Abstract Title Page

Title: Reciprocal Effects of Reading and Mathematics? Beyond the Cross-Lagged Panel Model

Authors and Affiliations:
Drew Bailey, University of California, Irvine
Yoonkyung Oh, University of Texas Health Science Center at Houston[2]
George Farkas, University of California, Irvine
Paul Morgan, Pennsylvania State University
Marianne Hillemeier, Pennsylvania State University

Abstract

Background:

The cross-lagged panel model (CLPM) has long been a popular methodological approach for investigators with observational panel data seeking to estimate the mutual effects of two variables on one another. For example, in a highly cited paper, Duncan et al. (2007) regressed later achievement measures on earlier math and reading achievement, as well as controls. They found that math achievement predicted later reading achievement more strongly than early reading achievement predicted later math achievement. If interpreted causally, these findings are counter-intuitive.

This pattern has been interpreted as evidence of likely confounding in cross-lagged estimates from modeling of correlational data (Bailey et al., 2018). Such confounding could occur if children's school-entry math skills are more reflective than school-entry reading skills of confounding factors (including personal and contextual characteristics) that influence children's domain-general learning throughout school.

Purpose:

We re-estimate cross-lagged relations between reading and mathematics achievement using alternative models that better allow for the possibility that the achievement measures are differentially affected by unmeasured confounds throughout the elementary grades. If confounding is a substantial problem in regression models estimating the effects of earlier academic skills on later academic skills, these models should more precisely estimate the strength of the inter-relations between children's reading and mathematics achievement over time. Precisely estimating the strength of these inter-relations has strong theoretical and practical implications, including regarding which types of early interventions can be expected to affect which types of later outcomes.

Data:

We analyzed data from the public version of the Early Childhood Longitudinal Study – Kindergarten Class of 2010-11 (ECLS-K: 2011), which is a large-scale, longitudinal study of a nationally representative sample of U.S. children who were in kindergarten during the 2010-11 academic year and were followed through the fifth grade. (We analyze the K – 3rd grade data that were available at the time of analysis.) The current analyses used the longitudinal panel sample consisting of 9,612 children who participated in the ECLS-K: 2011 in the springs of kindergarten and first, second and third grades data collection with non-missing data on at least one test.

Measures:

We analyzed reading and mathematics test scores measured at each of the four time points (i.e., spring of kindergarten, spring of 1st grade, spring of 2nd grade, and spring of 3rd grade). The reading and mathematics achievement assessments were created through multistage panel review processes, and are based on the 2009 National Assessment of Educational Progress (NAEP) reading and 2005 NAEP mathematics frameworks. Both were individually administered, and scores were scaled using IRT to be comparable across years. Measure reliabilities ranged from .73 to .95 across study waves (Tourangeau et al. 2018).

Analyses:

Several studies have reported on the likelihood of biased estimates when using CLPM to estimate reciprocal effects of variables from longitudinal data (Bailey et al. 2018; Berry & Willoughby 2017; Curran & Bauer 2011; Hamaker, Kuiper, & Grasman 2015; Rogosa, 1980). The problem occurs when the variables have individual and/or contextual influences that are stable over time.

Some of this recent work has adopted approaches within a multilevel, structural equation modeling framework. A key feature of these models is the inclusion of unmeasured stable factors, so that the between-individual and within-individual effects are separated. In one model, this is accomplished by including a random intercept term to account for the individual's over-time mean, in a model known as the random intercept cross-lagged panel model (RI-CLPM). A closely related model is the state-trait cross-lagged panel model (ST-CLPM). Figure 1 displays parameterizations of these models.

We present estimates from a total of 8 models: 1) a CLPM with autoregressive and cross-lagged paths freely estimated, 2) another CLPM with autoregressive and cross-lagged paths held constant across waves, 3-4) a ST-CLPM and a RI-CLPM with autoregressive and cross-lagged paths freely estimated but factor loadings held constant across waves, 5-6) a ST-CLPM and a RI-CLPM with autoregressive and cross-lagged paths constrained across waves but factor loadings freely estimated across waves, and 7-8) a ST-CLPM and a RI-CLPM with autoregressive and cross-lagged paths and factor loadings held constant across waves.

