Supplementary information for: Estimators for longitudinal latent exposure models: examining measurement model assumptions

Brisa N. Sánchez[*,1], **Sehee Kim**[1], **Mary D. Sammel**[2]

A. Score equations for MLE

Let $\theta = (\theta_0, \theta_1, \theta_2, \theta_3)$ represent all model parameters, with $\theta_0 = (\alpha_y, \gamma^\top, \sigma^2_\epsilon)^\top$, and $\theta_1 = (\alpha_{\xi_i}, \alpha_x)$, $\theta_2 = (\Lambda)$, and $\theta_3 = (\Sigma_\xi, \Sigma_\delta)$. Assuming normality of ϵ_i, δ_i, and ξ_i, the observed marginal likelihood is $L(\theta) = \prod_{i=1}^N f_{\mathrm{Y|X}}(Y_i|X_i;\theta) f_{\mathrm{X}}(X_i;\theta_1,\theta_2,\theta_3)$, where $f_{\mathrm{Y|X}}, f_{\mathrm{X}}$ are normal densities. Letting $\ell(\theta) = \log L(\theta)$, and θ_{jr} represent the r^{th} unique parameter contained in $\theta_j, j = 1,2,3$, the ith's subject contribution to the likelihood score equations is

$$\frac{\partial \ell_i}{\partial \alpha_y} = (Y_i - \mu^i_{\mathrm{Y|X}})/\sigma^2_{\mathrm{Y|X}} \tag{1}$$

$$\frac{\partial \ell_i}{\partial \gamma} = \widehat{\xi}^{\mathrm{LV}}_i (Y_i - \mu^i_{\mathrm{Y|X}})/\sigma^2_{\mathrm{Y|X}} + \frac{\Sigma_{\xi|\mathrm{X}}\gamma}{2(\sigma^2_{\mathrm{Y|X}})^2}\left[(Y_i - \mu^i_{\mathrm{Y|X}})^2 - \sigma^2_{\mathrm{Y|X}}\right] \tag{2}$$

$$\frac{\partial \ell_i}{\partial \sigma^2_\epsilon} = \frac{1}{(\sigma^2_{\mathrm{Y|X}})^2}[(Y_i - \mu^i_{\mathrm{Y|X}})^2 - \sigma^2_{\mathrm{Y|X}}] \tag{3}$$

$$\frac{\partial \ell_i}{\partial \theta_{1r}} = \gamma^\top \frac{\partial \widehat{\xi}^{\mathrm{LV}}_i}{\partial \theta_{1r}} (Y_i - \mu^i_{\mathrm{Y|X}})/\sigma^2_{\mathrm{Y|X}} + \left(\frac{\partial \mu^i_{\mathrm{X}}}{\partial \theta_{1r}}\right)^\top \Sigma^{-1}_{\mathrm{X}}(X_i - \mu^i_{\mathrm{X}}) \tag{4}$$

$$\frac{\partial \ell_i}{\partial \theta_{2r}} = \frac{1}{2(\sigma^2_{\mathrm{Y|X}})^2}[(Y_i - \mu^i_{\mathrm{Y|X}})^2 - \sigma^2_{\mathrm{Y|X}}] + \gamma^\top \frac{\partial \widehat{\xi}^{\mathrm{LV}}_i}{\partial \theta_{2r}} (Y_i - \mu^i_{\mathrm{Y|X}})/\sigma^2_{\mathrm{Y|X}} \tag{5}$$

$$+\frac{1}{2}\mathrm{tr}\left\{\frac{\partial \Sigma_{\mathrm{X}}}{\partial \theta_{2r}}\Sigma^{-1}_{\mathrm{X}}[(X_i - \mu^i_{\mathrm{X}})(X_i - \mu^i_{\mathrm{X}})^\top - \Sigma_{\mathrm{X}}]\Sigma^{-1}_{\mathrm{X}}\right\} + \left(\frac{\partial \mu_x}{\partial \theta_{2r}}\right)^\top \Sigma^{-1}_{\mathrm{X}}(X_i - \mu^i_{\mathrm{X}}) \tag{6}$$

