advances.in/psychology

June 18, 2025
Doi: 10.56296/aip00037
ISSN: 2976-937X

METHODS | SPECIAL ISSUE: NETWORK PSYCHOMETRICS

Cross-lagged panel networks

Anna Wysocki [1], Ian McCarthy [1], Riet van Bork [2], Angélique O. J. Cramer [3,4], & Mijke Rhemtulla [1]

Received: December 17, 2023| **Accepted:** June 3, 2025| **Published:** June 18, 2025| **Edited by:** Hudson Golino

[1]Department of Psychology, University of California, Davis, [2]Department of Psychology, University of Amsterdam, the Netherlands, [3]Department of Psychiatry, Faculty of Medicine, Amsterdam UMC (AMC), Amsterdam, the Netherlands, [4]Centre for Urban Mental Health, University of Amsterdam, Amsterdam, the Netherlands
* Mijke Rhemtulla, Department of Psychology, University of California, Davis, One Shields Avenue, Davis CA 95616. E-mail: mrhemtulla@ucdavis.edu. This article is published under the Creative Commons BY 4.0 license. Users are allowed to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

Network theory and accompanying methodology are becoming increasingly popular as an alternative to latent variable models for representing and, ultimately, understanding psychological constructs. The core feature of network models is that observed variables (e.g., symptoms of depression) directly influence one another over time (e.g., low mood → concentration problems), resulting in an interconnected dynamical system. The dynamics of such a system might result in certain states (e.g., a depressive episode). Network modeling has been applied to cross-sectional data and intensive longitudinal designs (e.g., data collected using an Experience Sampling Method). In this paper, we present a cross-lagged panel network model to reveal item-level longitudinal effects that occur within and across constructs that are measured at a small set of measurement occasions. The proposed model uses a combination of regularized regression estimation and structural equation modeling to estimate auto-regressive and cross-lagged pathways that characterize the effects of observed components of psychological constructs on each other over time. We demonstrate the application of this model to longitudinal data on students' commitment to school and self-esteem.

Keywords: network models; cross-lagged panel design; cross-lagged networks

1. INTRODUCTION

Panel designs involve collecting data from a large sample of participants over multiple discrete measurement occasions that are often separated by months or years. Panel designs are common in developmental research because they allow researchers to investigate how variables are related to each other over a developmental period (e.g., early childhood or adolescence). Researchers increasingly analyze panel data using cross-lagged panel models (CLPMs) in a structural equation modeling (SEM) framework, with constructs of interest modeled as latent variables underlying a set of observed measures (Finkel, 2008; Little et al., 2007; Mackinnon et al., 2022; Sher et al., 1996). In a latent variable CLPM, the associations of interest are between variables represented as latent constructs, which are presumed to influence each other over time (e.g., the latent construct 'neuroticism' affects the latent construct 'depression')[1]. Network models have been presented as an alternative to latent variable models for many psychological constructs. The key feature of network models is that relations are modeled at the level of the behaviors, thoughts, and feelings, that are queried in individual items rather than at the level of latent variables (Borsboom, 2008; Borsboom & Cramer, 2013; Cramer et al., 2010, 2012). For example, instead of a person's level on a latent "extraversion" construct causing that person both to dislike large groups of people and also to dislike parties, the network modeling approach hypothesizes liking groups of people and liking parties to be directly related (e.g., disliking groups of people makes a person dislike parties).

Thinking about constructs in a network manner makes the most sense—from a conceptual standpoint—when logic, theory, and/or empirical evidence indicates that the covariance between items is better explained by a direct relation (e.g., two symptoms of major depression: insomnia → fatigue) than by a latent variable (insomnia ← LV → fatigue). Investigating constructs from a network perspective is aligned with a growing body of empirical research suggesting that predictive power can be greater when items are considered as individual predictors rather than in aggregate (Revelle, 2024; Seeboth & Mõttus, 2018; Stewart et al., 2022).

In the present paper, we introduce the cross-lagged network model, which extends the network modeling approach to a model for longitudinal panel data. This model is essentially an exploratory cross-lagged panel model that examines longitudinal associations among individual elements of a construct rather than among the constructs. As the network modeling approach has been adopted in many applied fields (e.g., clinical and personality psychology; Borsboom et al., 2021; Costantini et al., 2020; McNally et al., 2015) where panel data are widely available, the cross-lagged network model presents a necessary addition to the network modeling toolbox. In the remainder of this paper, we will first introduce cross-lagged panel models and network models. We then present the crossed-lagged panel network as a combination of these two methods and distinguish it from other network

[1] It is also common to see cross-lagged panel models fit to composite scores, such as sum scores or scale averages. While these are not the same as latent variables, they also model the variables of interest as high-level aggregate concepts.

approaches to panel data. Throughout this article, we use a dataset of longitudinal panel data on two constructs, commitment to school and self-esteem, as an empirical example and will apply the proposed crossed-lagged panel network to these data. Finally, we discuss the strengths and limitations of the model more generally and offer suggestions for its use.

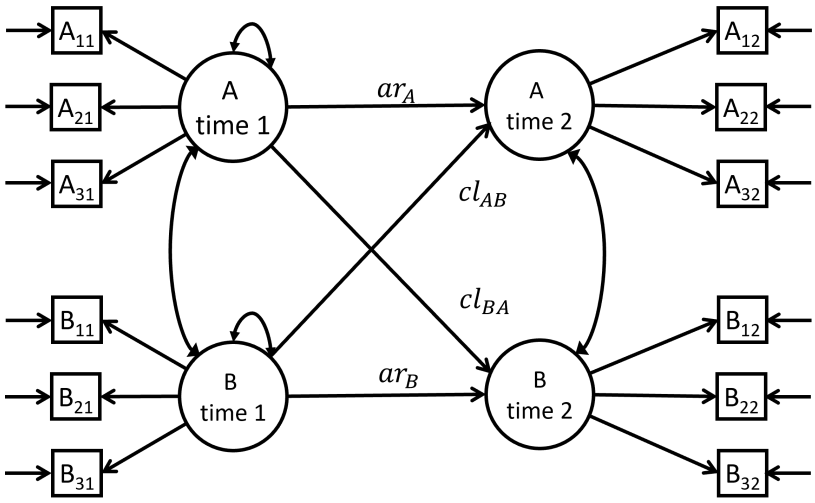
1.1 Cross-lagged Panel Models

In a cross-lagged panel design, two or more constructs are measured at two or more discrete measurement occasions that are typically widely spaced in time (e.g., months or years apart). By regressing the set of variables at each occasion on the set of variables at the previous occasion, one can estimate the "cross-lagged" effect of each variable on the other over a particular time lag (i.e., whatever time lag separates the two measurement occasions), while controlling for the auto-regressive effect of each variable on itself. Because cross-lagged model parameters do not support causal inference (Hamaker et al., 2015), we interpret cross-lagged paths in terms of prediction and not in causal terms (a subject we will return to in the Discussion). That is, a regression path from variable X at occasion t to a variable Y at future occasion $t+1$ suggests that a change in X predicts a change in Y over the course of that specific time lag. Such predictive relations over time may generate causal hypotheses (Epskamp & Fried, 2018; Godfrey-Smith, 2010; Williamson, 2006): If X really causes Y, then regressing Y at occasion $t+1$ on X at occasion t, controlling for Y at occasion t, should produce a non-zero regression coefficient. Finding a non-zero regression coefficient is thus an important but not sufficient condition for determining causality (Barnett et al., 2009).

In a traditional CLPM, the variables at each measurement occasion are scale scores (i.e., sum scores of a set of items on some psychometric scale). In recent years, it is increasingly common for developmental researchers to use latent variables to represent the constructs of interest in a panel model (e.g., Johnson et al., 2017). Figure 1 shows a CLPM with two latent variables and two measurement occasions. The cross-lagged paths (CL_{BA} and CL_{AB}) represent the predictive strength of each construct on the other at a later measurement occasion, controlling for the auto-regressive effects of each variable on itself (AR_A and AR_B).

CLPMs are most appropriate when one has data on several constructs at a few discrete measurement occasions from a large group of individuals, and when the research questions center on predictive or potentially causal effects of these constructs over time. The CLPM has known limitations that we return to later, but despite these, it continues to be a popular and useful model to investigate relations among constructs with data gathered at a limited number of measurement occasions.

Figure 1
Cross-lagged panel model based on latent variables, with two constructs and two measurement occasions.



1.2 Network Models

The network approach to multivariate psychological data is based on the idea that high-level attributes (disorders, traits, abilities) may arise via lower-level processes by which individual attitudes, behaviors, symptoms, beliefs, and abilities interact with each other and, as such, form dynamic systems that, over time, settle into stable states (e.g., an episode of major depression, a high state of neuroticism, or a particular level of intelligence; Borsboom, 2008; Cramer et al., 2010; van der Maas et al., 2019; van der Maas et al., 2006). Network models target these processes by modeling direct relations between lower-level variables, rather than representing them as reflections of a single underlying high-level attribute as in a latent variable model. Applications of network models to psychological data include personality (Beck & Jackson, 2020; Cramer et al., 2012; Taylor et al., 2021), depression (Aalbers et al., 2019; Cramer et al., 2016; Zavlis et al., 2022), schizophrenia (Abplanalp et al., 2022; Isvoranu et al., 2017), quality of life (Kossakowski et al., 2016), subjective well-being (Deserno et al., 2017), obsessive-compulsive disorder (Cervin et al., 2020), post-traumatic stress disorder (Armour et al., 2020; McNally et al., 2015), complicated grief (Robinaugh et al., 2016), and substance abuse (Rutten et al., 2021).

