.cbv.com.br/v1/superliga1415/estatisticas-novo.asp?gen=m. Accessed January 20, 2016.

Egozcue, J.J., Daunis-I-Estadella, J., Pawlowsky-Glahn, V., Hron, K., and Filzmoser, P., 2011. Simplicial regression. The normal model. Journal of Applied Probability and Statistics 6, 87-108.

Filzmoser, P., Hron, K., and Reimann, C., 2009. Principal component analysis for compositional data with outliers. Environmetrics 20, 621-632.

Genton, M.G., 2004. Skew-elliptical distributions and their applications: A journey beyond normality. Chapman & Hall, London.

Guedes, T.A., Rossi, R.M., Martins, A.B.T., Janeiro, V., Carneiro, J.W.P., 2014. Applying regression models with skew-normal errors to the height of bedding plants of Stevia rebaudiana (Bert) Bertoni. Acta Scientiarum. Technology 36, 463-468.

Gupta, R.D., and Gupta, R.C., 2008. Analyzing skewed data by power normal model. Test 17, 197-210.

Henningsen, A., Toomet, O., 2011. maxLik: a package for maximum likelihood estimation in R. Computational Statistics 26, 443-458, doi: 10.1007/s00180-010-0217-1.

Hron, K., Filzmoser, P., and Thompson, K., 2012. Linear regression with compositional explanatory variables. Journal of Applied Statistics 39, 1115-1128.

Joe, H., 2014. Dependence Modeling with Copulas. Chapman & Hall, London.

Maiti, S., and Dey, M., 2012. Tilted normal distribution and its survival properties. Journal of Data Science 10, 225-240.

Marshall, A.W., and Olkin, I., 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika 84, 641-652.

Martins, A.B.T., Janeiro, V., Guedes, T.A., Rossi, R.M., and Gonçalves, A.C.A., 2014. Modeling asymmetric compositional data. Acta Scientiarum. Technology 36, 307-313.

Mateu-Figueras, G., and Pawlowsky-Glahn, V., 2007. The skew-normal distribution on the simplex. Communications in Statistics - Theory and Methods 36, 1787-1802.

Migon, H.S., Gamerman, D., and Louzada, F., 2014. Statistical Inference: An Integrated Approach. Chapman & Hall, London.

Nelsen, R.B., 2006. An Introduction to Copulas. Springer, New York.

Pawlowsky-Glahn, V., and Buccianti, A., 2011. Compositional Data Analysis: Theory and Applications. Wiley, New York.

Pawlowsky-Glahn, V., Egozcue, J.J., and Tolosana-Delgado, R., 2015. Modeling and Analysis of Compositional Data. Wiley, New York.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rousseeuw, P.J., and van Zomeren, B.C., 1990. Unmasking multivariate outliers and leverage points. Journal of the American Statistical Association 85, 633-639.

Rousseeuw, P.J., and Leroy, A.M., 2003. Robust Regression and Outlier Detection. Wiley, New York.

Schwarz, G.E., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461-464.

van Den Boogaart, K.G., and Tolosana-Delgado, R., 2013. Analyzing Compositional Data with R. Springer, Berlin.

Appendix A. Score Functions and Hessian Matrix

In this appendix, we first show the elements of the score function vector (i.e. the score functions) of the log-likelihood function (1).

From (1), we obtain the following quantities:

$$U(\beta_{1j}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_{1j}} = \frac{1}{\sigma_j}\sum_{i=1}^{n}\left[w_{ij} + 2(1-\gamma_j)k_{ij}\phi(w_{ij})\right] z_{i1},$$

$$\vdots$$

$$U(\beta_{pj}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_{pj}} = \frac{1}{\sigma_j}\sum_{i=1}^{n}\left[w_{ij} + 2(1-\gamma_j)k_{ij}\phi(w_{ij})\right] z_{ip},$$

$$U(\sigma_j) = \frac{\partial l(\boldsymbol{\theta})}{\partial \sigma_j} = -\frac{1}{\sigma_j}\sum_{i=1}^{n}\left[1 - w_{ij}^2 - 2(1-\gamma_j)k_{ij}\phi(w_{ij})w_{ij}\right],$$

$$U(\gamma_j) = \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma_j} = \frac{n}{\gamma_j} - 2\sum_{i=1}^{n} k_{ij}\left[1 - \Phi(w_{ij})\right],$$

where $w_{ij} = \left(y_{ij} - \boldsymbol{z}_i'\boldsymbol{\beta}_j\right)/\sigma_j$ and $k_{ij} = [1 - (1-\gamma_j)\{1 - \Phi(w_{ij})\}]^{-1}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, d$.