$$\frac{\partial \ell_i}{\partial \theta_{3r}} = \frac{1}{2(\sigma^2_{\mathrm{Y|X}})^2}[(Y_i - \mu^i_{\mathrm{Y|X}})^2 - \sigma^2_{\mathrm{Y|X}}] + \gamma^\top \frac{\partial \widehat{\xi}^{\mathrm{LV}}_i}{\partial \theta_{3r}} (Y_i - \mu^i_{\mathrm{Y|X}})/\sigma^2_{\mathrm{Y|X}} \tag{7}$$

$$+\frac{1}{2}\mathrm{tr}\left\{\frac{\partial \Sigma_{\mathrm{X}}}{\partial \theta_{3r}}\Sigma^{-1}_{\mathrm{X}}[(X_i - \mu^i_{\mathrm{X}})(X_i - \mu^i_{\mathrm{X}})^\top - \Sigma_{\mathrm{X}}]\Sigma^{-1}_{\mathrm{X}}\right\} \tag{8}$$

[1] *Department of Biostatistics, University of Michigan, Ann Arbor, MI USA 48109*
[2] *Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, USA*
[*] *Correspondence to: brisa@umich.edu, Department of Biostatistics, University of Michigan, Ann Arbor, MI USA 48109*

where $\widehat{\xi}_i^{\text{LV}} = E(\xi_i|X_i) = \alpha_{\xi_i} + \Sigma_\xi \Lambda_x^\top \Sigma_{\text{X}}^{-1}(X_i - \mu_{\text{X}}^i)$ is the expected value of the latent variable gi ven the items; $\mu_{\text{Y|X}}^i = \alpha_y + \gamma^\top \widehat{\xi}_i^{\text{LV}}$; $\sigma_{\text{Y|X}}^2 = \gamma^\top \Sigma_{\xi|\text{X}} \gamma + \sigma_\epsilon^2$; $\Sigma_{\xi|\text{X}} = \text{var}(\xi_i|X_i) = \Sigma_\xi - \Sigma_\xi \Lambda_x^\top \Sigma_{\text{X}}^{-1} \Lambda_x \Sigma_\xi$; $\mu_{\text{X}}^i = E(X_i) = \alpha_x + \Lambda_x \alpha_{\xi_i}$; $\Sigma_{\text{X}} = Var(X_i) = \Lambda_x \Sigma_\xi \Lambda_x^\top + \Sigma_\delta$.

B. Estimating equations for Regression Calibration approach

The estimating equations for IV1 are:

$$S_{\alpha_y} = \sum_{i=1}^{N} (Y_i - \mu_{\text{Y,IV}}^i) \tag{9}$$

$$S_\gamma = \sum_{i=1}^{N} \widehat{\widetilde{X}}_{i1} (Y_i - \mu_{\text{Y,IV}}^i) \tag{10}$$

$$S_{G^*} = \sum_{i=1}^{N} \boldsymbol{V}_i^\top (\widetilde{X}_{i1} - \boldsymbol{V}_i G^*) \tag{11}$$

where $\mu_{\text{Y,IV}}^i = \alpha_y + \gamma^\top \widehat{\widetilde{X}}_{i1}$, $\widehat{\widetilde{X}}_{i1} = (\boldsymbol{V}_i \widehat{G^*})^\top$, G is the matrix of regression coefficients in the first stage regression, G^* is a vector of the stacked columns of G, $\boldsymbol{V}_i = I_\ell \otimes V_i$ where V_i is the ith row of $\boldsymbol{V}$, and I_ℓ is an identity matrix of dimension ℓ. The variance of $\widehat{\theta}_{\text{IV}} = (\widehat{\alpha_y}, \widehat{\gamma}^\top, (\widehat{G^*})^\top)^\top$ can be obtained from the corresponding diagonal elements of $\widehat{var}(\widehat{\theta_{\text{IV}}}) = B^{-1} A B^{-\top}$, where $A = 1/N \sum_{i=1}^N S_i S_i^\top$ and $B = 1/N \sum_{i=1}^N \partial S_i / \partial \theta_{\text{IV}}$, and $S_i = (S_{\alpha_y}, S_\gamma^\top, S_{G^*}^\top)^\top$. In particular, variances for $\widehat{\gamma}$ obtained using the estimating equations approach can also incorporate the variability inherent in the estimation of $\widehat{Z}_i$.