In a typical network model, a set of items measuring a construct (e.g., the items on an extraversion scale) are represented by a set of network nodes. The relations between these nodes are represented by edges. Networks estimated from cross-sectional data (i.e., data gathered at a single measurement occasion from a large group of individuals) are typically undirected conditional association networks; that is, each edge represents a partial correlation between the nodes (variables) that it connects, controlling for all other nodes in the network. Conditional association networks can be estimated using l1-regularized regression techniques, which push many edges to exactly zero, resulting in a sparse network that can still account for associations in the data. Alternatively, conditional associa-

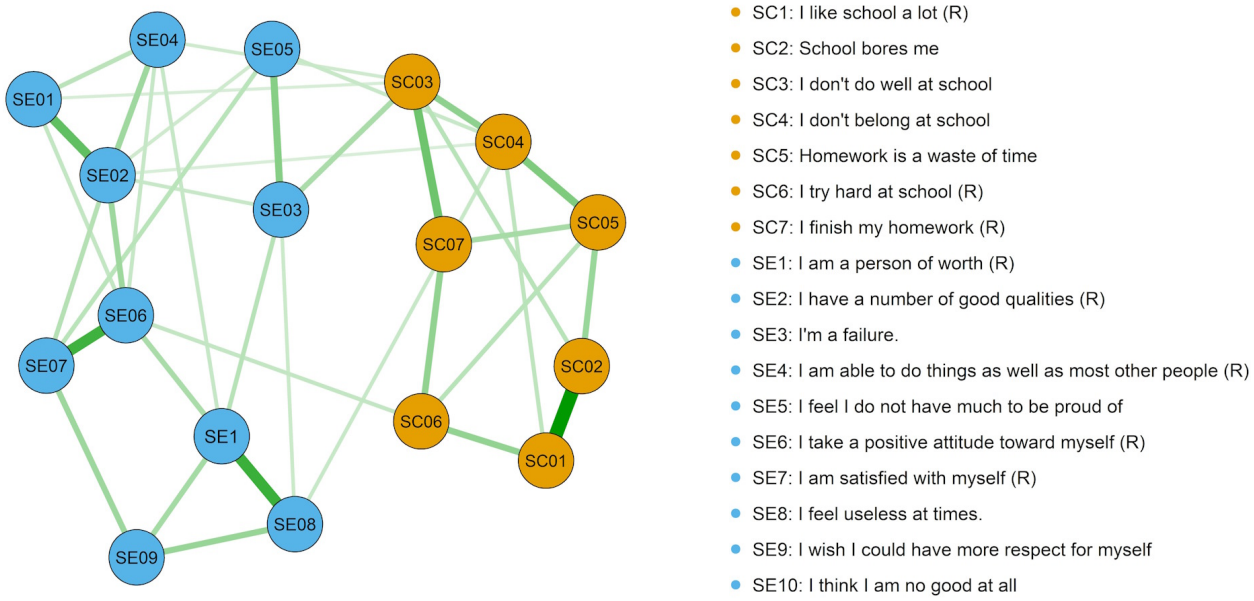
tion networks can be estimated using non-regularized techniques where sparsity is induced using edge significance levels or BIC model selection [2].

Figure 2 shows a partial correlation network for the constructs commitment to school (7 items) and self-esteem (10 items), based on data from the Iowa Youth and Families Project (Conger et al., 2011) at the first wave of data collection (i.e., students in 7th grade; data details are presented in the "Data" section below). Green lines represent conditional positive relations among variables; thicker and darker lines represent stronger relations. The layout of nodes is chosen by an algorithm that places more strongly connected nodes closer together (Fruchterman & Reingold, 1991). Thus, one can see from the thick green lines and the placement of the variables that the items within each construct cluster together, suggesting that these constructs are conceptually coherent. Within each construct, certain pairs of items are much more strongly connected than others; for example, the items "I like school a lot" (reverse coded) and "School bores me" are more strongly related than either of those items are to the item "I try hard at school" (reverse coded). This network also reveals some connections across the two constructs, including a link between the commitment to school item, "I don't do well at school" and the self-esteem item, "I'm a failure". These cross-network connections generate hypotheses about the pathways by which the two constructs could influence each other over time (Epskamp et al., 2018, 2022).

Network models have also been developed for intensive longitudinal time-series data, that is, data in which a single individual or a group of individuals provide many (e.g., 100 or more) data points that are typically closely spaced in time (e.g., seconds, minutes, or hours apart). Time-series network models include vector auto-regressive (VAR) models that model variables at occasion $t + 1$ as a function of those at occasion t within a single individual (e.g., Epskamp et al., 2018; Gates & Molenaar, 2012). In particular, Epskamp (2020a) and Epskamp et al. (2018) proposed a "graphical VAR" approach that estimates a temporal network containing auto-regressive and cross-lagged paths, as well as a contemporaneous network that describes associations among variables within a measurement occasion. This model can be fit to intensive time-series data from a single person, or to time-series data of multiple people. When multiple people are modeled, the graphical VAR method also produces a between-subjects network of random intercepts. Bringmann et al. (2013) also developed a multilevel modeling VAR approach for data with many time points and many participants. The VAR approach is a powerful model for discovering predictive relations—and hence, potential causal processes—within individuals (Epskamp et al., 2018, 2022).

[2] R packages for estimating partial correlation networks include qgraph (i.e., the EBICglasso or ggmModSelect functions for Gaussian data; Epskamp et al., 2012), IsingFit for binary data (Van Borkulo et al., 2014), mgm for mixed binary, continuous, and categorical data (Haslbeck & Waldorp, 2020), GGMnonreg for non-regularized estimation of network models with continuous, binary, and categorical data (Williams, 2021), and BGGM for Bayesian estimation of network models (Williams & Mulder, 2020).

Figure 2
Cross-Sectional Partial Correlation Network of Commitment to School and Self-Esteem



Note. Orange nodes are variables related to Commitment to School; blue nodes are variables related to Self-Esteem. The network was estimated using the ggmModSelect function in *qgraph*, which performs a search over unregularized models to find one with minimum EBIC.

1.3 Network Models for Panel Data

While the graphical VAR approach was designed for individual time-series data, it can be applied to panel data (where it is called “panel VAR”) with three or more measurement occasions. The temporal network then contains estimates of autoregressive and cross-lagged paths across measurement occasions after controlling for person-level stable means in each variable. One notable feature of the panel VAR model is that it assumes stationarity across measurement occasions (i.e., equal means and covariances at each measurement occasion, and equal autoregressive and cross-lagged path coefficients across each pair of subsequent occasions). No matter how many measurement occasions are present in the data, only a single temporal network is estimated. The assumption of stationarity may often be appropriate for intensive repeated measurement designs in which measurements are tightly spaced, but it might be more likely to be violated when measurement occasions are further apart. Stationarity may be especially likely to be violated in, for example, developmental research where measurement occasions span multiple developmental periods (e.g., from early childhood until adolescence) or research in which the time span between the different measurement moments is very different (e.g., two measurement points with several months between them and a third measurement point a year later).

In the following section, we present a model that can be used to estimate cross-lagged predictive relations among the individual components of multiple constructs over time for non-intensive panel data. As an empirical example, we use panel data on two constructs, commitment to school and self-esteem. In addition to highlighting predictive relations, the proposed model will yield information about which variables in the network are most central in terms of predictive ability.

1.4 Cross-Lagged Panel Networks

In the previous sections, we described CLPMs, which highlight pathways among constructs over a few fixed measurement occasions, and network models, which reveal unique pathways among the constituent variables of a construct. We combine these two approaches to form cross-lagged panel networks (CLPNs). The key feature of a CLPN is that the relations among individual items are modeled as directed paths across time, reflecting the variance shared between a variable at occasion t and another (or the same) variable at occasion $t + 1$, controlling for all other variables at occasion t (see also Conlin et al., 2022). The interpretation of these directed paths is subject to the same constraints as those of the traditional CLPM described above; that is, paths may be interpreted as predictive relations and/or as causal-hypothesis-generating. Similarly, the CLPN cannot estimate differences in predictive patterns across individuals (i.e., every person is assumed to have the same pattern of predictive relations over time) but it can allow different predictive patterns over time (e.g., predictive patterns from 8th grade to 9th grade may differ from those from 9th to 10th grade).

The proposed CLPN model has theoretical and pragmatic benefits. Results of the model can suggest mechanisms by which constructs sustain themselves (e.g., beliefs and behaviors that make up self-esteem may influence each other over time, creating positive feedback loops that allow self-esteem to persist as a stable trait) and may lead to changes in other constructs over time (e.g., beliefs and behaviors related to self-esteem may influence a person's ability to make and maintain friendships over time, resulting in a positive association between the broad constructs of self-esteem and friendship quality). Because a CLPM with latent variables makes the strong assumption that all predictive effects in the observed data are explained by relations among the latent constructs, the latent variable CLPM has no potential to discover which individual components of such constructs have the strongest influence on longitudinal change and stability. In contrast, the CLPN can uncover the individual behaviors, beliefs, attitudes, traits, or symptoms that are responsible for these auto-regressive and cross-lagged longitudinal construct associations. For example, consider two items of the commitment to school scale: success in school and enjoyment of school. The CLPN allows that not doing well in school may affect a student's later enjoyment of school over and above that student's mean-level change in school commitment. Moreover, the CLPN allows individual nodes belonging to one construct to affect nodes belonging to a different construct at a later time. For example, not doing well in school may affect a student's later feelings of being a failure over and above the broader effect

of school commitment on self-esteem. If individual components of one construct have unique effects on other constructs, the CLPN will highlight these effects. Learning which components are primarily responsible for the cross-time associations among constructs may generate more fine-grained descriptions and theories of how constructs relate to each other. A key benefit of this approach is that we do not need the assumption of stationarity to estimate the model, and, as such, can identify when the patterns of effects vary across occasions. The cost of this flexibility is that the CLPN has many more potential paths to estimate and interpret, both compared with a latent-variable CLPM and compared with a panel VAR approach. For example, whereas a CLPM with two latent constructs, each measured by 10 items at 3 measurement occasions, would produce 4 auto-regressive coefficients and 4 cross-lagged paths; a CLPN based on the same data would produce 40 auto-regressive paths (i.e., those connecting each variable at each measurement occasion to the same variable at the next occasion) and 740 cross-lagged paths (connecting each variable at each occasion to each other variable at the subsequent occasion). While it may be possible to estimate 800 regression coefficients, it is not feasible to interpret them individually.

The CLPN deals with the problem of model complexity in two ways. The first is by using an exploratory estimation procedure that sets many of the paths to zero, resulting in fewer paths to visualize and interpret. The second is by offering outcome statistics that summarize information in the estimated model, such as the degree to which each variable predicts and is predicted by variables within and outside of its construct.

In the next section, we briefly describe the technical details of our proposed model, and then present a simulation study to evaluate its performance relative to the panel VAR model and relative to two alternative estimation methods. We then apply the model to the constructs commitment to school and self-esteem and describe each step of the analysis and offer suggestions for how the results may be interpreted.

2. METHODOLOGICAL DETAILS

The process of producing a CLPN model can be broken down into four steps. The first step involves collapsing overlapping items to reduce the set of variables submitted to the CLPN analysis. The second step involves fitting a series of regularized regression models to select the initial CLPN model. The third step is to re-estimate the model selected in step 2 as a structural equation model (SEM) obtaining non-regularized estimates for the cross-lagged and auto-regressive coefficients across time. Finally, the fourth step is to summarize the results by producing plots and computing summary statistics such as nodewise in-prediction and nodewise out-prediction. We describe these steps in the following sections.

