Web-based Supplementary Materials for "Marginal Regression Models for Clustered Count Data Based on Zero-Inflated Conway-Maxwell-Poisson Distribution with Applications"

Hyoyoung Choo-Wosoba[1,*], Steven M. Levy[2] and Somnath Datta[3]

[1]Department of Bioinformatics and Biostatistics,
University of Louisville, Louisville, KY 40202, U.S.A.

[2]Department of Preventive & Community Dentistry
Department of Epidemiology,
University of Iowa, Iowa City, IA 52242, U.S.A.

[3]Department of Biostatistics
University of Florida, Gainesville, FL 32610, USA

[*]*email:hchoo01@cardmail.louisville.edu*

Web Appendix A.1: Estimating functions based on the MES algorithm for the zero degenerated part and the count part (CMP)

Equation (7) is the estimating function of the zero part and below is given for the derivative term with respect to γ.

$$\frac{\partial p_i}{\partial \gamma} = \frac{X_{\gamma,i} e^{X_{\gamma,i}\gamma}}{(1+e^{x_{\gamma,i}\gamma})^2},$$

and the derivative term for the count part from Equation (8) is given as

$$\frac{\partial E(y_i)^T}{\partial \beta} = \begin{pmatrix} \frac{\partial E(y_i)^T}{\partial \beta_1} \\ \frac{\partial E(y_i)^T}{\partial \beta_2} \\ \cdots \\ \frac{\partial E(y_i)^T}{\partial \beta_{p_\beta}} \end{pmatrix}_{p_\beta \times n_i}$$ and each component is calculated as

$$\frac{\partial E(y_{ij})}{\partial \beta_k} = x_{\beta,ijk}\left(\sum_{s=0}^{\infty} \frac{s^2 \lambda_{ij}^s}{(s!)^v} / Z(\lambda_{ij}, v) - \left[\sum_{s=0}^{\infty} \frac{s \lambda_{ij}^s}{(s!)^v} / Z(\lambda_{ij}, v)\right]^2\right)$$

$$Var(y_{ij}) = \sum_{s=0}^{\infty} \frac{s^2 \lambda_{ij}^s}{(s!)^v} / Z(\lambda_{ij}, v) - \left[\sum_{s=0}^{\infty} \frac{s \lambda_{ij}^s}{(s!)^v} / Z(\lambda_{ij}, v) \right]^2.$$

Here, $x_{\beta,ijk}$ is the k^{th} covariate corresponding to the count part in the j^{th} element within the i^{th} subject and $Var(y_{ij})$ is the j^{th} diagonal element of D_i.

Web Appendix A.2: Estimating a dispersion parameter, v, based on the MPL method for a CMP distribution of y

We estimating the dispersion parameter, v, from Equation (6) where

$$\frac{\partial \ell_{ij}^c}{\partial v} = \Bigg(- log(y_{ij}!) + \sum_{s=0}^{\infty} \frac{\lambda_{ij}^s log(s!)}{(s!)^v} / Z(\lambda_{ij}, v) \Bigg) (1 - u_{ij}),$$

$$\frac{\partial \ell_{ij}^c}{\partial v^2} = \Bigg(- \sum_{s=0}^{\infty} \frac{\lambda_{ij}^s (log(s!))^2}{(s!)^v} / Z(\lambda_{ij}, v) + \left[\sum_{s=0}^{\infty} \frac{\lambda_{ij}^s log(s!)}{(s!)^v} / Z(\lambda_{ij}, v) \right]^2 \Bigg) (1 - u_{ij}).$$

Web Appendix B.1: Estimating a correlation coefficient, $\boldsymbol{\delta}$, based on the MES algorithm for zero-degenerated distribution of $\boldsymbol{u}$

Estimation of the common off-diagonal correlation coefficient δ is carried out from Equation (9) leading to

$$\widehat{\delta} = \frac{\frac{1}{N*}\sum_{i=1}^{N}\sum_{s<t}\frac{(u_{is}-p_{is})(u_{it}-p_{it})}{\sqrt{p_{is}(1-p_{is})p_{it}(1-p_{it})}}}{\frac{1}{N_{total}}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{(u_{ij}-p_{ij})^2}{p_{ij}(1-p_{ij})}},$$

where $N* = \sum_{i=1}^{N} n_i(n_i-1)/2$ and $N_{total} = \sum_{i=1}^{N} n_i$. See, e.g., Kong *et al.* (2014).