Results:

Table 1 displays descriptive statistics and correlations for all math and reading achievement measures. Table 2 displays model fit statistics. The CLPM models fit the data worst, as indexed by CFI, TLI, and RMSEA (the RMSEA are approximately .08 for models 1A and 1B). The best fitting models on the basis of RMSEA were RI-CLPM models 2A and 2C, and ST-CLPM model 3C. The alternative models fit the data better than the CLPM models.

Table 3 displays parameter estimates for all models. In the CLPMs, we replicate previous findings of asymmetric prediction from Duncan et al. (2007). The CLPM pattern of relatively large cross-effects, with math-reading larger than reading-math, disappears in all RI-CLPM and ST-CLPM models, and even reverses in models 2A and 2B.

In all of the alternative models, loadings on the math and reading factors are high, and the loadings on the math factor are higher than the loadings on the reading factor, consistent with the explanation that these math tests more strongly reflect factors common to academic performance across development than do the reading tests.

Notably, in all ST-CLPM and RI-CLPM models, the correlation between stable math and reading factors is near unity (in all cases, $\boldsymbol{\varphi} > .90$), consistent with the interpretation that the empirically stable individual and environmental factors affecting children's math and reading achievement similarly throughout this developmental period are substantially overlapping.

Conclusions:

In sum, the modifications of the CLPM estimated here fit the longitudinal math and reading test score data better than the original CLPM and provide very different substantive interpretations than were provided by those models. The cross-lagged effects of math at time t to reading at $t+1$, or of reading at time t to math at $t+1$ are generally small but positive; to the extent they are asymmetrical, the effects from earlier reading to later math appear stronger than the effects from earlier math to later reading, a reversal of what is found in the CLPM in the same data.

Appendices

Appendix A. References

Bailey, D. H., Duncan, G. J., Watts, T., Clements, D., & Sarama, J. (2018). Risky business: Correlation and causation in longitudinal studies of skill development. *American Psychologist, 73,* 81-94.

Berry, D. & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development, 88*, 1186-1206. doi: 10.1111/cdev.12660

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583-619.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446. doi:http://dx.doi.org/10.1037/0012-1649.43.6.1428

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological methods*, *20*, 102-116.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245-258.

Tourangeau, K., Nord, C., Le, T., Wallner-Allen, K., Vaden-Kiernan, N., Blaker, L., & Najarian, M. (2018). Early Childhood Longitudinal Study, Kindergarten class of 2010–11 (ECLS-K: 2011). User's manual for the ECLS-K: 2011 kindergarten-third grade data file and electronic codebook, public version. NCES 2018-034. Washington, DC: National Center for Education Statistics.

Appendix B. Tables and Figures

Table 1. Descriptive Statistics and Correlations

	Mean	SD	Reading at S-K	Reading at S-1	Reading at S-2	Reading at S-3	Math at S-K	Math at S-1	Math at S-2	Math at S-3
Reading at S-K	68.08	14.17	1.00							
Reading at S-1	93.09	18.02	0.79	1.00						
Reading at S-2	107.40	15.51	0.71	0.86	1.00					
Reading at S-3	116.80	14.82	0.64	0.78	0.85	1.00				
Math at S-K	49.91	12.84	0.73	0.71	0.69	0.67	1.00			
Math at S-1	73.87	17.64	0.66	0.73	0.71	0.70	0.83	1.00		
Math at S-2	90.65	16.66	0.62	0.70	0.74	0.73	0.78	0.86	1.00	
Math at S-3	102.87	15.63	0.59	0.67	0.69	0.72	0.75	0.82	0.88	1.00

Note: *N*=9612

Table 2. Model Fit Statistics (Full panel sample: *N*= 9612)