2.1 Collapsing Nodes

The guiding intuition behind network modeling of psychological constructs is that interactions among variables may occur at a much lower level of specificity than the broad constructs that are typically considered meaningful (Cramer et al., 2012). But there is no guarantee that the items used to assess a construct are at the appropriate level of generality to capture these causal effects. In particular, for many cognitive and affective constructs, it is not unusual to have sets of items that are highly overlapping. While a typical assessment of psychopathology may contain a single item to represent each symptom, a self-esteem questionnaire might include all the items, "I feel good about myself", "I am satisfied with myself", and "I like myself the way I am". Given a high degree of conceptual overlap, a reasonable expectation is that these three items behave in the same way (i.e., have the same effects and predictors) in the overall network architecture. Thus, we recommend collapsing semantically similar items (by summing or averaging scores on these items) to achieve a set of variables that capture meaningful units of interaction. Doing so will simplify the model, reduce the possibility of spurious relations, and render the results more interpretable.

We stress that the goal of this analysis step is not to reduce a large set of items to a small set of underlying dimensions, but simply to remove the redundancy that arises from nearly-identical items that query the same manifest behavior. As such, we recommend that researchers examine item wording and use their own expertise to identify redundant items, that they do this prior to data collection or analysis, and that they transparently report on their rationale for collapsing items. In the absence of content expertise or for researchers who prefer an automated approach, unique variable analysis (implemented in the EGAnet R package; Christensen et al., 2023) identifies pairs of nodes that are highly empirically overlapping, but it is not guaranteed to detect items that overlap in their meaning. Another promising approach leverages large language models to identify pairs of items with high overlap (Wulff & Mata, 2025).

Dimension-reduction techniques such as principal components analysis, singular value decomposition, and factor analysis aim to reduce a large set of items to just a few broad dimensions. This aim is antithetical to the goals of the network approach, which are to understand the connections among low-level behaviors, beliefs, attitudes, abilities, symptoms, or traits. Thus, while it would be possible to use statistical dimension-reduction techniques to decide which items to collapse, we recommend against it. As a central goal of network analysis is to move away from abstractions and toward explaining constructs in terms of concrete observable variables, applying statistical grouping methods is likely to be counterproductive.

We further recommend that researchers err on the side of not collapsing too much, for the same reason: the more variables are collapsed, the more removed the composite gets from a concrete observable behavior, belief, or attitude. As the goal is to foster an understanding of the relations

among low-level variables, it is important that only variables that are highly semantically similar be collapsed.

2.2 Regularized Regression

Once the data have been condensed into a set of appropriate variables, the next step is to estimate auto-regressive and cross-lagged paths; that is, linear regression coefficients of each variable on itself and each other variable at the previous occasion. We present equations using just the first two measurement occasions ($T1$ and $T2$); with more occasions these equations will repeat for each subsequent pair of adjacent occasions (e.g., $T2$ to $T3$, $T3$ to $T4$, etc.). For each variable and each measurement occasion after the initial one, we fit the model

$$\mathbf{x}_{T2} = \beta_0 + \mathbf{B}\mathbf{x}_{T1} + \epsilon, \tag{1}$$

where $\mathbf{x}_{T1}$ and $\mathbf{x}_{T2}$ are vectors containing all p variables at $T1$ and $T2$, respectively, $\beta_\mathbf{0}$ is a vector of intercepts for the p variables at $T2$, **B** is a $p \times p$ matrix of linear regression coefficients[3] of all p variables at $T2$ on those at $T1$, and ϵ is a vector of residuals for each variable. The model can be extended to more measurement occasions by regressing the variables at each subsequent measurement occasion on the previous occasion.

If the sample size is large enough—at a minimum, sample size must be greater than the total number of variables at all measurement occasions for the covariance matrix to be positive definite—the model can be estimated in a single step in an SEM framework as a traditional cross-lagged panel model with p unique variables measured at t occasions. But the set of parameters to estimate in such a model includes variances and covariances at each of t occasions in addition to all the auto-regressive and cross-lagged paths, giving $q = (tp)(p+1)/2 + p^2(t-1)$ parameters to estimate. This number quickly inflates with many variables and/or occasions, for example, with 20 variables and 4 occasions there are 2040 parameters to estimate. Given that conventional SEM wisdom entails having many more participants than parameters (i.e., the $N:q$ ratio should be high; Bentler & Chou, 1987; Jackson, 2003; Tanaka, 1987), and that estimation of cross-sectional networks commonly uses lasso regularization to shrink the parameter space of large networks (Epskamp & Fried, 2018), we opted to estimate the model first using lasso-regularized regression models. In the Simulation that follows we will compare the SEM-only approach to the regularized regression approach described here. To proceed, we fit a series of univariate linear regression models of the form $x_{(j,T2)} = \beta_{(0,j)} + \beta'_j x_{T1} + \epsilon_j$, where $x_{(j,T2)}$ is the jth variable at $T2$, β_0 is its intercept, $\beta_\mathbf{j}$ is the corresponding vector of regression coefficients that predict it from the vector $\mathbf{x}_{\mathbf{T1}}$ containing all p variables at $T1$, and ϵ_j is its residual.

[3] We assume that the data are continuous and normally distributed. However, the model can readily be extended to allow for alternative distributions by using generalized linear model with a link function, e.g., a logit link for binary variables.

At this step, regression coefficients are estimated using lasso regression, a penalized regression approach that applies an l_1 penalty to the estimated regression coefficients (Friedman et al., 2010). This estimation technique has the effect of shrinking small regression paths to exactly zero, thus achieving a sparse solution. The l_1 penalty is just one possible penalty that can be used to address the problem of overfitting, which is important when estimating a regression with a large number of predictors (the l_2 penalty, i.e., ridge regression, is another possibility). The rationale for using lasso regression is that, on the assumption that some paths are truly zero, it will set to zero those paths that are most likely not to exist, in addition to estimating coefficients for non-zero paths. The result is a sparse network, in which many of the paths from variables at $T1$ to variables at $T2$ (and from $T2$ to $T3$, etc.) will be estimated as exactly zero.

The regularized regression estimates minimize $\frac{1}{N}\sum_{i=1}^{N}\left[l(x_{(i,j,T2)}, \hat{\beta}_{(0,j)} + \hat{\beta}_j' x_{(i,T1)}) + \lambda_j \|\hat{\beta}_j\|_1\right]$, where $i = 1, \ldots, N$ denotes individuals and $j = 1, \ldots, p$ denotes variables. Thus, what is minimized is the sum of individual likelihoods plus a penalty, $\lambda_j \|\hat{\beta}_j\|_1$ in which $\|\hat{\beta}_j\|_1$ denotes the sum of absolute values of the coefficients in $\hat{\beta}_j$, and λ_j is a tuning parameter that determines the strength of the penalty. When λ_j is zero, the penalty drops away and we are left with ordinary least squares (OLS) regression. As λ_j increases, all the coefficients in $\hat{\beta}_j$ will eventually be shrunk to zero. Though there are many criteria that can be applied to choose λ_j, here we use 10-fold cross-validation (CV): the data are partitioned into 10 non-overlapping subgroups. Within each set of 9 of these subgroups, the regression model is estimated 100 times, using a series of 100 values of λ_j, and for each value the prediction error is computed on the left-out subgroup. Prediction error is averaged across all 10 subgroups for each λ_j, and the value with the lowest average prediction error is chosen. The CV criterion has been shown to be the most conservative criterion in that it typically shrinks the fewest coefficients to zero, allowing more parameters to be estimated compared to other criteria (e.g., EBIC; Bien & Tibshirani, 2011; Wysocki & Rhemtulla, 2021).

Lasso regression can help overcome estimation challenges in high-dimensional or near high-dimensional settings by shrinking the parameter space (Fan et al., 2016), and it minimizes overfitting leading to more generalizable results (i.e., greater predictive accuracy) than OLS estimates (McNeish, 2015). However, lasso regression also biases the non-zero edges towards zero. Additionally, recent research has shown that the CV approach to selecting λ_j has extremely high sensitivity (i.e., of the true existing edges, a very high proportion is correctly included in the model). The trade-off is that the CV approach often has a high false positive rate and this false positive rate does not diminish with increased sample size (Kuismin & Sillanpää, 2016). In other words, paths that are estimated to be zero in CV are very likely to be true zeroes, but many paths that are not set to zero may be false positives. Using regularized regression can help initially estimate the model. But, given the high false positive rate and biased estimates, we want a further step to prune the model and get non-regularized estimates for interpretation.

2.3 Model Pruning and Estimating Edge Weights in SEM

The regularized regressions from the previous step provide us with a pruned starting model in which a subset of variables at each occasion predict a subset of variables at the subsequent occasion. We then fit this starting model as a path model in the SEM framework, fixing to zero all paths that were estimated to be zero in the regularized regression step. In the full SEM model, all variables' residuals at each measurement occasion are allowed to covary with each other. The SEM approach also makes it easy to deal with missing values by estimating the model using full-information maximum likelihood, and it will enable model comparisons using likelihood ratio tests to test whether cross-time constraints hold.

2.4 Cross-Time Constraints

When more than two measurement occasions are available to model, it may be of interest to investigate whether the predictive relations across successive occasions are equal (e.g., if the paths from 7th to 8th grade are equivalent to those from 8th to 9th grade). This question can be answered by fitting two nested models within an SEM framework and comparing their fit. The first SEM model is an unconstrained model that simultaneously estimates auto-regressive and cross-lagged paths across all neighboring pairs of measurement occasions. This model is "unconstrained" because it allows paths to take on different values across subsequent intervals. It is important that zeroes be imposed consistently across each pair of measurement occasions in the SEM model to ensure that the next (constrained) model is nested within it. There is no guarantee that the initial lasso regression estimation step will produce a consistent pattern of zeroes across occasions (e.g., the path from variable 1 to variable 2 may be non-zero from $T1$ to $T2$, but zero from $T2$ to $T3$), so it is necessary to use some decision rule to choose which paths to set to zero consistently across all measurement intervals. A conservative rule would be to constrain a path to zero only if it is estimated to be zero across every neighboring pair of occasions. A more liberal rule would be to constrain a path to zero if it is estimated to be zero across the majority of intervals. We recommend using the conservative rule unless there are good reasons to prefer a more sparse/parsimonious network. Finally, when estimating the full model with multiple measurement occasions, residuals within each measurement occasion should be allowed to covary, to capture shared variance between each pair of variables in the network that can be attributed to effects within a time point (e.g., direct causal effects and common-cause effects that occur between measurement intervals).