C. Constraints on G

Enforcing constraints on G is straightforward by modifying the design matrix $\boldsymbol{V}_i$ in (11). Specifically, let V_{i1} be the elements of V_i that have constrained coefficients, and V_{i2} be the remaining coefficients. Then, instead of having

$$\boldsymbol{V}_i = \begin{pmatrix} V_i & 0 & 0 \\ 0 & V_i & 0 \\ 0 & 0 & V_i \end{pmatrix}$$

as above, we have

$$\boldsymbol{V}_i^c = \begin{pmatrix} V_{i1} & V_{i2} & 0 & 0 \\ V_{i1} & 0 & V_{i2} & 0 \\ V_{i1} & 0 & 0 & V_{i2} \end{pmatrix}.$$

In this case, $G_c^* = (G_1^*, G_2^*)$ is the vector of coefficients with G_1^* being the coefficients for V_{1i} that are constrained to be equal across scaling items (e.g., block diagonal elements of the matrix G), and G_2^* being the unconstrained coefficients.

Alternative ways to impose restrictions on G is to regress scaling items only on items measured at the same visit (e.g., x_{i11} regressed only on $x_{i12}, ..., x_{i1k}$). This effectively constraints G to be a block diagonal matrix. One can further constrain the blocks of G to be equal, similar to the time invariance assumption for Λ. However although these constraints seem intuitive, they yielded biased estimates when serial independence is present, since constraining G to be block diagonal effectively assumes items measured at other time points are marginally uncorrelated (simulations not shown).

Table S1. Descriptive Statistics for ELEMENT dataset on lead exposure during pregnancy and child mental development

	Time	N	Mean	Std
Child's mental development	24mpp	318	91.4	11.30
Mother's log2 Plasma lead	T1	153	3.8	0.96
	T2	169	3.4	0.86
	T3	157	3.5	0.79
Mother's whole blood lead				
Laboratory 1	T1	155	6.9	4.70
	T2	173	6.2	3.10
	T3	159	6.7	3.40
Laboratory 2	T1	172	7.6	4.60
	T2	198	6.6	3.20
	T3	304	6.9	3.60
Child's log2 Cord blood lead	Birth	238	2.1	0.90
Child's Blood Lead	24mpp	318	4.8	3.40
Covariates				
Mother's IQ	24mpp	341	89.6	18.60
Age	Scr	341	25.9	5.10
Child's gender	Birth	341	0.5	

Abbreviations:Tt=Trimester t; $x$mpp=x months post partum; Scr=Screening

Figure S1. Percent bias in $\widehat{\gamma}_1$ (top), $\widehat{\gamma}_2$ (middle), and bias in $\widehat{\gamma}_3$ (bottom) for instrumental variable and maximum likelihood estimators with various working models(x-axis) when data are generated under a variety of true models (denoted by different symbols in the legend) and true $\gamma = (-2, -2, 0)$.