Finally, the elements of Hessian matrix, which are computed by second derivatives of the log-likelihood function (1), are given by

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{1j}^2} = \frac{1}{\sigma_j^2}\sum_{i=1}^{n}\left\{-1 + 2(1-\gamma_j)\phi(w_{ij})k_{ij}[w_{ij} + k_{ij}(1-\gamma_j)\phi(w_{ij})]\right\} z_{i1}^2,$$

$$\vdots$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{pj}^2} = \frac{1}{\sigma_j^2}\sum_{i=1}^{n}\left\{-1 + 2(1-\gamma_j)\phi(w_{ij})k_{ij}[w_{ij} + k_{ij}(1-\gamma_j)\phi(w_{ij})]\right\} z_{ip}^2,$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{1j}\partial \sigma_j} = -\frac{2}{\sigma_j^2}\sum_{i=1}^{n}\left\{w_{ij} - (1-\gamma_j)k_{ij}\phi(w_{ij})\left[k_{ij}(1-\gamma_j)\phi(w_{ij})w_{ij} + w_{ij}^2 - 1\right]\right\} z_{i1},$$

$$\vdots$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{pj}\partial \sigma_j} = -\frac{2}{\sigma_j^2}\sum_{i=1}^{n}\left\{w_{ij} - (1-\gamma_j)k_{ij}\phi(w_{ij})\left[k_{ij}(1-\gamma_j)\phi(w_{ij})w_{ij} + w_{ij}^2 - 1\right]\right\} z_{ip},$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{1j}\partial \gamma_j} = -\frac{2}{\sigma_j}\sum_{i=1}^{n} k_{ij}\phi(w_{ij})\left[k_{ij}(1-\gamma_j)\{1 - \Phi(w_{ij})\} + 1\right] z_{i1},$$

$$\vdots$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_{pj}\partial \gamma_j} = -\frac{2}{\sigma_j}\sum_{i=1}^{n} k_{ij}\phi(w_{ij})\left[k_{ij}(1-\gamma_j)\{1 - \Phi(w_{ij})\} + 1\right] z_{ip},$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma_j\partial \sigma_j} = -\frac{2}{\sigma_j}\sum_{i=1}^{n} k_{ij}\phi(w_{ij})w_{ij}\left\{1 + (1-\gamma_j)k_{ij}\left[1 - \Phi(w_{ij})\right]\right\},$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma_j^2} = -\frac{n}{\gamma_j^2} + 2\sum_{i=1}^{n} \left[k_{ij}\{1-\Phi(w_{ij})\}\right]^2,$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma_j^2} = \frac{1}{\sigma_j^2}\sum_{i=1}^{n} \left\{1 - 3w_{ij}^2 + 2(1-\gamma_j)k_{ij}\phi(w_{ij})w_{ij}\left[k_{ij}(1-\gamma_j)\phi(w_{ij})w_{ij} + w_{ij}^2 - 2\right]\right\}.$$

Hence, the elements of the observed information matrix can be calculated by considering the Hessian matrix H, i.e. $I(\boldsymbol{\theta}) = -H$.

Appendix B. Data Set of Players' Points at the Brazilian Men's Volleyball Super League 2014/2015

Player	Attack	Block	Serve	z	Player	Attack	Block	Serve	z
1	383	37	13	0	65	89	14	1	0
2	352	35	15	1	66	58	35	7	0
3	315	37	29	1	67	72	18	9	0
4	260	73	39	1	68	69	18	11	0
5	331	20	10	0	69	56	36	6	0
6	281	19	22	0	70	67	27	3	0
7	251	32	38	1	71	79	11	5	0
8	249	32	20	0	72	69	19	7	0
9	228	31	40	0	73	78	11	6	1
10	238	16	11	1	74	63	26	4	0
11	243	9	13	0	75	82	5	4	0
12	226	21	8	1	76	84	2	3	0
13	226	25	4	0	77	80	4	4	1
14	219	21	15	0	78	73	11	3	0
15	148	87	18	1	79	64	12	5	1
16	225	17	10	0	80	68	8	4	0
17	224	16	8	0	81	52	22	4	0
18	216	16	9	0	82	64	8	3	0
19	197	28	16	0	83	61	4	5	0
20	203	23	8	1	84	63	3	1	0
21	193	27	11	0	85	25	29	12	0
22	156	49	21	0	86	48	8	7	1
23	200	22	4	0	87	52	8	2	1
24	177	25	17	0	88	56	4	1	0
25	160	35	22	1	89	55	5	1	0
26	139	58	9	0	90	31	21	4	0
27	137	58	11	0	91	39	14	1	1
28	143	37	24	0	92	33	14	6	1
29	151	37	15	1	93	37	9	2	1
30	129	51	19	1	94	37	8	1	0
31	163	29	4	0	95	22	15	8	0
32	135	45	12	0	96	39	3	3	1
33	126	47	11	0	97	38	4	2	0
34	147	12	19	0	98	38	2	1	0
35	160	13	5	1	99	22	13	4	0
36	140	18	18	1	100	22	15	2	0
37	122	45	9	0	101	25	7	4	1
38	150	21	3	0	102	16	14	5	0
39	124	37	13	0	103	28	1	5	1
40	113	54	3	0	104	24	7	2	0
41	144	16	9	0	105	11	5	16	1
42	98	60	8	0	106	26	4	2	0
43	130	21	13	0	107	13	12	6	0
44	134	14	13	1	108	26	4	1	0
45	141	17	3	0	109	20	7	2	0
46	133	18	6	0	110	21	5	2	0
47	95	40	18	0	111	3	19	6	0
48	125	10	17	0	112	22	3	2	0
49	107	36	5	0	113	9	7	7	0
50	133	10	3	0	114	10	10	2	1
51	128	14	1	0	115	5	13	4	0
52	124	8	6	1	116	15	5	2	0
53	95	21	19	0	117	15	4	1	0
54	97	29	4	1	118	5	8	5	0
55	90	26	9	0	119	14	1	1	0
56	108	14	1	1	120	8	4	3	1
57	109	6	8	0	121	4	6	5	0
58	104	11	2	0	122	8	3	3	0
59	68	40	6	0	123	8	3	1	0
60	98	9	6	0	124	3	3	5	0
61	97	11	3	0	125	3	5	3	0
62	74	31	4	0	126	4	3	2	0
63	70	29	10	0	127	3	2	4	1
64	77	26	1	1					