Web Appendix B.2: Estimating a correlation coefficient, $\boldsymbol{\rho}$ based on the MES algorithm for a CMP distribution of $\boldsymbol{y}$

Estimation of the common off-diagonal correlation coefficient ρ is carried out from Equation (10) leading to

$$\widehat{\rho} = \frac{\frac{1}{N*}\sum_{i=1}^{N}\sum_{s<t}\frac{(1-u_{is})(1-u_{it})(y_{is}-E(y_{is}))(y_{it}-E(y_{it}))}{\sqrt{Var(y_{is})Var(y_{it})}}}{\frac{1}{N_{total}}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{(1-u_{ij})^2(y_{ij}-E(y_{ij}))^2}{Var(y_{ij})}},$$

where $N* = \sum_{i=1}^{N}\sum_{s<t}(1-u_{is})(1-u_{it})$ and $N_{total} = \sum_{i=1}^{N}\sum_{j=1}^{n_i}(1-u_{ij})^2$. See, e.g., Kong *et al.* (2014).

Web Appendix C.1: (a modified/adjusted) Sandwich covariance matrix of $\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}$ and $\widehat{\boldsymbol{v}}$

based on the MPL method

The partial derivatives of ℓ with respect to β, γ and v from Equations (12) and (13) are given follows:

- For β,

$$\frac{\partial \ell}{\partial \beta_k} = \sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}=0) x_{\beta,ijk} \frac{-(1-p_{ij})/Z(\lambda_{ij},v)^2 \sum_{s=0}^{\infty} \frac{s\lambda_{ij}^s}{(s!)^v}}{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)} +$$

$$\sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}\geq 1) x_{\beta,ijk}(y_{ij} - \sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v)).$$

$$\frac{\partial^2 \ell}{\partial\beta\partial\beta^T} = X_\beta^T diag\Big[I(y_{ij}=0) \times$$

$$\frac{\Big[\Big(2/Z(\lambda_{ij},v)^3 \sum_{s=0}^{\infty} \frac{s\lambda_{ij}^s}{(s!)^v} - (1-p_{ij})/Z(\lambda_{ij},v)^2 \sum_{s=0}^{\infty}\frac{s^2\lambda_{ij}^s}{(s!)^v}\Big)\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\} - \Big\{\frac{(1-p_{ij})}{Z(\lambda_{ij},v)^2}\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}\Big\}^2}{\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}^2}\Big] X_\beta -$$

$$X_\beta^T\, diag\Big[I(y_{ij}\geq 1)\Big[\sum_{s=0}^{\infty}\frac{s^2\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v) - \Big(\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v)\Big)^2\Big]\Big] X_\beta$$

where X_β is a $\left(\sum_{i=1}^{N} n_i \times p_\beta\right)$ matrix.

- For γ,

$$\frac{\partial \ell}{\partial \gamma_k} = \sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}=0)x_{\gamma,ijk}\frac{p_{ij}(1-p_{ij})(1-1/Z(\lambda_{ij},v))}{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)} - \sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}\geq 1)x_{\gamma,ijk}p_{ij}.$$

$$\frac{\partial^2 \ell}{\partial \gamma \partial \gamma^T} =$$

$$X_\gamma^T\, diag\Bigg[I(y_{ij}=0)\frac{p_{ij}(1-p_{ij})(1-1/Z(\lambda_{ij},v))(1-2\times p_{ij})(p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)) - p_{ij}^2(1-p_{ij})^2(1-1/Z(\lambda_{ij},v))^2}{\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}^2}\Bigg]X_\gamma -$$

$$X_\gamma^T\, diag\Big[I(y_{ij}\geq 1)p_{ij}(1-p_{ij})\Big]X_\gamma,$$

where X_γ is a $\left(\sum_{i=1}^{N} n_i \times p_\gamma\right)$matrix.

