Online Appendix

To accompany Pek, J. & Chalmers, R. P. (2015). Diagnosing Nonlinearity with Confidence Envelopes for a Semiparametric Approach to Modeling Bivariate Nonlinear Relations among Latent Variables. *Structural Equation Modeling*. doi = 10.1080/10705511.2014.937790

Geometry of Parametric Bootstrap Confidence Envelopes

Consider the simple case of constructing a $(1 - \alpha)100\%$ confidence envelope (CE) for a linear latent regression where $g(\boldsymbol{\theta}_p) = E[\eta_2|\eta_1] = \alpha + \beta\eta_1$, and $\boldsymbol{\theta}_p = (\alpha, \beta)$, such that the subset of ML parameters of interest is $p = 2$. The left panel of Figure A depicts the parameter space of the model where the abscissa represents the estimated intercept $\hat{\alpha}$, the ordinate represents the estimated slope $\hat{\beta}$, and the black dot in the center of the ellipse represents the ML estimate. The 95% Wald-type confidence ellipse is represented by the solid line, and shows that $\hat{\alpha}$ and $\hat{\beta}$ are negatively correlated. The construction of the 95% confidence ellipse begins from defining a unit circle centered at $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta})$, scaling the circle by the critical value $\sqrt{\chi_{2,.95}^2} = 2.45$, and shaping the circle by the estimated variance-covariance matrix of the parameter estimates $\widehat{VAR}(\hat{\boldsymbol{\theta}}_p)$.

The right panel of Figure A depicts the function space, which is a one-to-one mapping from the parameter space via $g(\boldsymbol{\theta}_p)$. The abscissa represents the latent predictor η_1 and the ordinate represents the latent outcome η_2 ; the bolded black line is the predicted linear function of $E[\eta_2|\eta_1]$, and the two solid curves form the 95% Wald-type CE. Observe that all possible values of $\hat{\boldsymbol{\theta}}_p$ lying within the 95% Wald-type confidence ellipse are associated with all possible predicted regression lines lying within the 95% Wald-type CE. As an example, the point represented by the darker colored cross in the parameter space maps onto the dashed darker colored line in the function space, and both estimates are located within their respective confidence sets.

In constructing the 95% parametric bootstrap CE, B bootstrap vectors of parameter estimates $\hat{\boldsymbol{\theta}}_p^b$, where $b = 1, \dots, B$, are sampled from the parameter space. Next, these B bootstrap parameter estimates are passed through the function $g(\boldsymbol{\theta}_p)$ to obtain B regression lines. Finally, the boundary formed by these B bootstrapped lines are taken to form the 95% CE. In the left panel of Figure A, each of the $B = 85$ bootstrap vectors $\hat{\boldsymbol{\theta}}_p^b$ are represented as crosses that map onto the $B = 85$ regression lines in the right panel of Figure A. The 95% bootstrap CE for $E[\eta_2|\eta_1]$ is depicted by the open circles in the right panel of Figure A, forming the boundary of the overlap of the $B = 85$ lines. The 95% Wald-type and bootstrap CEs for this simulated example share much overlap, and the very slight departure between the two CEs is due to the delta method assuming that the sampling distribution of $E[\eta_2|\eta_1]$ is normal and therefore symmetric.

Figure A: Wald-type and parametric bootstrap confidence envelopes for a linear function. The left panel depicts the parameter space, and the right panel depicts the corresponding function space. Every bootstrap estimate in the parameter space maps onto a line in the function space, as depicted by the darker colored cross in the parameter space, and the dashed lined in the function space.

Table 1: Draws for 95% and 90% bootstrap confidence envelopes.

p	$R = 2\sqrt{\chi_{p,.95}^2}$	$B \approx E[R]$	$R = 2\sqrt{\chi_{p,.90}^2}$	$B \approx E[R]$
8	7.88	14,255	7.31	4,586
9	8.23	29,945	7.66	9,232
10	8.56	61,976	8.00	18,352
11	8.87	126,671	8.31	36,095
12	9.17	256,130	8.61	70,351
13	9.46	513,079	8.90	136,042
14	9.73	1,019,382	9.18	261,257
15	10.00	2,010,562	9.45	498,651
16	10.26	3,939,636	9.70	946,541
17	10.50	7,674,148	9.95	1,787,852
18	10.75	14,868,770	10.20	3,361,809
19	10.98	28,667,700	10.43	6,295,567
20	11.21	55,024,760	10.66	11,745,310
21	11.43	105,177,600	10.88	21,836,900
22	11.65	200,272,600	11.10	40,469,420
23	11.86	379,989,400	11.31	74,777,710
24	12.07	718,584,600	11.52	137,789,200
25	12.27	1,354,673,000	11.73	253,241,500
26	12.47	2,546,395,000	11.93	464,303,700
27	12.67	4,773,392,000	12.12	849,338,400
28	12.86	8,924,985,000	12.32	1,550,345,000
29	13.05	16,646,750,000	12.50	2,824,215,000

Note. p = number of parameters, B = number of bootstrap draws, and R = range, where $B \approx E[R] \equiv \int_{-\infty}^{\infty} 1 - \Phi(x)^B - (1 - \Phi(x))^B dx$ and $\Phi(x)$ is the cumulative distribution function for the normal random variable x .