Appendix C. Graphs of Simulations

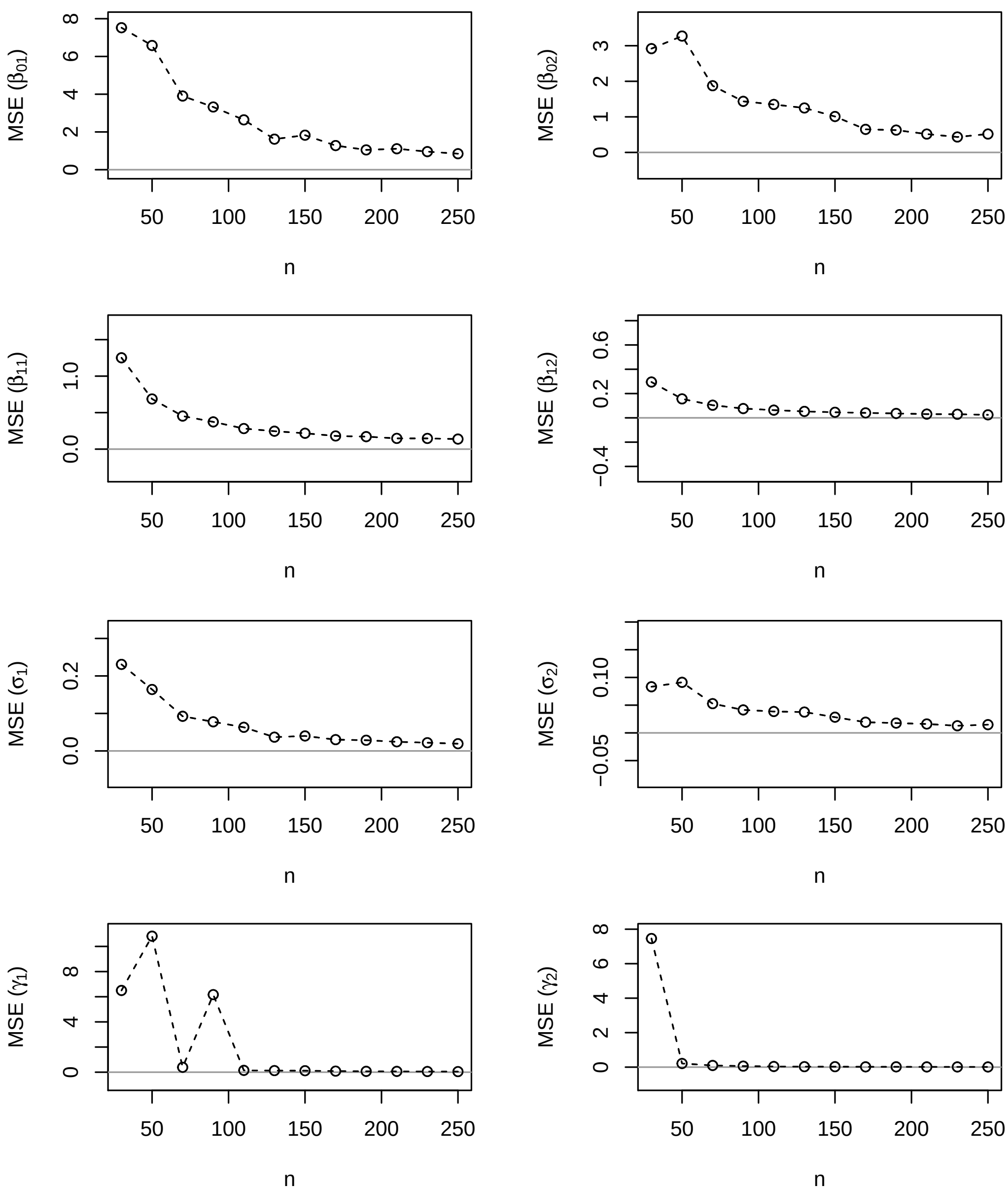

Figure 4. MSE for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 0.2$.