- For v,

$$\frac{\partial \ell}{\partial v} = \sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}=0)\frac{(1-p_{ij})\sum_{s=0}^{\infty}\frac{log(s!)\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v)^2}{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)} +$$

$$\sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}\geq 1)\Big(-log(y_{ij}!)+\sum_{s=0}^{\infty}\frac{\lambda_{ij}^{s} logs!}{(s!)^{v}}/Z(\lambda_{ij},v)\Big).$$

$$\frac{\partial^2 \ell}{\partial v^2}=\sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}=0)\frac{\Big[\mathcal{H}_{ij}\times\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}-\Big\{(1-p_{ij})^2\sum_{s=0}^{\infty}\frac{log(s!)\lambda_{ij}^{s}}{(s!)^{v}}\Big\}^2/Z(\lambda_{ij},v)^4\Big]}{\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}^2}-$$

$$\sum_{i=1}^{N}\sum_{j=1}^{n_i} I(y_{ij}\geq 1)\Big[\sum_{s=0}^{\infty}\frac{(log(s!))^2\lambda_{ij}^{s}}{(s!)^{v}}/Z(\lambda_{ij},v)-\Big(\sum_{s=0}^{\infty}\frac{log(s!)\lambda_{ij}^{s}}{(s!)^{v}}/Z(\lambda_{ij},v)\Big)^2\Big],$$

where

$$\mathcal{H}_{ij}=\Big[2(1-p_{ij})\Big\{\sum_{s=0}^{\infty}\frac{log(s!)\lambda_{ij}^{s}}{(s!)^{v}}\Big\}^2/Z(\lambda_{ij},v)^3-(1-p_{ij})\sum_{s=0}^{\infty}\frac{(log(s!))^2\lambda_{ij}^{s}}{(s!)^{v}}/Z(\lambda_{ij},v)^2\Big].$$

- For off-diagonal elements,

$$\frac{\partial^2 \ell}{\partial\beta\partial\gamma^T}=X_{\beta}^{T} diag\Big[I(y_{ij}=0)$$

$$\left.\frac{p_{ij}(1-p_{ij})\Big\{\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v)^2\Big\}\Big\{(p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v))+(1-p_{ij})(1-1/Z(\lambda_{ij},v))\Big\}}{\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}^2}\right]X_{\gamma}.$$

$$\frac{\partial^2\ell}{\partial\beta_k\partial v}=\sum_{i=1}^{N}\sum_{j=1}^{n_i}I(y_{ij}=0)x_{ijk,\beta}\frac{\Big\{\mathcal{W}_{ij}\times\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}-(1-p_{ij})^2/Z(\lambda_{ij},v)^4\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}\sum_{s=0}^{\infty}\frac{\lambda_{ij}^s logs!}{(s!)^v}\Big\}}{\Big\{p_{ij}+(1-p_{ij})/Z(\lambda_{ij},v)\Big\}^2}$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{n_i}I(y_{ij}\geq 1)x_{ijk,\beta}\Big[\sum_{s=0}^{\infty}\frac{slog(s!)\lambda_{ij}^s}{(s!)^v}/Z(\lambda_{ij},v)-\sum_{s=0}^{\infty}\frac{log(s!)\lambda_{ij}^s}{(s!)^v}\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}/(Z(\lambda_{ij},v))^2\Big],$$

where

$$\mathcal{W}_{ij}=\;-2(1-p_{ij})/Z(\lambda_{ij},v)^3\sum_{s=0}^{\infty}\frac{\lambda_{ij}^s logs!}{(s!)^v}\sum_{s=0}^{\infty}\frac{s\lambda_{ij}^s}{(s!)^v}-(1-p_{ij})/Z(\lambda_{ij},v)^2\sum_{s=0}^{\infty}\frac{slog(s!)\lambda_{ij}^s}{(s!)^v}.$$

$$\frac{\partial^2 \ell}{\partial \gamma_k \partial v} =$$

$$\sum_{i=1}^{N} \sum_{j=1}^{n_i} I(y_{ij} = 0) x_{\gamma, ijk} \frac{\Big\{ -p_{ij}(1-p_{ij})/Z(\lambda_{ij}, v)^2 \sum_{s=0}^{\infty} \frac{\lambda_{ij}^s logs!}{(s!)^v} \Big\} \Big\{ p_{ij} + (1-p_{ij})/Z(\lambda_{ij}, v) + (1-p_{ij})(1 - 1/Z(\lambda_{ij}, v)) \Big\}}{\Big\{ p_{ij} + (1-p_{ij})/Z(\lambda_{ij}, v) \Big\}^2}.$$