The second (constrained) model is identical to the first but with cross-time constraints imposed, such that the set of paths from $T1$ to $T2$ is constrained to be equal to the set of paths from $T2$ to $T3$, and so on. Once both unconstrained and constrained models have been fit to the data, they can be compared using a nested chi-square difference test, in addition to approximate fit measures (e.g., $RMSEA_D$;

Savalei et al., 2024) and information criteria (e.g., AIC and BIC). Based on the results of these fit comparisons, the final selected model may be either unconstrained (different predictive relations over occasions) or partly constrained (the same predictive relations over time points).

2.5 In-Prediction and Out-Prediction

The final selected model (either unconstrained across occasions, or constrained, depending on the results of the model comparison in the previous step) can be summarized into two measures of variable centrality: In-prediction is the extent to which each variable is predicted by other variables in the network, and out-prediction is the extent to which each variable predicts other variables in the network. In-prediction is simply the proportion of variance in each variable at a measurement occasion that is accounted for by the complete set of variables at the previous occasion, and as such it can range in value from 0 to 1:

$$\text{inPred}_j = \frac{\text{var}(\hat{x}_{(j,T2)})}{\text{var}(x_{(j,T2)})} = \frac{\text{var}(\hat{\beta}_j' x_{T1})}{\text{var}(x_{(j,T2)})}. \tag{2}$$

In-prediction may be computed on a subset of predictors at the previous occasion, which can reveal more specific information. For example, *cross-lagged in-prediction* can be computed as the proportion of variance accounted for by all other variables at the previous measurement occasion, that is, cross-lagged in-prediction excludes the auto-regressive path. *Within-construct in-prediction* is the proportion of variance accounted for by all the variables within the same construct, excluding the auto-regressive effect. *Cross-construct in-prediction* is the proportion of variance accounted for by all variables that belong to a different construct at the previous measurement occasion.

Out-prediction captures the average proportion of variance across all variables at the next measurement occasion (e.g., $T2$) that is accounted for by a single target variable at the previous occasion (e.g., $T1$):

$$\text{outPred}_j = \frac{1}{p} \sum_{k=1}^{p} \frac{\text{var}(\hat{\beta}_{(x_{(k,T2)} x_{(j,T1)})} x_{(j,T1)})}{\text{var}(x_{(k,T2)})}, \tag{3}$$

where j indexes the target variable and $k = 1, \ldots, p$ indexes the set of outcome variables. Like in-prediction, out-prediction values can range from 0-1, though they will typically be substantially lower than in-prediction (an out-prediction value of 1 would indicate that a variable accounts for all of the variance in all variables at the subsequent measurement occasion, which would require complete multicollinearity among the whole set of variables!). As with in-prediction, we can define narrower versions of out-prediction: *cross-lagged out-prediction* averages the predictive effects over all variables except the one that is doing the predicting, *within-construct out-prediction* averages the predictive effects across only those variables that belong to the same construct (excluding the auto-

regressive predictive effect) and *cross-construct out-prediction* averages the predictive effects across only those variables that belong to a different construct. Each of these in- and out-prediction measures can be computed from the model estimates produced in the previous step; no new models are estimated.

If the model includes more than two constructs, in-prediction and out-prediction could be computed separately for the effect of each variable on/from variables belonging to each other construct. All measures of in-prediction and out-prediction can be computed for each pair of adjacent measurement occasions. If cross-time constraints have been imposed, then these measures will be identical across occasions.

Knowing the stability of an estimate (e.g., its fluctuation in response to sampling variability) is important for calibrating interpretations, and as such, we can use bootstrapping to obtain stability intervals around each of the in- and out-prediction estimates. A bootstrapping approach starts by drawing (with replacement) M samples of size n from the original dataset, where n is the size of the original data set. In and out-prediction indices are calculated from each of the M samples to create a bootstrapped sampling distribution of these indices. These sampling distributions are then used to compute confidence intervals around each estimate[4].

3. SIMULATIONS

In the previous section, we outlined a proposed two-step estimation approach for the CLPN. We refer to the proposed method as the ℓ_1 SEM approach, because it follows a regularized regression step (ℓ_1 GLM) with SEM estimation. To test the performance of this approach, we compare it to each step individually: ℓ_1 GLM refers to the lasso regularized regression step alone (i.e., the final model estimates are the regularized linear regression coefficients without any re-estimation), and SEM refers to an approach in which SEM is used to estimate all possible coefficients of the full model (with no zeroes imposed), after which nonsignificant ($p > .01$) edges are pruned. The SEM model is not re-estimated after pruning. We also include a comparison to the panel VAR approach available in the psychonetrics package in R (Epskamp, 2020b; R Core Team, 2024). This latter comparison is important for a few reasons: First, panel VAR is already implemented in widely-used software. Second, panel VAR fits a model that is both more flexible and more constrained than the CLPN: it is more flexible in that it allows for stable trait-level variance (i.e., random intercepts) to explain between-person variances and covariances among the variables, and it is more constrained in that it does not allow temporal effects to be different across different pairs of time points. To assess the effects of these two features of the model, we conducted two simulation studies. In the first, no stable individual differences were generated; that is, all between-person associations were due to associations across time in the measured variables. In the second, we included additional stable variance in the population generating

[4] Researchers could also use this approach to get bootstrapped edge weights in addition to stability estimates.

model. In both studies, we compared populations in which the stationarity assumption held and did not hold.

3.1 Design

Each simulation study included two data generating conditions (stationary, nonstationary), two values of network density (*m* = 0.5 or 0.8), three values of model size (*p* = 5, 9, or 16 variables), and two values of sample size (*N* = 500 or 2000), resulting in $2 \times 2 \times 3 \times 2 = 24$ conditions. We generated 4 waves of data (4 measurement occasions) in all conditions. All R code to reproduce the simulations is available at https://osf.io/9h5nj.

To generate data, population β matrices were randomly sampled from an edge weight bank with values ranging from $[-0.3, -0.1] \cap [0.1, 0.3]$ increasing in increments of 0.005, with probability equal to the network density (e.g., a network with $.8$ density had 80% of values in its β matrix filled and the rest set to 0). The stationary process had a single population β matrix and covariance matrix across all 4 time points, while the nonstationary process had three different β matrices for four occasions. The stationary covariance matrix was constructed such that $cov(x_t) = cov(x_{t-1})$. We generated 20 different populations (i.e., 20 sets of randomly drawn coefficients) and generated 100 sample data sets from each, resulting in 2000 replications per condition within each process (stationary and nonstationary).

Each simulated dataset was analyzed according to each of the four methods: ℓ_1 SEM (i.e., the approach described in the previous sections), and ℓ_1 GLM, SEM, and Panel VAR. For the ℓ_1 GLM and the first step of the ℓ_1 SEM estimation, the λ parameter of the regularized GLM step was chosen using 10-fold CV. For the SEM and second step of the ℓ_1 SEM approaches, models were estimated using maximum likelihood estimation, then nonsignificant ($\alpha > .01$) paths were removed from the estimated solution. The Panel VAR model was fit using maximum likelihood estimation, after which paths not significantly different from zero at $\alpha = .01$ were fixed to zero and the model was re-estimated with those zeroes in place.

We examined five outcomes for each condition. *Convergence rate* is the proportion of replications that converged to a proper solution. *Sensitivity* is the proportion of true nonzero edges that are estimated to be nonzero. *Specificity* is the proportion of true zero edges that are estimated to be zero. *Matthew's Correlation Coefficient* (MCC) indexes the overall accuracy of the estimated presence vs. absence of edges. It is equivalent to the phi correlation coefficient between two binary variables (i.e., edge presence vs. absence) and can be written:

$$MCC = \frac{(TP)(TN) - (FP)(FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \tag{4}$$

where true positive rate (TP) is the proportion of nonzero edges correctly estimated to be nonzero,

true negative rate (TN) is the proportion of zero edges correctly estimated to be zero, false positive rate (FP) is the proportion of zero edges incorrectly estimated to be nonzero, and false negative rate (FN) is the proportion of nonzero edges incorrectly estimated to be zero.

Finally, *mean absolute error* (MAE) indexes the average distance between the estimated and true edge weights (the "edge distance") for the estimated regression coefficients in each condition, and is computed:

$$MAE = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{N_m(k_m)}\sum_{n=1}^{N_m}\sum_{i=1}^{k_m}|\hat{\beta}_{imn} - \beta_{im}|\right), \tag{5}$$

where M is the number of population models (i.e., 20), N_m is the number of successfully converged replications within each model m, and k_m is the number of nonzero edges (regression coefficients) in population model m.

3.2 Results

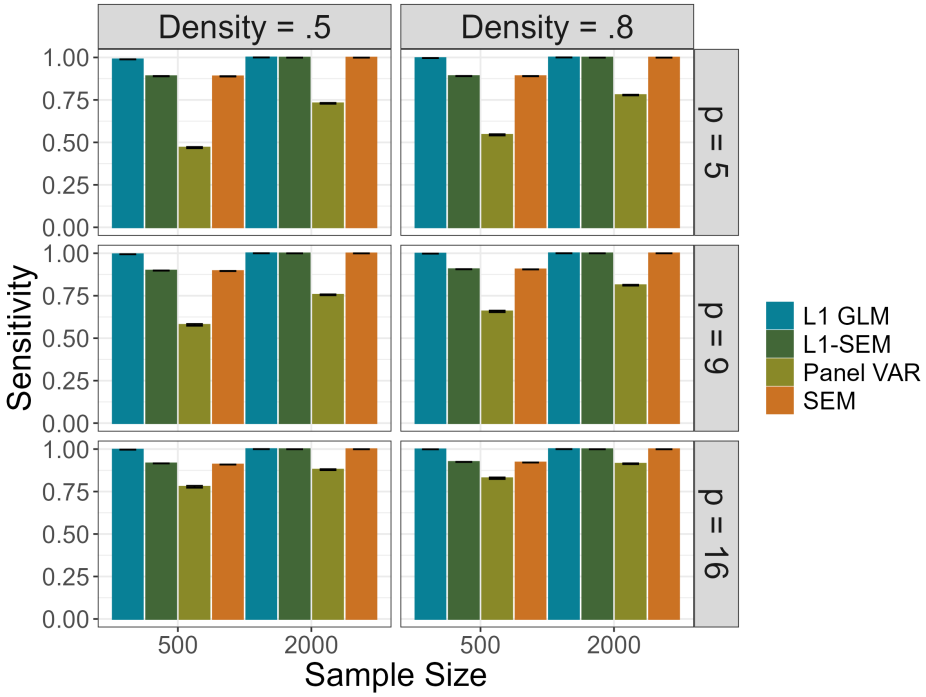
3.2.1 Convergence

The only method to show convergence problems was the Panel VAR approach, which did not produce estimation errors but produced extremely outlying parameter estimates. Across all conditions, the median absolute edge distance values ranged from 0.01 to 0.03, but the distributions across replications had long right tails. Edge distance values ranged as high as 595 in the nonstationary conditions, and 862 in the stationary conditions. To remove improper solutions, we excluded Panel VAR results that had a mean edge distance higher than 1.0. The proportion of excluded samples ranged from 0 to 3% and are fully displayed in Table S1 in the supplement (https://osf.io/9h5nj).