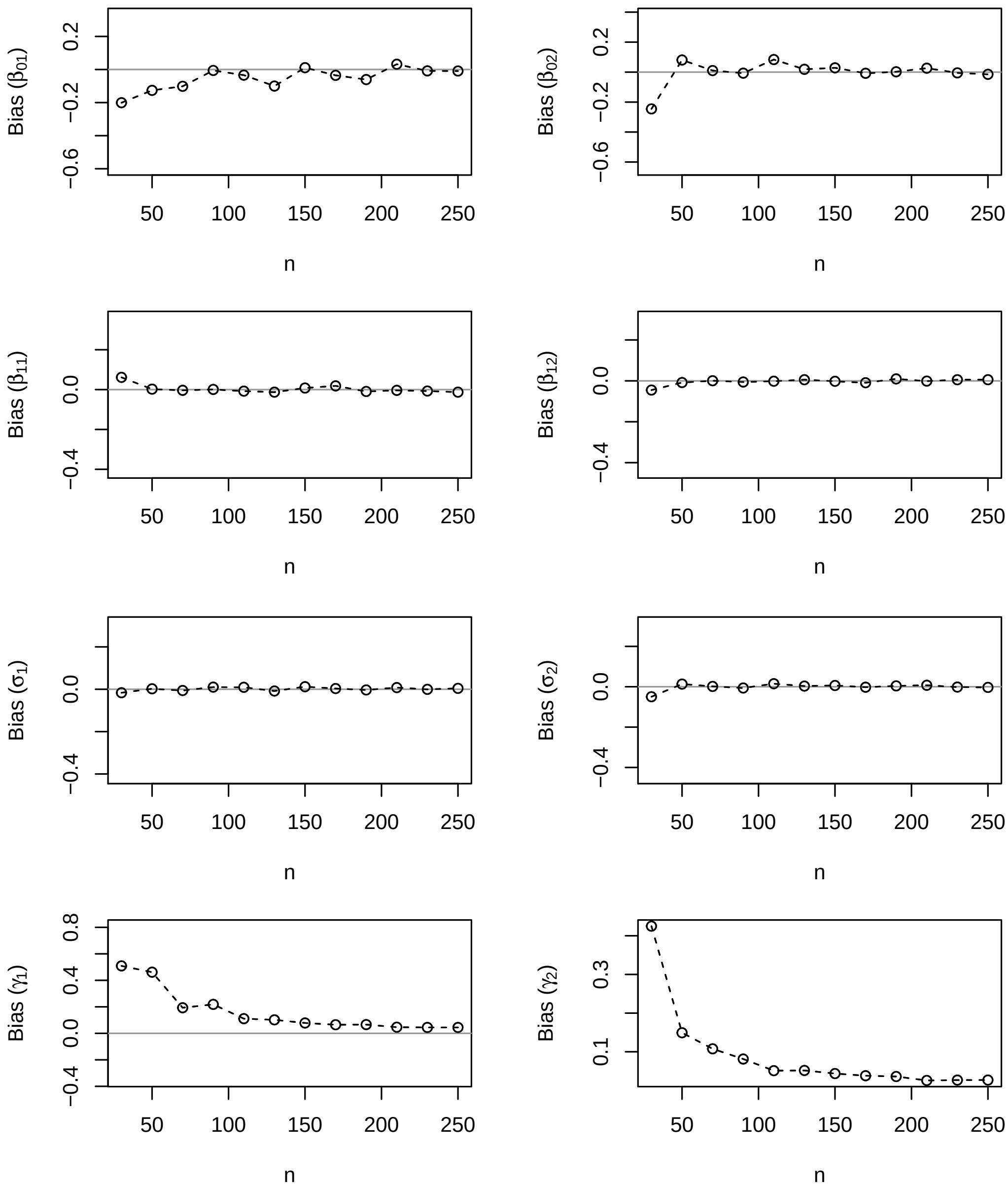

Figure 5. Bias for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 0.2$.