Web Appendix C.2: A modified(adjusted) sandwich covariance matrix of $\widehat{\beta}$ and $\widehat{\gamma}$ based on the MES algorithm

B_1 and B_2 in the modified sandwich variance matrices are given below. Here, B_2 accounts for variability in the u_i; see, e.g., Satten and Datta (2000).

$$B_1 = E\sum_{i=1}^{N}\begin{pmatrix} -\frac{\partial E(\boldsymbol{y}_i)^T}{\partial\beta}V_{y_i}^{-1}Diag(\mathbf{1}-\boldsymbol{u}_i)\frac{\partial E(\boldsymbol{y}_i)}{\partial\beta} & 0 & -\sum_{j=1}^{n_i}\frac{\partial^2\ell_{ij}^c}{\partial\beta\partial v} \\ 0 & -\frac{\partial \boldsymbol{p}_i(\gamma)^T}{\partial\gamma}V_{u_i}^{-1}\frac{\partial \boldsymbol{p}_i(\gamma)}{\partial\gamma} & 0 \\ -\frac{\partial}{\partial v}(\Psi_2(\beta))^T & 0 & -\sum_{j=1}^{n_i}\frac{\partial^2\ell_{ij}^c}{\partial v^2} \end{pmatrix}.$$

$$B_2 = E\sum_{i=1}^{N}\begin{pmatrix} -\frac{\partial E(\boldsymbol{y}_i)^T}{\partial\beta}V_{y_i}^{-1}Diag(\mathbf{1}-\boldsymbol{u}_i)(\boldsymbol{y}_i - E(\boldsymbol{y}_i(\beta))) \\ -\frac{\partial \boldsymbol{p}_i(\gamma)^T}{\partial\gamma}V_{u_i}^{-1}(\boldsymbol{u}_i-\boldsymbol{p}_i(\gamma)) \\ -(1-\boldsymbol{u}_i)\{-log(\boldsymbol{y}_i!)+\sum_{s=0}^{\infty}\frac{\boldsymbol{\lambda}_i^s logs!}{(s!)^v}/Z(\boldsymbol{\lambda}_i,v)\}^T \end{pmatrix}\widetilde{S}_i(\boldsymbol{u}_i,\boldsymbol{y}_i|\theta=\beta,\gamma,v) -$$

$$\sum_{i=1}^{N}\begin{pmatrix} -\frac{\partial E(\boldsymbol{y}_i)^T}{\partial\beta}V_{y_i}^{-1}Diag(\mathbf{1}-E(\boldsymbol{u}_i))(\boldsymbol{y}_i - E(\boldsymbol{y}_i(\beta))) \\ -\frac{\partial \boldsymbol{p}_i(\gamma)^T}{\partial\gamma}V_{u_i}^{-1}(E(\boldsymbol{u}_i)-\boldsymbol{p}_i(\gamma)) \\ -(1-E(\boldsymbol{u}_i))\{-log(\boldsymbol{y}_i!)+\sum_{s=0}^{\infty}\frac{\boldsymbol{\lambda}_i^s logs!}{(s!)^v}/Z(\boldsymbol{\lambda}_i,v)\}^T \end{pmatrix}\widetilde{S}_i(\boldsymbol{y}_i|\theta=\beta,\gamma,v),$$

where $\widetilde{S}_i(\boldsymbol{u}_i, \boldsymbol{y}_i|\theta = \beta, \gamma, v) = \frac{\partial log f(\boldsymbol{u}_i, \boldsymbol{y}_i|\theta)}{\partial \theta^T}$, $\widetilde{S}_i(\boldsymbol{u}_i|\theta = \beta, \gamma, v) = \frac{\partial log f(\boldsymbol{y}_i|\theta)}{\partial \theta^T} = \int log f(\boldsymbol{u}_i, \boldsymbol{y}_i|\theta^T) dF(u_{i1}, \ldots, u_{in_i}|y_{i1}, \ldots, y_{in_i})$, and $M_{\text{MES}} =$