3.2.2 Network Structure Recovery

Figures 3 to 6 display the results for the nonstationary process. In terms of capturing the true structure, the SEM and ℓ_1 SEM clearly outperform the ℓ_1 GLM and Panel VAR. These two methods have the greatest balance of sensitivity (Figure 3) and specificity ((Figure 4), where there is a slight tradeoff for sensitivity in low sample size ($n = 500$) and low density ($m = 0.5$) for a near perfect specificity in all conditions. Additionally, these methods display a very strong MCC (Figure 5) of at least 0.75 in all conditions and a near perfect MCC in a large sample size ($n = 2000$), indicating a strong, positive correlation with the true structure of the generating models. The ℓ_1 GLM has near perfect sensitivity in all conditions, but similar to the Panel VAR, its specificity is lower than the benchmark of 0.8 across every condition. The Panel VAR has extremely low sensitivity (~ 0.5) in small parameter ($p = 5$) and small sample size ($n = 500$) conditions, but approaches and exceeds this benchmark in the remaining conditions. Additionally, with an MCC value lower than 0.25 in all conditions, it struggles to attain a balanced tradeoff between sensitivity and specificity.

Figure 3
Sensitivity of Estimated Network Structure Under Nonstationarity



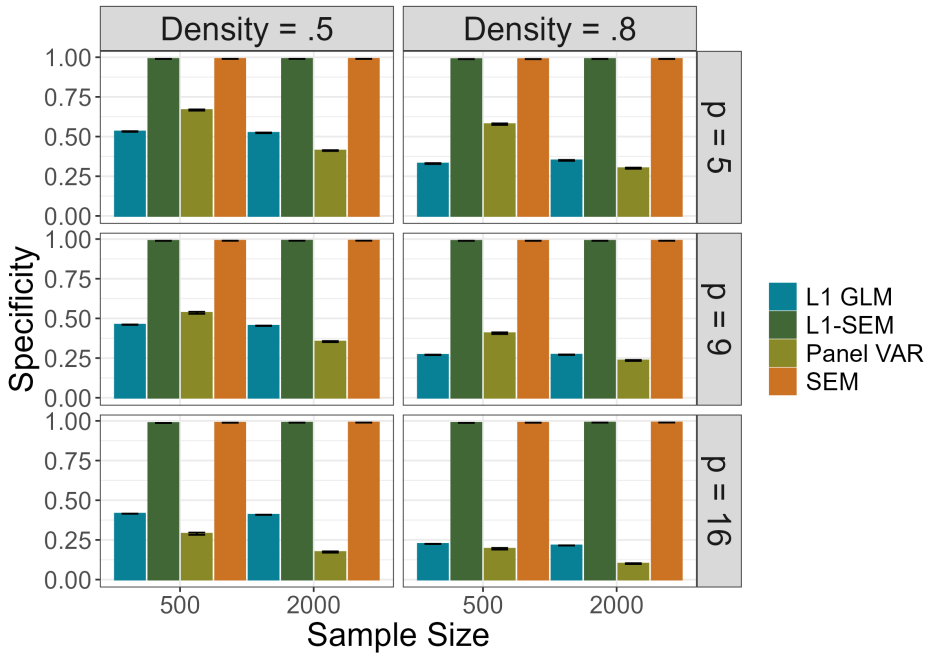
Note. Error bars represent ± one standard error from the mean.

For the MAE (Figure 6), we see the SEM and ℓ_1 SEM performing the best in a lower sample size ($n = 500$) and lower density ($m = 0.5$) setting by a small margin, with the ℓ_1 GLM catching up to perform equally well in a larger sample size ($n = 2000$). This indicates that these three methods generally estimate the true edge weights accurately across all simulated conditions. In contrast, the panel VAR model has at least double the MAE of the other methods in all conditions, thus indicating issue with accurately estimating true edge weights. This result is likely due to the violation of stationarity in the generating model; that is, while the true edges differ across time, Panel VAR estimates a single set of time-invariant edges and cannot capture those differences.

Figures 7 to 10 display the results for the stationary process. The ℓ_1 SEM, SEM, and ℓ_1 GLM all perform similarly in sensitivity (Figure 7), specificity (Figure 8), and MCC (Figure 9). The SEM and ℓ_1 SEM perform the best as a whole, but have slightly lower sensitivity with low density ($m = 0.5$) and small sample size ($n = 500$). They also lag behind Panel VAR in MCC for a small sample size ($n = 500$). While they are slightly outperformed in these conditions, they perform at least as well or better in the remaining conditions. Panel VAR's performance has significantly improved with the stationary assumption now satisfied. Across all conditions, it has a near perfect sensitivity and MCC, but its specificity drops below 0.8 when $p = 9$ and even lower than 0.5 when $p = 16$.

We see a similar performance in MAE for all but the Panel VAR. With a small parameter count (p = 5)

Figure 4
Specificity of Estimated Network Structure Under Nonstationarity



Note. Error bars represent ± one standard error from the mean.

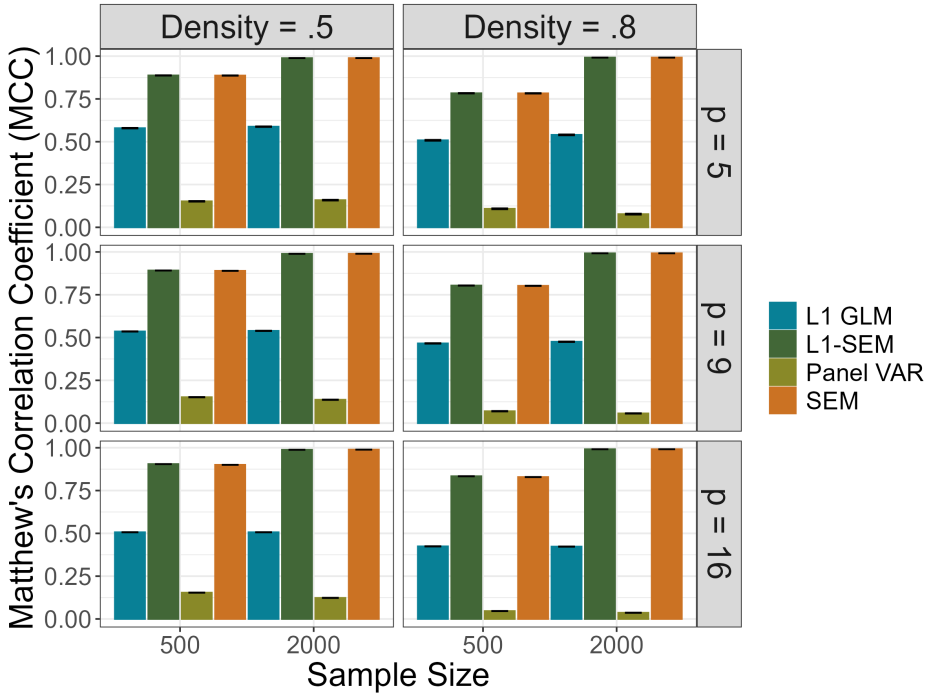
and a small sample size (n = 500), Panel VAR has the lowest MAE out of all of the methods. In p = 9, we can see that its edge distances begin to increase past the other methods, and it surges in p = 16. In these conditions that Panel VAR struggles in, the other methods perform nearly equally.

Thus, in all cases, the SEM and ℓ_1 SEM methods perform the best. They display the greatest balance between sensitivity and specificity, achieve a strong, positive MCC, and at least as good of a MAE as the competing methods. If one were to use the ℓ_1 GLM in a small parameter context (p = 5), it may be sufficient, but with an inadequate specificity and MCC, it is not the optimal choice. If the generating process is known to be stationary, then the Panel VAR would be a better choice compared to the ℓ_1 GLM, but only in this small parameter environment.

4. SIMULATION 2: STABLE PERSON MEANS

In the simulation study described in the previous section, we generated data from CLPN model parameters, that is, under the assumption that the CLPN is sufficient to accurately describe the patterns of variation in the world. But the CLPN, like the CLPM that it is based on, famously does not account for stable between-person variance. If individual people have different baseline levels of each variable (e.g., Angélique has a higher baseline enjoyment of school than Anna that contributes to their observed enjoyment of school at every occasion), the CLPN has no parameters to account for these

Figure 5
Matthew's Correlation Coefficient Under Nonstationarity



Note. Error bars represent ± one standard error from the mean.

effects, so those effects will lead to bias in the autoregressive and cross-lagged parameter estimates (Freichel et al., 2024; Hamaker et al., 2015; Lucas, 2023). Of the methods we tested, the panel VAR approach does include person-level means, so it would be expected to perform much better when stable means are included in the generating model. In fact, in the conditions we simulated, the panel VAR model may have been disadvantaged precisely because it had to estimate person-level means and covariances that were zero in the generating model.