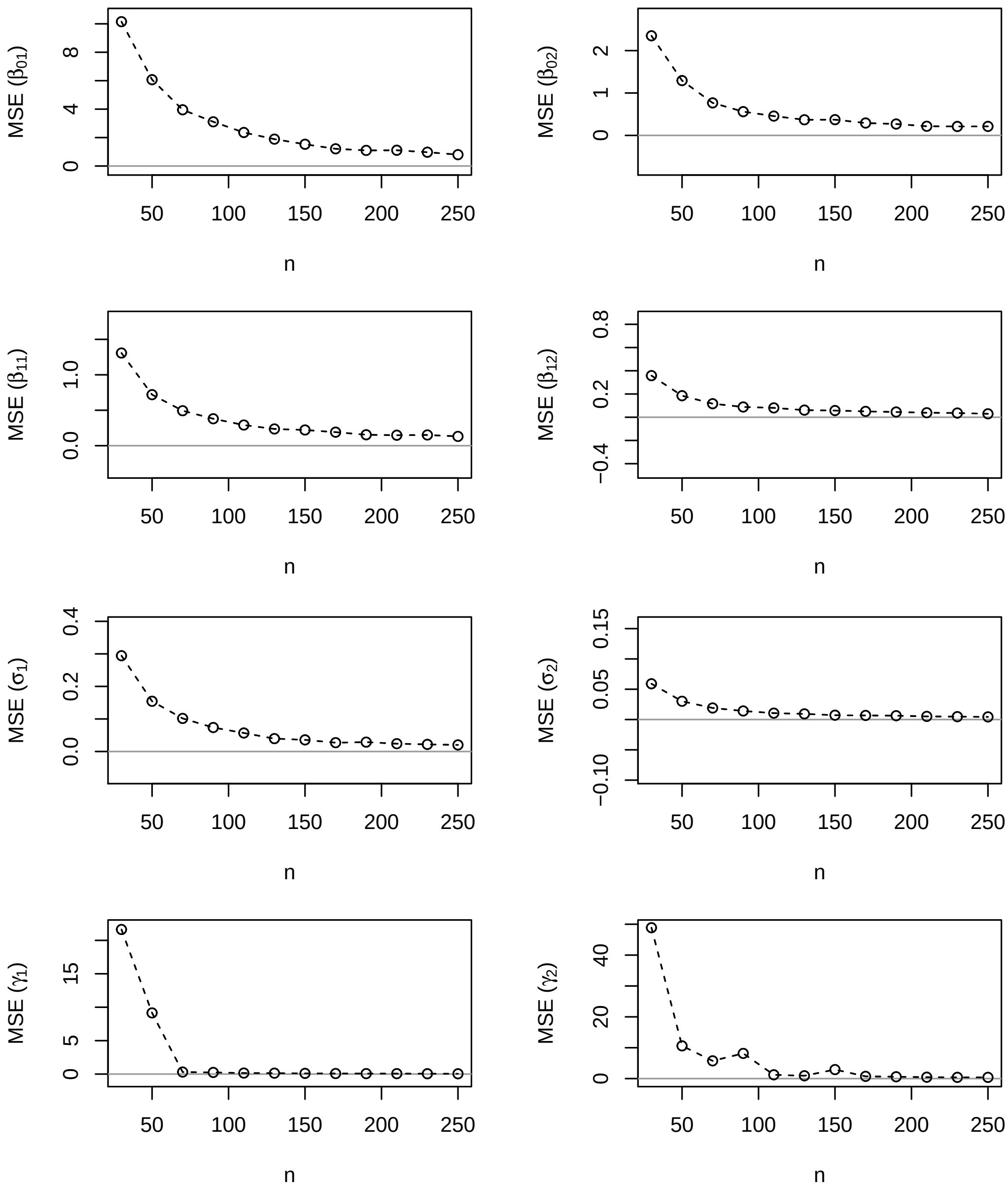

Figure 6. MSE for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 1.5$.

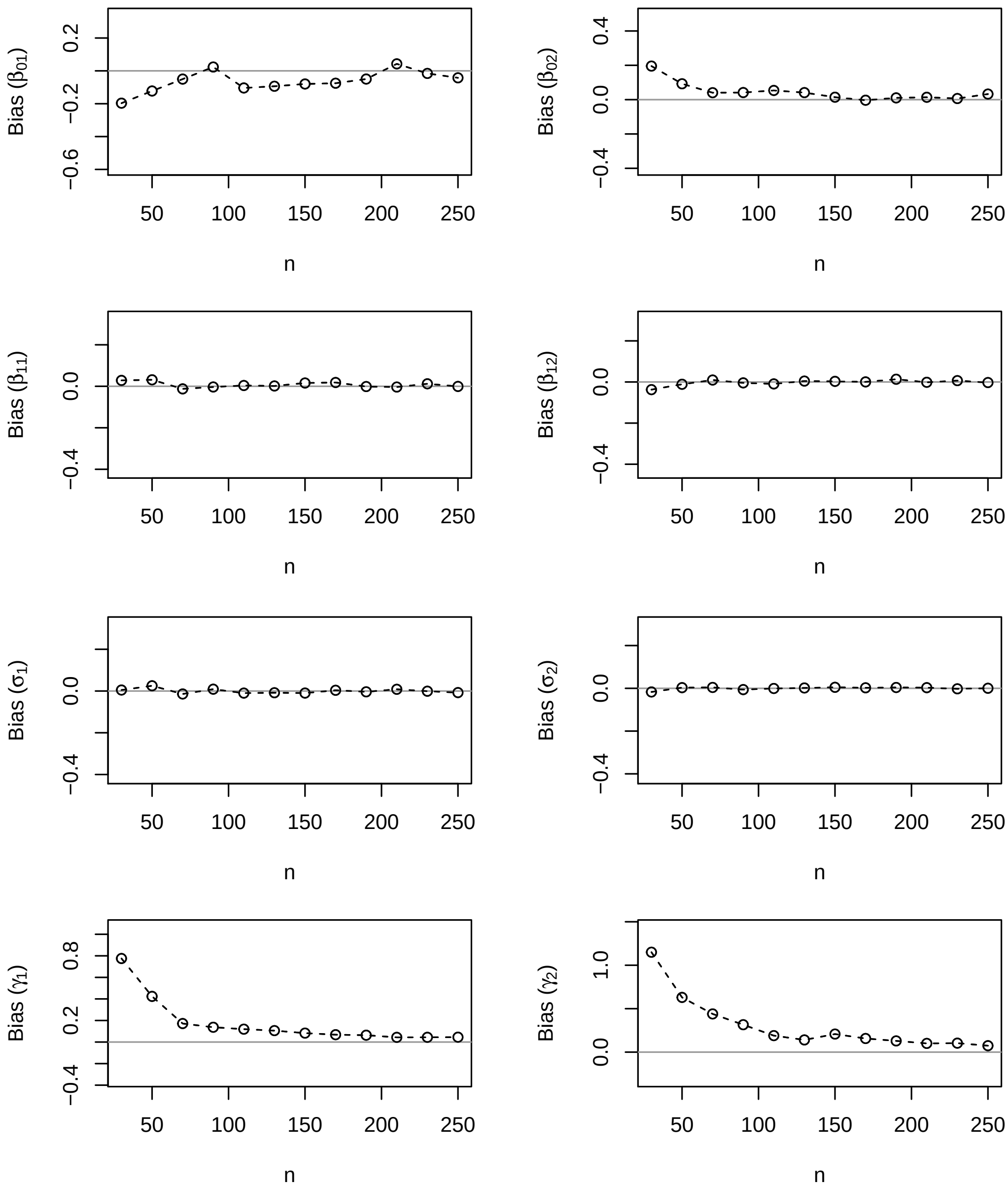

Figure 7. Bias for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 0.5, \gamma_2 = 1.5$.