$$\sum_{i=1}^{N} \begin{pmatrix} \frac{\partial E(\boldsymbol{y}_i)^T}{\partial \beta} V_{y_i}^{-1} Diag(\mathbf{1} - \boldsymbol{u}_i)(\boldsymbol{y}_i - E(\boldsymbol{y}_i(\beta))) \\ \frac{\partial \boldsymbol{p}_i(\gamma)^T}{\partial \gamma} V_{u_i}^{-1}(\boldsymbol{u}_i - \boldsymbol{p}_i(\gamma)) \\ \sum_{j=1}^{n_i} \frac{\partial \ell_{ij}^c}{\partial v} \end{pmatrix}_{\Big|(\widehat{\beta}, \widehat{\gamma}, \widehat{v})} \begin{pmatrix} \frac{\partial E(\boldsymbol{y}_i)^T}{\partial \beta} V_{y_i}^{-1} Diag(\mathbf{1} - \boldsymbol{u}_i)(\boldsymbol{y}_i - E(\boldsymbol{y}_i(\beta))) \\ \frac{\partial \boldsymbol{p}_i(\gamma)^T}{\partial \gamma} V_{u_i}^{-1}(\boldsymbol{u}_i - \boldsymbol{p}_i(\gamma)) \\ \sum_{j=1}^{n_i} \frac{\partial \ell_{ij}^c}{\partial v} \end{pmatrix}^T_{\Big|(\widehat{\beta}, \widehat{\gamma}, \widehat{v})} .$$

Note that we use the final estimate of the u_{ij} in all of the above.

WEB TABLE 1

Empirical bias and variance of our ZICMP estimators in a simulation study guided by the *airfreight breakage data*. The number of clusters is 30; the size of each cluster is 20. Each entry is based on 100 Monte Carlo iterations.

MPL									
		Low intraclass correlation				High intraclass correlation			
	True	Bias	SE	SE (BS)	SE (Adj_SW)	Bias	SE	SE (BS)	SE (Adj_SW)
β_0	13.8	0.2368	1.2988	1.2472	1.2059	1.3290	3.4226	2.8328	2.8935
β_1	1.3	0.0236	0.1388	0.1240	0.1200	0.1225	0.3239	0.2696	0.2741
γ_0	2	-0.0008	0.2553	0.2509	0.3063	0.0485	0.3975	0.6196	0.5227
γ_1	-3	-0.0174	0.2497	0.2428	0.2825	-0.0763	0.3960	0.6763	0.4995
v	5.7818	0.1004	0.5514	0.5201	0.5031	0.5567	1.4247	1.1796	1.2053

MES							
		Low intraclass correlation			High intraclass correlation		
	True	Bias	SE	SE (BS)	Bias	SE	SE (BS)
β_0	13.8	0.2383	1.3009	1.2592	1.3175	3.4271	2.3657
β_1	1.3	0.0230	0.1376	0.1395	0.1232	0.3217	0.4115
γ_0	2	0.0062	0.2553	0.2512	0.0587	0.4229	0.7440
γ_1	-3	-0.0141	0.2546	0.2397	-0.1044	0.3932	0.4364
v	5.7818	0.1005	0.5515	0.5338	0.5533	1.4188	0.9935

SE: Monte Carlo based empirical standard error; SE (BS): bootstrap estimated standard error, SE (Adj_SW): square root of adjusted sandwich variance estimate.

WEB TABLE 2

Empirical bias and variance of our ZICMP estimators in a simulation study guided by the *airfreight breakage data*. The number of clusters is 75; the size of each cluster is 15. Each entry is based on 100 Monte Carlo iterations.

MPL									
		Low intraclass correlation				High intraclass correlation			
	True	Bias	SE	SE (BS)	SE (Adj_SW)	Bias	SE	SE (BS)	SE (Adj_SW)
β_0	13.8	0.0991	0.9378	0.8927	0.8666	0.1844	1.8638	1.9151	1.8809
β_1	1.3	0.0105	0.0827	0.0874	0.0855	0.0177	0.1773	0.1799	0.1776
γ_0	2	0.0046	0.1683	0.1713	0.2094	0.0448	0.2876	0.2818	0.3433
γ_1	-3	-0.0223	0.1895	0.1769	0.2089	-0.0435	0.3031	0.2834	0.3359
υ	5.7818	0.0432	0.3874	0.3710	0.3608	0.0774	0.7797	0.7960	0.7827