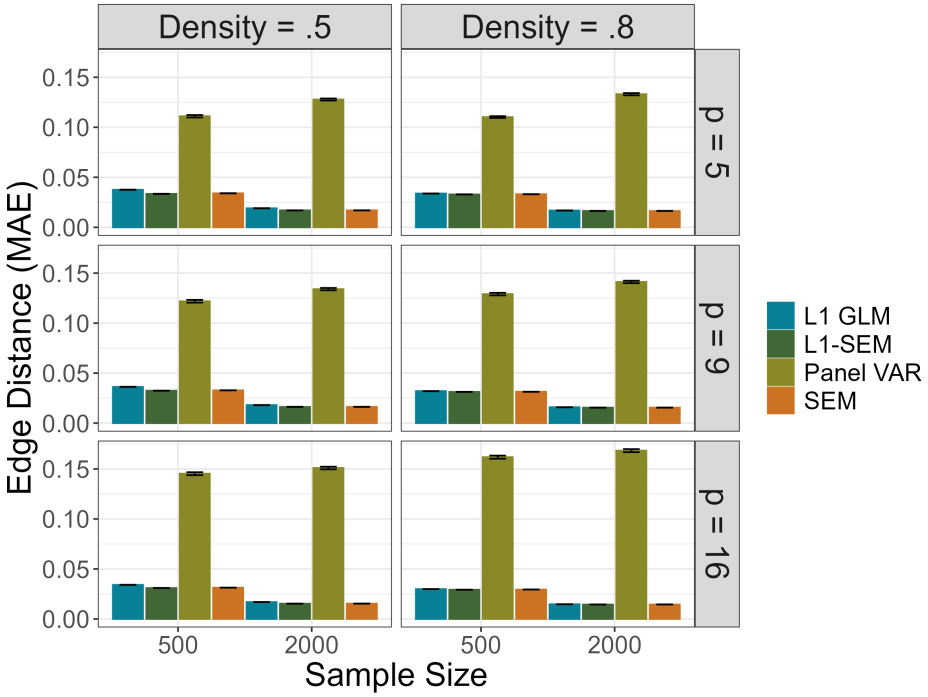
To test these hypotheses, we conducted a second simulation study that was identical to the first but with person-level effects in the generating model. In the generating model, the amount of between-person stable variance was equivalent to the variance accounted for by the CLPN parameterization, such that half the variance of each variable was unaccounted for by the CLPN model. Additionally, small between-person covariances were generated using the *clusterGeneration* R package (Joe, 2006; Qiu & Joe., 2023) based on a randomly drawn set of eigenvalues for each model.

4.1 Results

Figures depicting all results of Simulation 2 are in the supplemental material, and we describe the general trends here.

In terms of convergence, we saw fewer convergence failures when the between-person variance and

Figure 6
Mean Absolute Error of Edge weights Under Nonstationarity



Note. Error bars represent ± one Monte Carlo standard error from the mean.

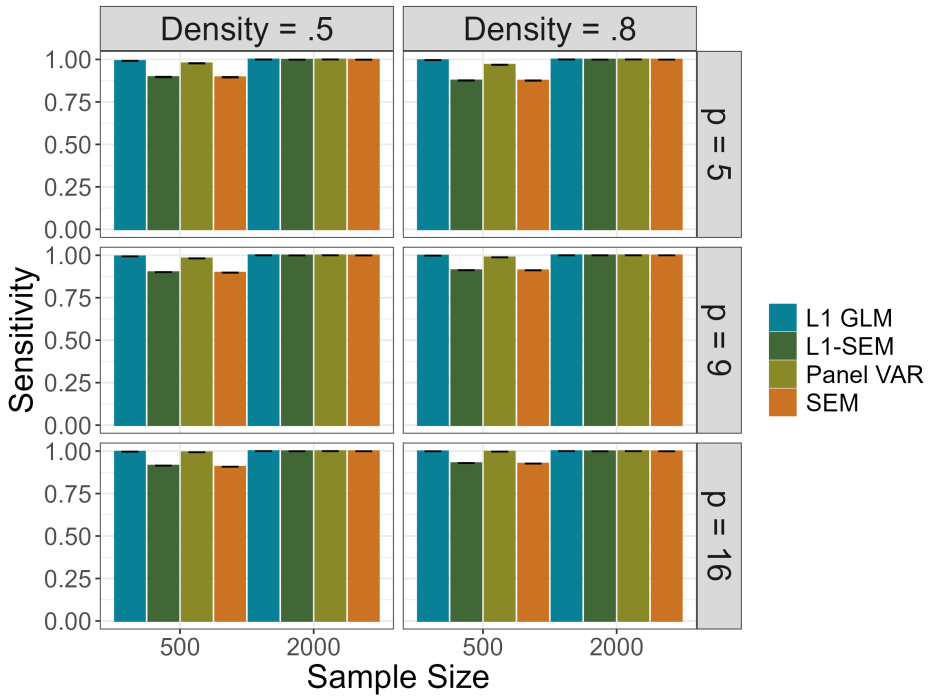
covariance components were not zero in the generating model (see Table S2). There were still some convergence failures, predominantly in the high density, high parameter conditions ($p = 16$, $m = 0.8$).

When the within-person process was nonstationary (Figures S2 to S5), we expect the Panel VAR to perform similarly to the first simulation, as its stationarity assumption is again not satisfied. We found that it performed very similarly across metrics measuring structural accuracy (sensitivity, specificity, and MCC). Its overall edge estimation accuracy was higher, with a much lower MAE in smaller parameter conditions ($p = 5, 9$). This increase in estimation performance did not appear with the other three methods, which perform worse across all metrics compared to the original nonstationary simulation.

With the stationary assumption satisfied (Figures S6 to S9), the Panel VAR clearly performs the best. It captures the overall structure of the population parameters very well, with a near-perfect sensitivity and very strong MCC across all conditions, and a specificity that passes the common benchmark of 0.8 in all conditions except for the high parameter (p = 16), high density conditions (m = 0.8), as well as the high parameter (p = 16), low density (m = 0.5), and low sample size condition (n = 500). It has the lowest MAE by far for all conditions except the high parameter (p = 16), high density conditions (m = 0.8). Similar to the nonstationary process, the other three methods perform worse across all metrics comparatively.

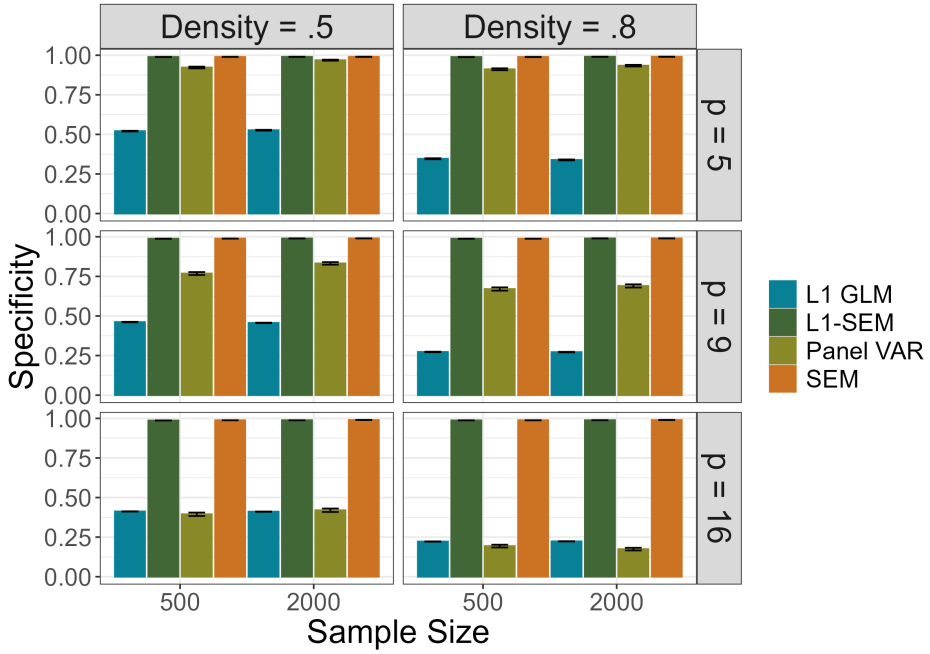
Ultimately, when there is a stable between-person variance and covariance structure for the Panel

Figure 7
Sensitivity of Estimated Network Structure Under Stationarity



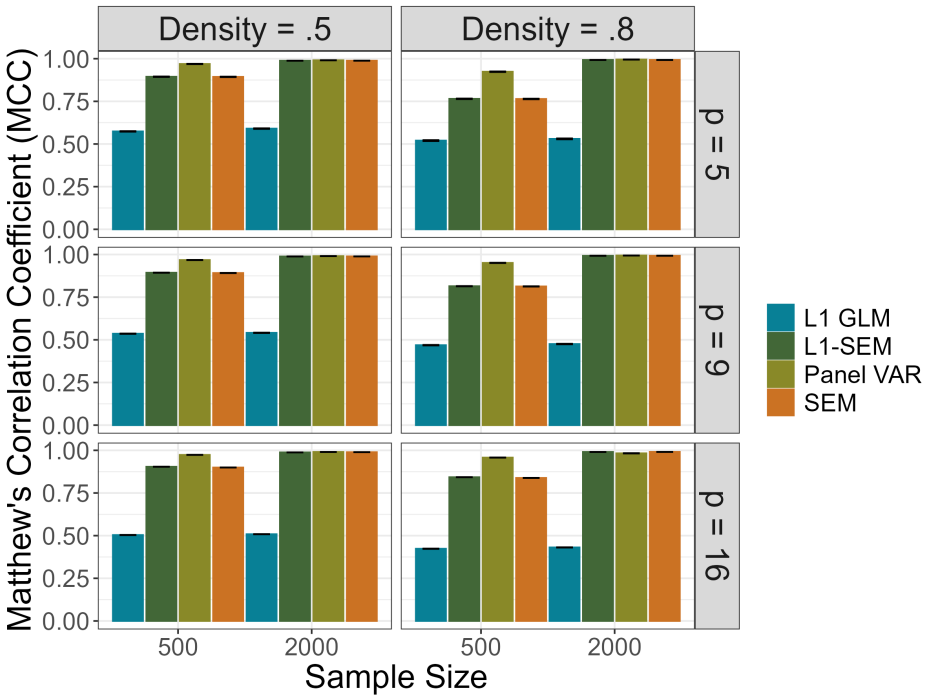
Note. Error bars represent ± one Monte Carlo standard error from the mean.