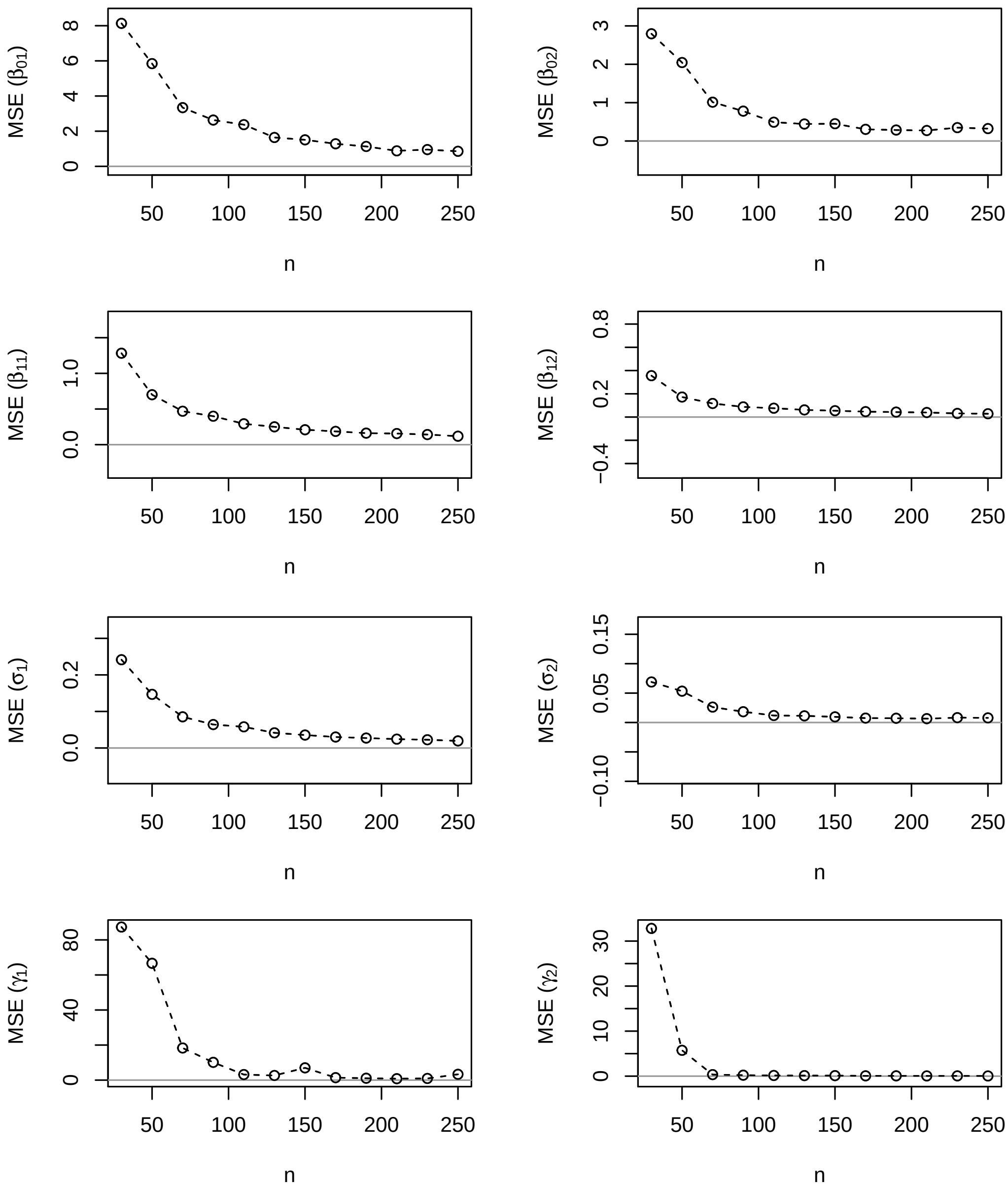

Figure 8. MSE for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2, \gamma_2 = 0.5$.