Model	Structure	Factor Loadings	AR & CL paths	χ^2	*p*-value	*df*	CFI	TLI	RMSEA	90% CI	SRMR
1A	CLPM	N/A	Free	1250.03	<.000	18	0.986	0.978	0.084	[0.080, 0.088]	0.023
1B	CLPM	N/A	Constrained	1504.36	<.000	26	0.983	0.982	0.077	[0.074, 0.080]	0.026
2A	RI-CLPM	Constrained	Free	269.78	<.000	15	0.997	0.995	0.042	[0.038, 0.047]	0.016
2B	RI-CLPM	Constrained	Constrained	651.99	<.000	23	0.993	0.991	0.053	[0.050, 0.057]	0.029
2C	RI-CLPM	Free	Constrained	337.00	<.000	17	0.996	0.994	0.044	[0.040, 0.048]	0.010
3A	ST-CLPM	Constrained	Free	537.60	<.000	15	0.994	0.989	0.060	[0.056, 0.065]	0.025
3B	ST-CLPM	Constrained	Constrained	812.11	<.000	23	0.991	0.989	0.060	[0.056, 0.063]	0.032
3C	ST-CLPM	Free	Constrained	346.49	<.000	17	0.996	0.994	0.045	[0.041, 0.049]	0.009

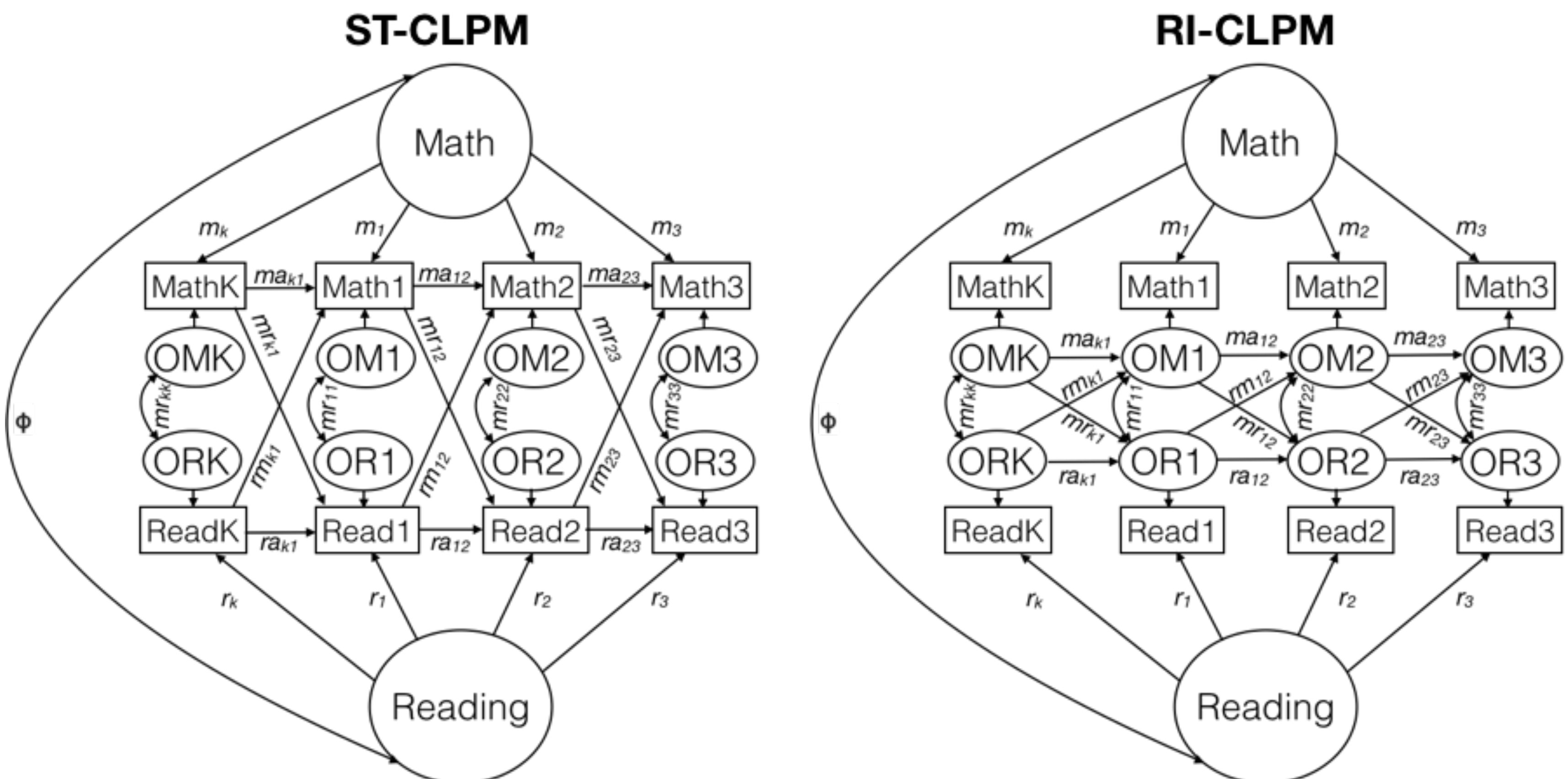
Note: AR = autoregressive; CL = cross-lagged. The RI-CLPM and ST-CLPM structures appear in Figure 1.