Figure 8
Specificity of Estimated Network Structure Under Stationarity



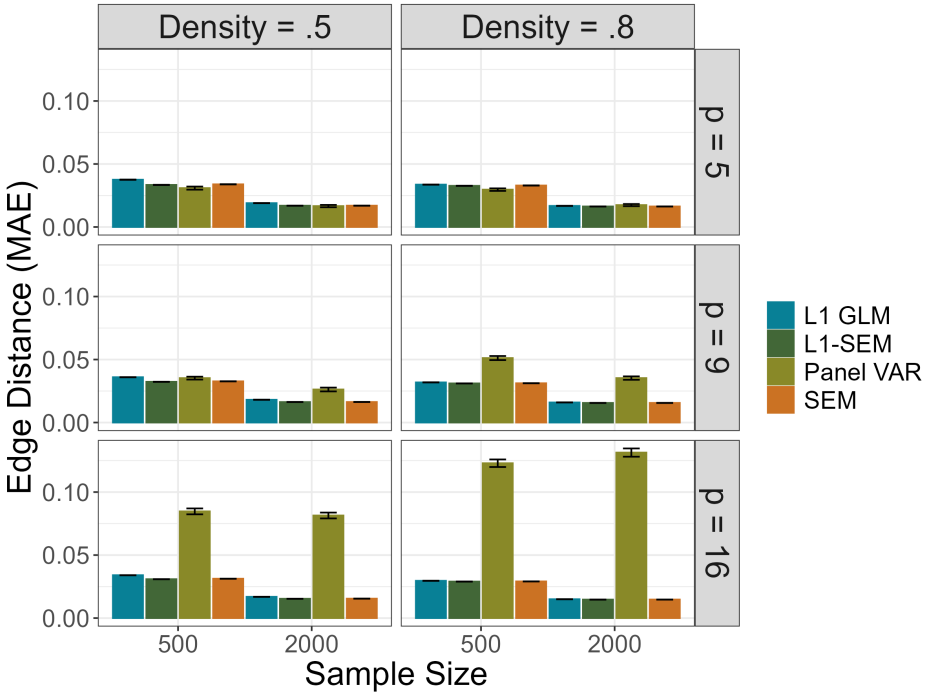
Note. Error bars represent ± one Monte Carlo standard error from the mean.

Figure 9
Matthew's Correlation Coefficient Under Stationarity



Note. Error bars represent ± one Monte Carlo standard error from the mean.

Figure 10
Mean Absolute Error of Edge weights Under Stationarity



Note. Error bars represent ± one Monte Carlo standard error from the mean.

VAR to estimate and the generating process is stationary, it is a very strong choice. Given its convergence issues and decreasing performance in high density, high parameter conditions, one should still proceed with caution in such an environment. If these conditions are not met, then it may be best to proceed with the ℓ_1 SEM or SEM methods.

5. EMPIRICAL EXAMPLE: COMMITMENT TO SCHOOL AND SELF-ESTEEM

To demonstrate estimation and interpretation of a cross-lagged network model, we used publicly available data from the Iowa Youth and Families Project (Conger et al., 2011) on the Commitment to School and Self-Esteem scales. We chose these constructs because they each plausibly feature causal interactions among the attitudes that are measured by individual scale items. For example, the Commitment to School scale measures the extent to which children do well in school ("I don't do well at school") and their enjoyment of school ("I like school a lot"), which are likely to affect each other directly.

5.1 Data

The Iowa Youth and Families Project collected yearly data from an initial sample of 451 students who were in 7th grade in 1989 until 1992. Commitment to School was measured by seven items developed by Robert Conger for the Iowa Youth and Families Project, and Self-Esteem was measured with the Rosenberg Self-Esteem Scale (Rosenberg, 1965). Participants answered each item on a 5-point scale on which the endpoints corresponded to "Strongly Agree" and "Strongly Disagree"; items were coded such that lower scores correspond to higher levels of commitment to school or self-esteem. Table S3 in the supplement displays the final set of items that were included in the CLPN analysis.

Data on these constructs are available at 4 waves, each separated by one year (1989 – 1992). Of the 451 students who participated in the first wave, 424, 407, and 403 participated in waves 2-4, respectively; 395 students provided data at all 4 waves, 11 of these students were missing data on a single item at one wave. Because *glmnet* does not accommodate incomplete cases, in the first step of the analysis (regularized regression analyses with *glmnet*) we remove all cases with missing data. In subsequent steps (SEM analyses) missing data is handled using full information maximum likelihood estimation.

5.2 Analyses and Results

All analyses were done in R (R Core Team, 2024). CLPN regressions were estimated with the *glmnet* package (Friedman et al., 2010) and the *lavaan* package for SEM (Rosseel, 2012), and networks were plotted using the *qgraph* package (Epskamp et al., 2012). R code to reproduce all analyses and CLPN figures is available at osf.io/9h5nj.

5.2.1 Latent Variable CLPM

For the sake of providing a baseline comparison, we begin by presenting the traditional CLPM based on latent variables. We defined two latent variables, Commitment to School and Self-Esteem, each indicated by the set of items presented in Table S3, at each of the four measurement occasions. To ensure that the latent variables are comparable over time, we imposed weak measurement invariance; that is, the unstandardized factor loadings were constrained to be the same at all four measurement occasions. We fit the model in lavaan, using full information maximum likelihood to deal with missing data, and robust corrections to the standard errors and test statistic to deal with violations of the normality assumption (estimator = MLR). The weak invariance constraint did not significantly reduce fit compared to the model that allowed loadings to differ over time, $\chi^2_\Delta(df = 45) = 47.75$, $p = .36$. We further constrained the auto-regressive and cross-lagged paths to be constant across each pair of measurement occasions; this model fit significantly worse than the unconstrained model, but an examination of approximate fit indices and information criteria suggested that its fit was not substantially worse, and the BIC suggested superior fit of the constrained model: $\chi^2_\Delta(df = 8) = 25.00$, $p = .002$, $\Delta AIC = 11.46$, $\Delta BIC = -21.43$, $RMSEA_D = .069$.

As is typical in latent variable CLPM analyses, we allowed each item's residual to covary across measurement occasions, to account for stable item-specific variance that is unrelated to the latent factor (Little, 2013; Little et al., 2007; Mayer, 1986). The resulting model with cross-time constraints is depicted in Figure 11 (only two measurement occasions are shown because all path coefficients are identical across time). The model fit was reasonable by conventional criteria: $\chi^2(df = 2145) = 3984.08$, $p < .001$, $RMSEA = .047(.044, .049)$, $CFI = .86$, $TLI = .85$. This analysis revealed small cross-lagged coefficients, only one of which was significantly different from zero at the $\alpha = .05$ level ($SE \rightarrow SC$, $\beta = .09$, $p = .046$; $SC \rightarrow SE$, $\beta = .04$, $p = .14$). On the basis of this analysis, one would conclude that the unique predictive effects of school commitment on self-esteem and vice-versa are weak to non-existent.

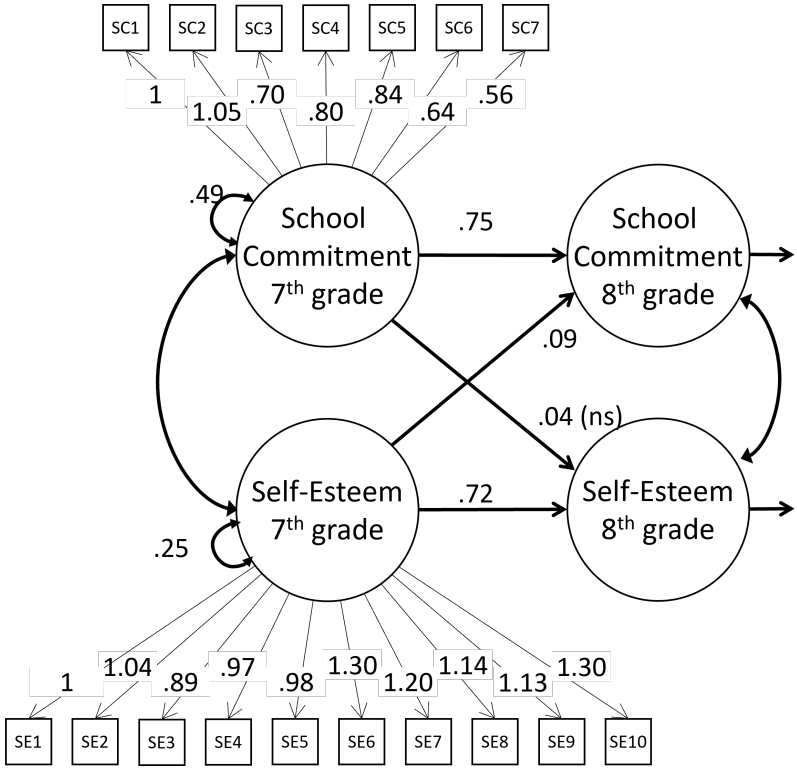
5.2.2 CLPN

Next, we followed the steps presented in the Methodological Details section to fit a cross-lagged panel network model. We opted not to collapse any overlapping variables, because we did not see items that appeared to be semantically redundant. Following the steps outlined above, we first fit a series of regularized regressions to obtain our initial model. Next, we re-estimated it in a SEM framework allowing the residuals of variables within the same occasion to covary[5]. The model displayed good approximate fit; $\chi^2(df = 1277) = 1653.76$, $p < .001$; $RMSEA = .025(.02, .029)$; $CFI = .97$, $TLI = .95$.

The significant ($p < .01$) paths from the non-regularized analysis are plotted in Figure 12 as three di-

[5]These covarying residuals capture shared variance between each pair of variables in the network that can be attributed to effects within a time point (e.g., direct causal effects and common-cause effects that occur between measurement intervals.