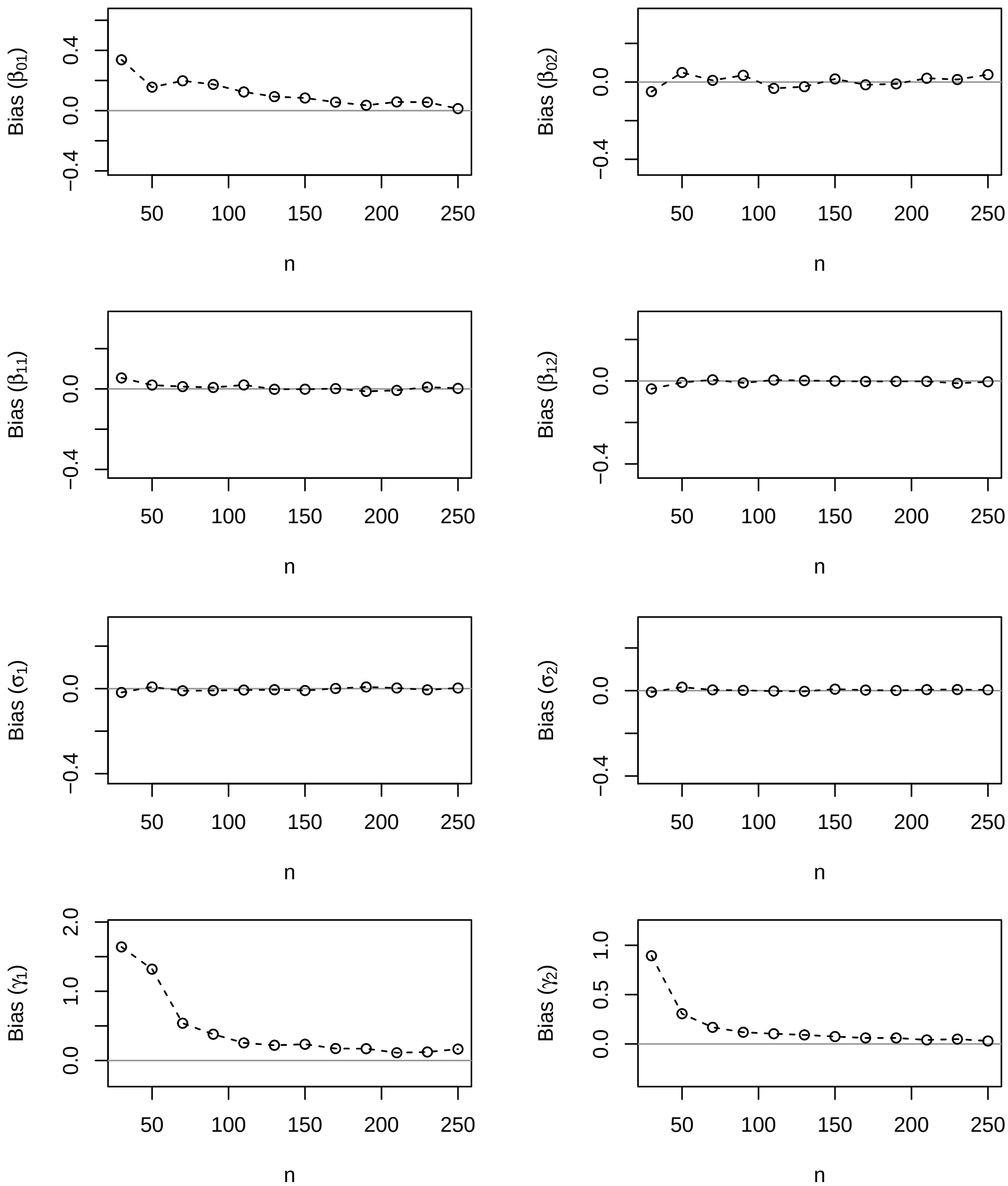

Figure 9. Bias for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2, \gamma_2 = 0.5$.