The number of B draws is based on the duality between parameter and function space, where the boundary points making up the 95% confidence set for a $p = 2$ parameter model in both spaces share a range of $R = 2\sqrt{\chi_{2,.95}^2}[\widehat{VAR}(\hat{\theta}_p)]^{1/2}$ (see Figure A). In general, $B \approx E[R] \equiv \int_{-\infty}^{\infty} 1 - \Phi(x)^B - (1 - \Phi(x))^B dx$, where $\Phi(x)$ is the cumulative distribution function for the normal random variable x (Pearson & Hartley, 1956). In this application, $g(\hat{\theta}_p)$ takes the place of x and the standard error of estimate $[\widehat{VAR}(\hat{\theta}_p)]^{1/2}$ is a scaling constant which does not influence size of B .

Table 2: Population generating models for the eight latent regressions.

Function	Latent Regression	$VAR(\zeta_{2i})$	$VAR(\epsilon_{1i})$	$VAR(\epsilon_{2i})$
Linear 1	$\eta_{2i} = 1 + .5\eta_{1i}$	.250	1/3	1/6
Linear 2	$\eta_{2i} = 1 + .5\eta_{1i}$	.750	1/3	1/3
Quadratic 1	$\eta_{2i} = 5 - .25\eta_{1i}^2$	.125	1/3	.083
Quadratic 2	$\eta_{2i} = 2 - \sqrt{.025}\eta_{1i}^2$	.050	1/3	1/30
Quadratic 3	$\eta_{2i} = 2 - \sqrt{.025}\eta_{1i}^2$	.150	1/3	1/15
Exponential 1	$\eta_{2i} = 5 + .04[1 - \exp(-1.5\eta_{1i})]$	.120	1.000	.250
Exponential 2	$\eta_{2i} = 2 + .5[1 - \exp(-.5\eta_{1i})]$	.090	1/3	.060
Exponential 3	$\eta_{2i} = 2 + .5[1 - \exp(-.5\eta_{1i})]$	.270	1/3	.12

Note. The measurement models for each function are defined in Equation 1 with $\boldsymbol{\nu}_1 = \boldsymbol{\nu}_2 = \mathbf{0}_6$, and $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \mathbf{1}_6$. The latent predictor for the nonlinear latent functions follows the distribution $\eta_{1i} \sim N(0, 1)$. All other random variables in the model follow a normal distribution with zero means, and their variances are defined in the table. The latent predictor explains 50% (Linear 1, Quadratic 1 and 2, Exponential 1 and 2) or 25% (Linear 2, Quadratic 3 and Exponential 3) of the variance in the latent outcome, and the latent variables explain 75% of the variances of the measured variables.

All seven population generating models made use of the measurement model below:

$$\begin{aligned} \mathbf{y}_{1i} &= \boldsymbol{\nu}_1 + \boldsymbol{\lambda}_1\eta_{1i} + \boldsymbol{\epsilon}_{1i} \\ \mathbf{y}_{2i} &= \boldsymbol{\nu}_2 + \boldsymbol{\lambda}_2\eta_{2i} + \boldsymbol{\epsilon}_{2i}, \end{aligned} \tag{1}$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are vectors of observed variables measuring the latent predictor η_1 and latent outcome η_2 for individual i . The subscripts 1 and 2 denote the predictor (or exogenous) and outcome (or endogenous) aspects of the model. Intercepts and slopes (or loadings) for the regression of the observed variables on the latent variables are denoted by vectors $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$, respectively. The residuals $\boldsymbol{\epsilon}$ have joint zero mean vector and covariance matrix $\boldsymbol{\Theta}$, and follow a multivariate normal distribution. Typically, observed variables are assumed to be independent, after conditioning on the latent variables such that $\boldsymbol{\Theta}$ is specified as a diagonal matrix. Each latent variable was indicated by six measured variables having means $\boldsymbol{\nu}_1 = \boldsymbol{\nu}_2 = \mathbf{0}_6$ and factor loadings $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \mathbf{1}_6$. The structural model for the latent outcome η_2 varied across the eight functions as summarized above. Note that the linear functions were generated to have skewness of .75 and kurtosis of 1.2 using the generalized lambda distribution system (Ramberg & Schmeiser (1974), cited in Fan, Felsovalyi, Sivo, & Keenan (2002)), so as to avoid the degenerate solution of $K = 1$ mixture components.

Figure B: Absolute bias for for all eight population functions across 1,000 replications.

Figure C: Confidence envelopes (CEs) of a single replication for all eight population functions ($N = 500$).

References

- Fan, X., Felsovalyi, A., Sivo, S. A., & Keenan, S. C. (2002). *SAS[®] for monte carlo studies: A guide for quantitative researchers*. Cary, NC: SAS Institute Inc.
- Pearson, E. S., & Hartley, H. O. (1956). *Tables of the incomplete beta-function*. London, UK: Cambridge University Press.