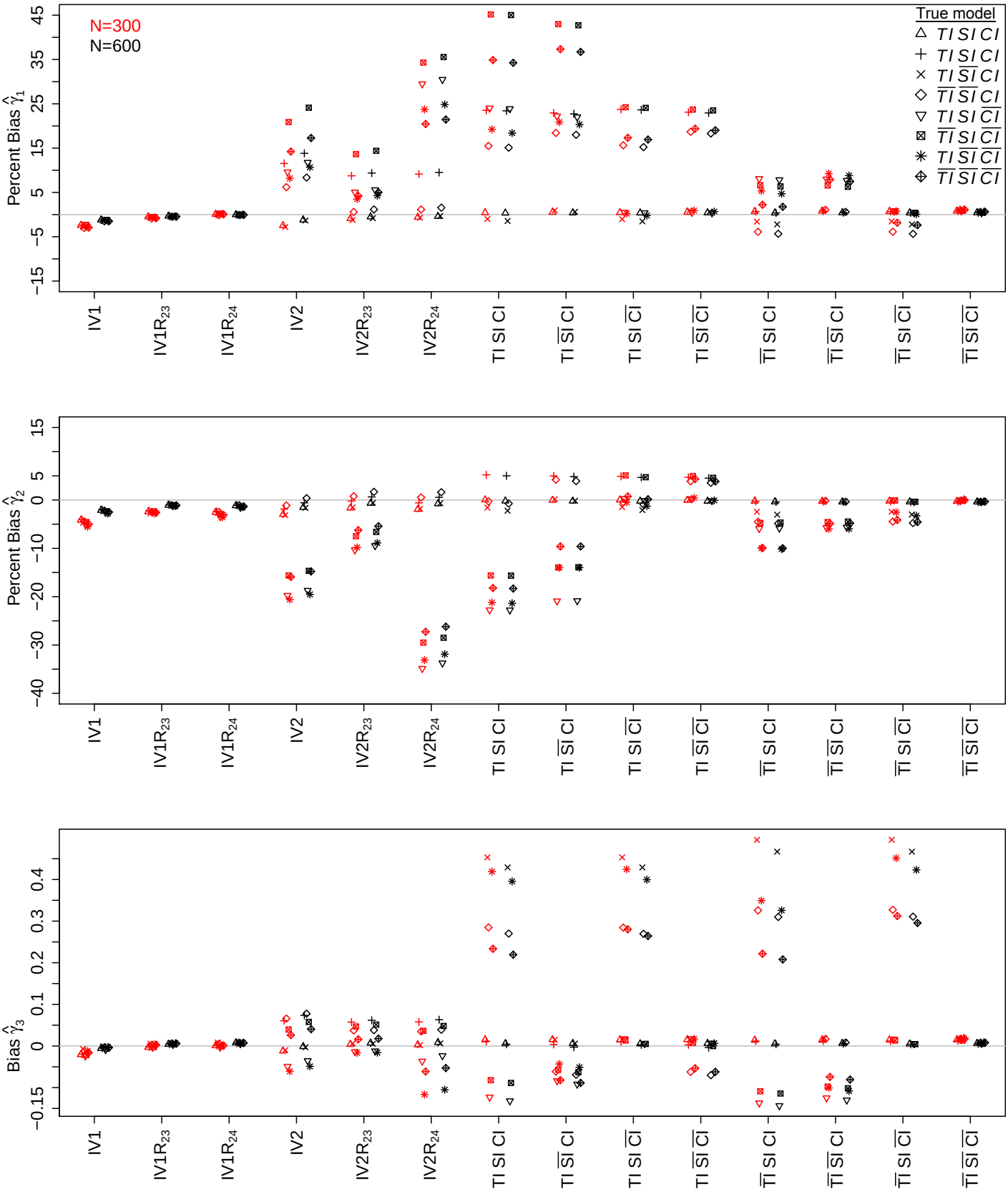

Figure S2. MSE of $\widehat{\gamma}_1$ (top), $\widehat{\gamma}_2$ (middle), $\widehat{\gamma}_3$ (bottom) for instrumental variable and maximum likelihood estimators when data are generated from true $\gamma = (-2, -2, -2)$ and various degrees of lack of time invariance (left); various degrees of serial dependence (middle); various degrees of lack of conditional independence (right).