Table 3. Model Estimates (Full panel sample: *N*= 9612)

Model:	1A	1B	2A	2B	2C	3A	3B	3C
Loadings								
$m_{k,}$ $m_{1,}$ $m_{2,}$ m_3	N/A	N/A	.86, .89, .88, .87	.84, .86, .88, .89	.82, .88, .90, .91	.82, .77, .78, .77	.82, .77, .77, .77	.83, .61, .62, .62
r_k, $r_{1,}$ $r_{2,}$ r_3	N/A	N/A	.73, .76, .76, .77	.69, .70, .73, .74	.70, .80, .85, .86	.70, .67, .67, .66	.69, .67, .66, .64	.71, .49, .48, .48
Autoregressive Paths								
ma_{k1}, $ma_{12,}$ ma_{23}	.74, .74, .81	.75, .77, .79	.23, .29, .50	.38, .37, 36	.38, .35, .34	.16, .19, .26	.19, .20, .20	.32, .32, .32
ra_{k1}, $ra_{12,}$ ra_{23}	.59, .73, .69	.66, .69, .68	.53, .65, .61	.63, .68, .66	.55, .53, .48	.35, .37, .29	.34, .34, .34	.46, .47, .46
Cross-Lagged Paths								
mr_{k1}, $mr_{12,}$ mr_{23}	.28, .18, .22	.22, .23, .23	.05, .03, .09	.06, .06, .06	.04, .03, .02	-.07, -.11, -.04	-.07, -.07, -.07	$-.01^{ns}$, $-.01^{ns}$, $-.01^{ns}$
rm_{k1}, $rm_{12,}$ rm_{23}	.12, .16, .10	.12, .13, .13	.11, .14, .06	.12, .13, .13	.04, .04, .04	$.01^{ns}$, -.05, -.12	-.04, -.04, -.04	$.00^{ns}$, $.00^{ns}$, $.00^{ns}$
Occasion Covariances								
mr_{kk}, $mr_{11,}$ mr_{22}, mr_{33}	.73, .34, .30, .22	.73, .34, .30, .22	.44, .23, .26, .18	.50, .26, .26, .16	.48, .17, .19, .10	.43, .08, .06, .12	.43, .11, .09, .13	.43, .11, .09, .13
φ	N/A	N/A	0.94	0.96	0.91	0.92	0.92	0.93

Note: All paths are standardized. All are statistically significant (p < .05), except for those denoted nonsignificant (ns).

Figure 1. Parameterization of ST-CLPM and RI-CLPM Structural Models



Note: Subscripts indicate grades (K = kindergarten, 1 = 1st Grade, 2 = 2nd Grade, 3 = 3rd Grade). The RI-CLPM model in the right panel is equivalent to a cross-lagged panel model when the variances of the Math and Reading intercepts are constrained to 0.