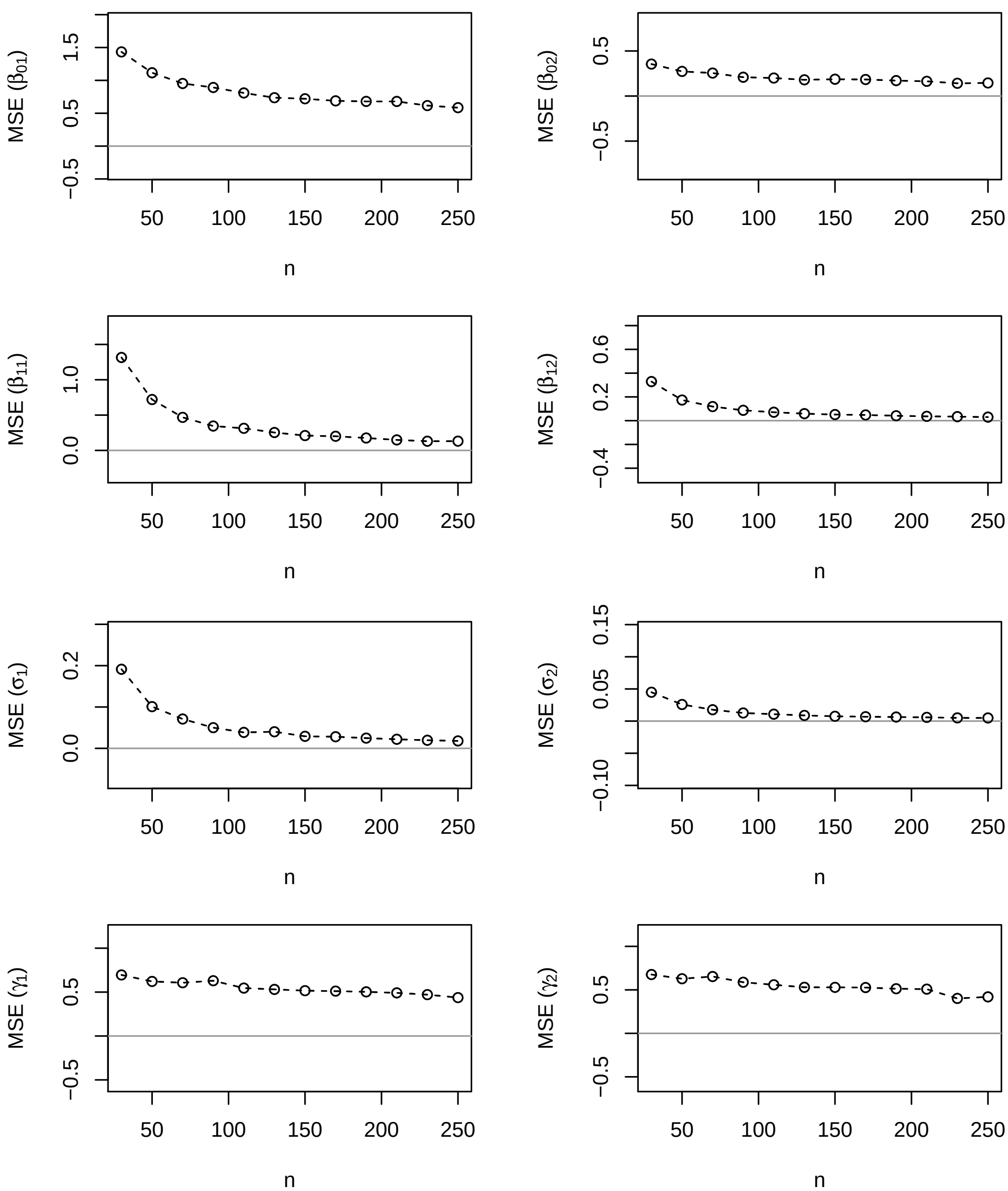

Figure 10. MSE for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2, \gamma_2 = 2$.

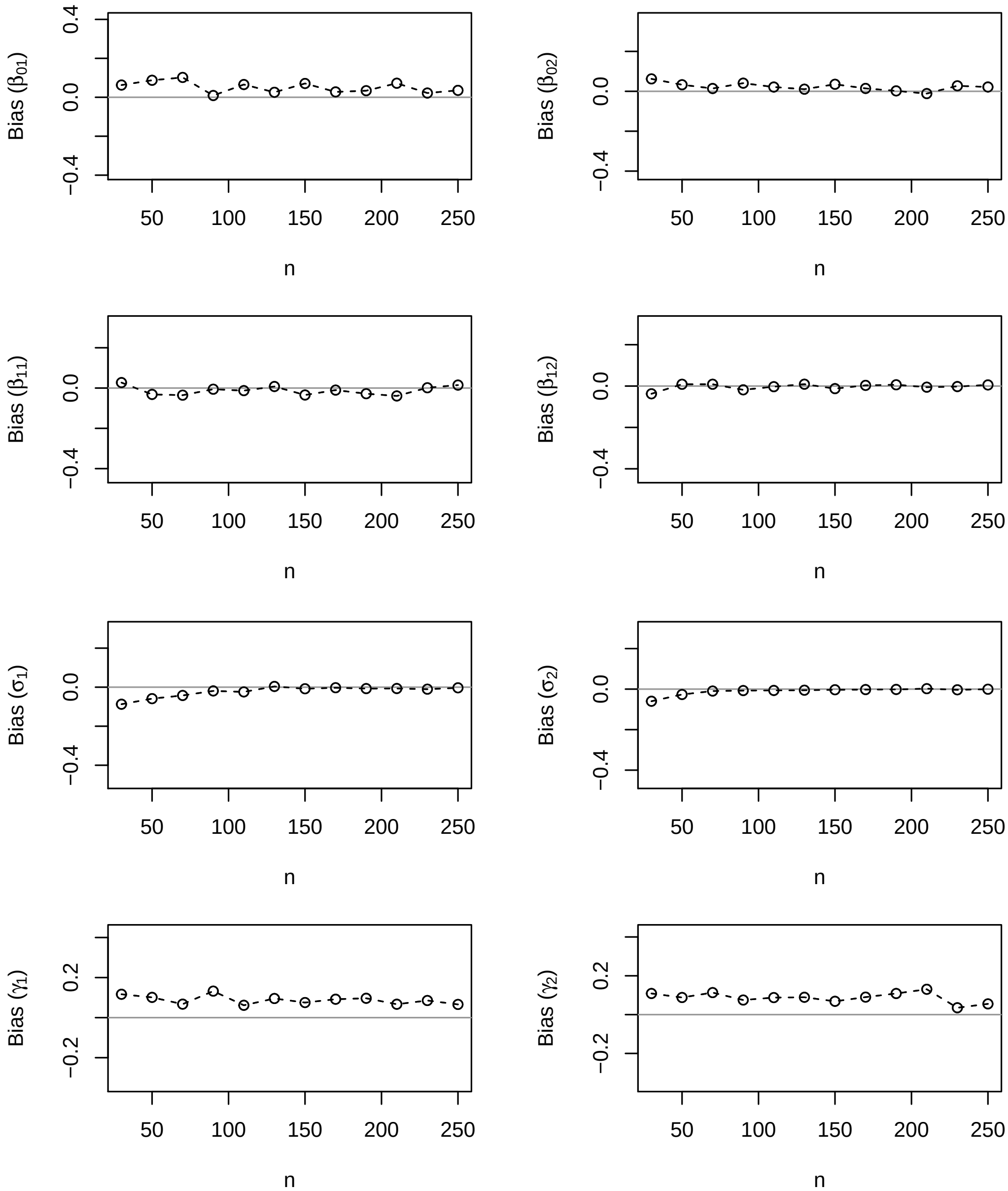

Figure 11. Bias for the estimates of $\beta_{01} = 2, \beta_{02} = -8, \beta_{11} = \beta_{12} = 1, \sigma_1 = 4, \sigma_2 = 2, \gamma_1 = 2, \gamma_2 = 2$.