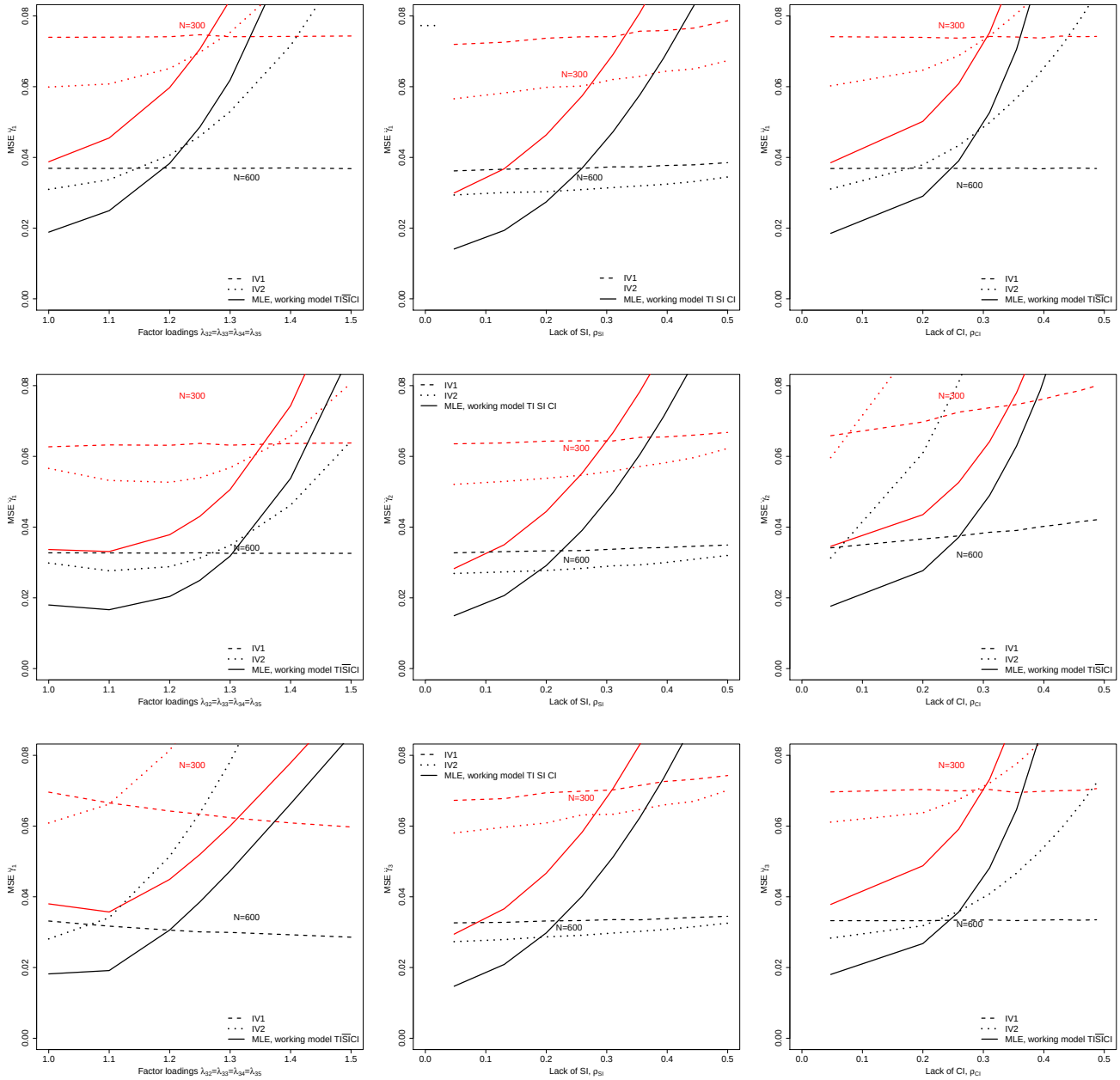

Figure S3. MSE of $\widehat{\gamma}_1$ (top), $\widehat{\gamma}_2$ (middle), $\widehat{\gamma}_3$ (bottom) for instrumental variable and maximum likelihood estimators when data are generated from true $\gamma = (-2, -2, 0)$ and various degrees of lack of time invariance (left); various degrees of serial dependence (middle); various degrees of lack of conditional independence (right).