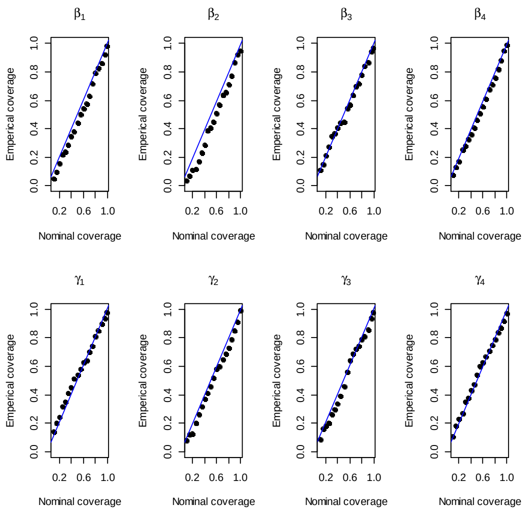
MES							
		Low intraclass correlation			High intraclass correlation		
	True	Bias	SE	SE (BS)	Bias	SE	SE (BS)
β_0	13.8	0.1004	0.9380	0.8607	0.1993	1.8663	1.7956
β_1	1.3	0.0100	0.0821	0.0860	0.0160	0.1783	0.2519
γ_0	2	0.0072	0.1679	0.1676	0.0382	0.2604	0.2848
γ_1	-3	-0.0233	0.1891	0.1731	-0.0405	0.2927	0.2826
υ	5.7818	0.0433	0.3874	0.3587	0.0825	0.7873	0.7761

SE: Monte Carlo based empirical standard error; SE (BS): bootstrap estimated standard error, SE (Adj_SW): square root of adjusted sandwich variance estimate.

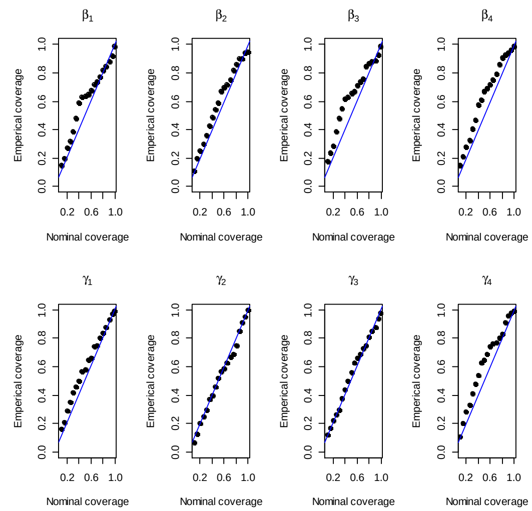
WEB FIGURE 1

Empirical coverage of the confidence intervals in a simulation study guided by the dental data of nine-year-old children from the Iowa Fluoride Study. The p-p plots are for $N = 200$ and $n = 15$ when intra-cluster correlation is low. Three sets of plots are provided for the regression parameters corresponding to the four covariates: ZICMP/MPL (upper left panel), ZICMP/MES (upper right panel), ZIP (bottom left panel) and ZIP with Adj_SW (bottom right panel).

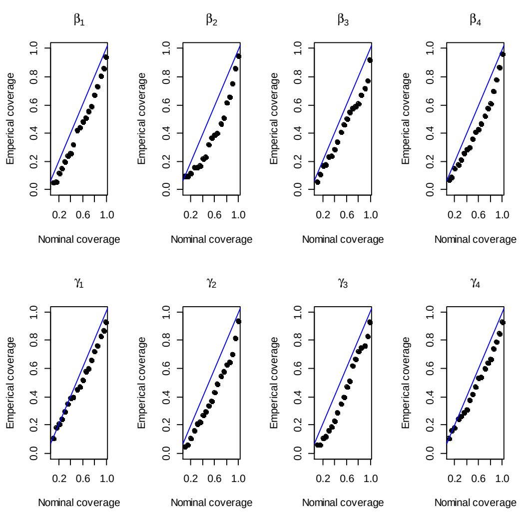
(a) Coverage probabilities based on ZICMP/MPL



(b) Coverage probabilities based on ZICMP/MES



(c) Coverage probabilities based on ZIP



(d) Coverage probabilities based on ZIP with Adj_SW