Figure 11
Latent Variable CLPM Results

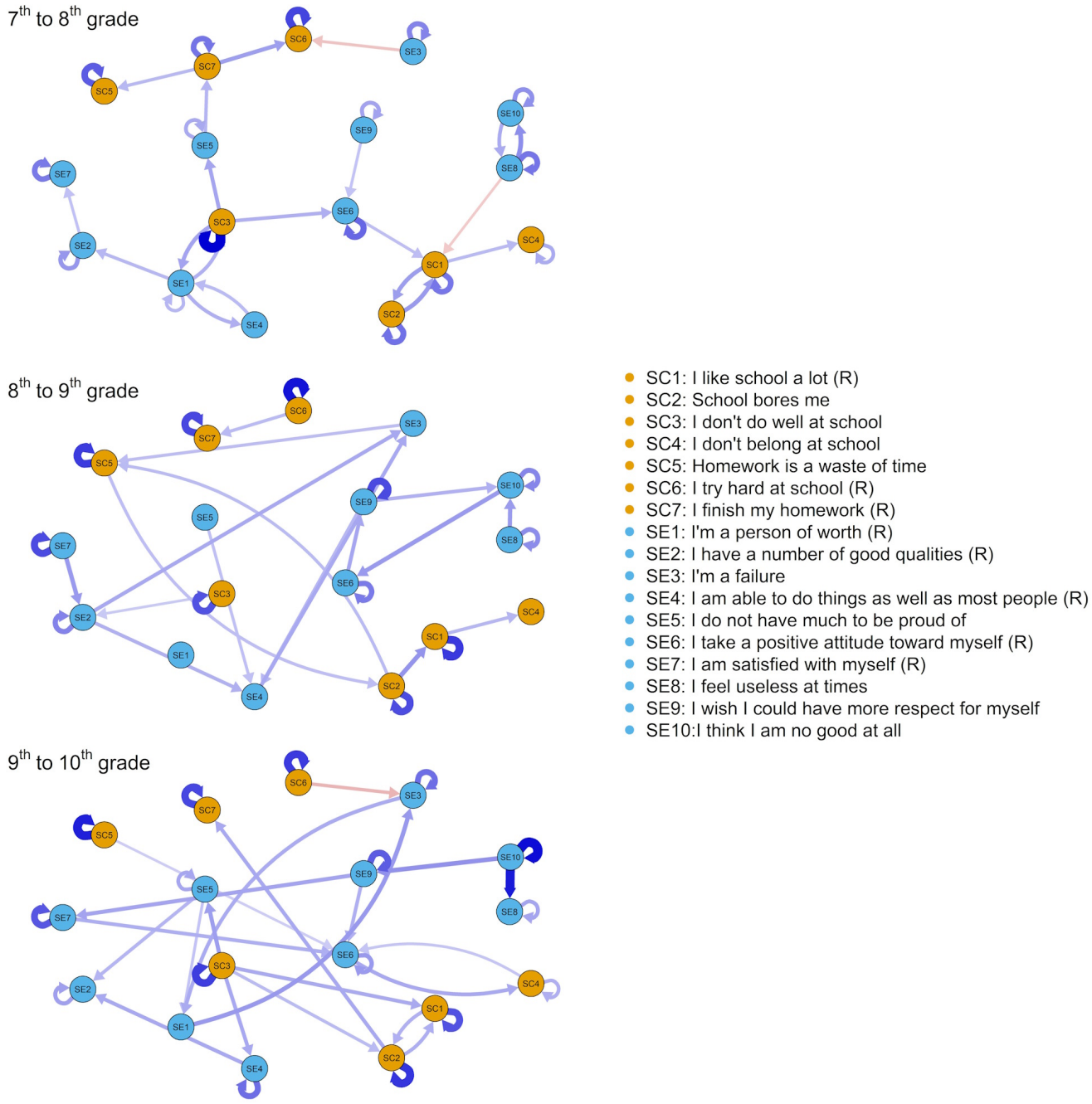


Note. For clarity, only the first two measurement occasions are depicted and the measurement model is only shown for the first occasion. All 4 years were included in the model; auto-regressive and cross-lagged paths were constrained to be equal across each pair of occasions. Unstandardized factor loadings were also constrained to be equal at each measurement occasion. Residuals of each item were allowed to covary across occasions (not shown). Residual variances were free to vary across occasions.

rected networks, in which arrows represent cross-time effects (e.g., $SE4 \rightarrow SC4$ represents a path from SE4 at one measurement occasion to SC4 at the subsequent occasion). The arrows that start and end at the same node represent auto-regressive paths (e.g., the path from SE4 at one measurement to SE4 at the subsequent measurement occasion). Arrow thickness/darkness represents the relative strength of these effects (thicker arrows represent stronger relations) and color represents the direction of the effect (blue arrows represent positive effects). The placement of the variables is determined by the Fruchterman-Reingold algorithm described earlier, applied to the first measurement interval (i.e. 7th to 8th grade) and fixed to have the same layout for the subsequent two intervals for ease of comparison.

Similar to the CLPM results, the auto-regressive paths (e.g., the effect of SC1 at 7th grade on SC1 at 8th grade) tended to be stronger than cross-lagged paths. The auto-regressive paths capture all stability in individual items over time, including construct-related stability as well as any stable unique variance (e.g., a response set that leads participants to answer questions in idiosyncratic ways may be stable across measurement occasions—such an effect would be expected to appear in the auto-regressive paths).

Figure 12
CLPN Estimated Using the Hybrid L1-SEM Approach



The next modeling step is to test whether cross-time constraints can be imposed across the 3 measurement intervals. To do this, we first fit an unconstrained CLPN model, where only the paths that were estimated to be zero at all three measurement intervals were fixed to zero, and every other cross-lagged and autoregressive path was freely estimated. We again used the full-information estimation method with robust corrections to the fit statistics. This model fit well; $\chi^2(df = 987) = 1483.22$,

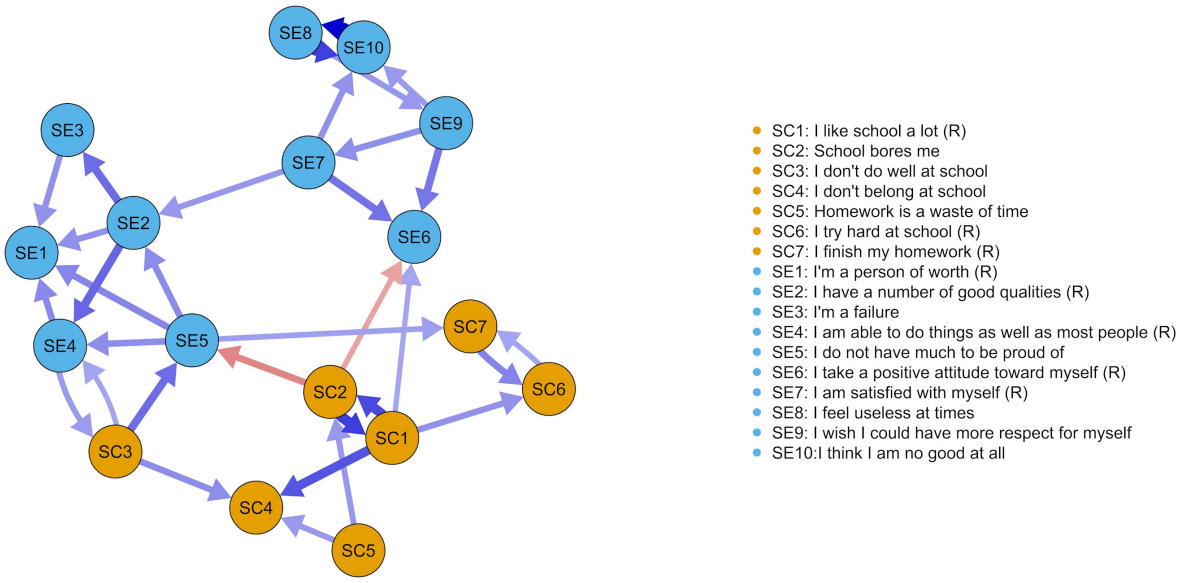
$p < .001$; $RMSEA = .033(.029, .037)$; $CFI = .96$, $TLI = .92$ (robust statistics reported). Second, we fit a model with cross-time constraints imposed, and found similarly good fit, $\chi^2(df = 1485) = 2197.63$, $p < .001$; $RMSEA = .033(.030, .036)$; $CFI = .94$, $TLI = .92$. Because the model with cross-time constraints is nested within the unconstrained model, we can compare the models statistically. The chi-square difference test finds a significant difference, $\chi^2_\Delta(df = 498) = 714.4$, $p < .001$. This significant difference suggests that adding the cross-time constraints significantly reduces model fit. However, as the chi-square test is known to be sensitive to large sample sizes, it is useful to also compare fit using approximate measures and information criteria. $RMSEA_D$, which indexes the degree of additional misfit induced by the constraints, is $.031$, well within the bounds of what is typically considered good fit (Savalei et al., 2023). Moreover, both the AIC and BIC, which weigh parsimony in addition to fit, suggested that the constrained model fits better. On the whole, thus, adding the cross-time constraints does not appear to appreciably worsen fit, and the more parsimonious model may improve the generalizability of the results. Thus, we accepted the model with cross-time constraints as our final model.

Figure 13 shows the model with cross-time constraints imposed. Due to these constraints, a single network picture depicts the predictive relations from each measurement occasion to the subsequent occasion. In this figure, paths not significantly different from 0 ($p > .01$) are not shown, nor are autoregressive paths.

Inspection of Figure 13 reveals several interesting cross-lagged paths that did not appear in the latent variable. These paths may shed light on the ways in which commitment to school can affect self-esteem and vice-versa. For example, boredom at school (SC2), school performance (SC3), self-worth (SE5), and attitude towards self (SE6) appear to be "bridge nodes" that may connect these two constructs. School performance is predicted by the beliefs that one has worth (SE1) and that one is competent (SE4), and it in turn predicts the belief that one doesn't have much to be proud of (SE5) and that one is a person of worth (SE1). School performance is also related to other components of commitment to school, for instance by predicting belongingness at school (SC4) and enjoying school (SC1), and being predicted by homework completion (SC7) and belongingness at school (SC4). Self-worth is predicted by school performance (SC3) and boredom at school (SC2) and predicts perceived value of homework (SC5) and effort at school (SC6). A model that treats all of these components as interchangeable markers of a single underlying construct will necessarily miss out on these nuances.

When interpreting the network pathways, it is important to remember that the paths represent conditional predictive relations, controlling for all other variables at the previous measurement occasion. Remembering this can help to interpret counterintuitive-seeming paths, like the significant negative paths from finding school interesting (SC2) to positive self-attitude (SE6) and being proud of oneself (SE5): holding constant other variables in the network (e.g., school performance and enjoyment), students who report being bored with school see themselves more positively than those who do not

Figure 13
CLPN Estimated Using the Hybrid L1-SEM Approach with Cross-Time Constraints Imposed



Note. Auto-regressive paths not shown. Relations across subsequent measurement occasions were constrained to be equal across each pair of subsequent measurement occasions.

report being bored. While these negative paths appear to contradict the general positive associations among commitment to school variables (e.g., finding school interesting strongly positively predicts and is predicted by enjoyment of school), they might reflect a tendency for students who see themselves as having mastered the material to feel bored in class. It is also important to keep in mind that these paths represent between-subjects associations and may not represent associations that hold within any individual over time (e.g., if bored students see themselves more positively, it does not follow that any individual student will see themselves more positively in contexts in which they feel more bored). We return to the important issue of between-person and within-person effects in the Discussion.

Finally, we note two pairs of variables that have strong predictive relations with each other: Feeling useless (SE8) and thinking of oneself as no good (SE10) and enjoying school (SC1) and being bored by school (SC2). These pairs of items are good candidates for variables that could perhaps have been collapsed into a single item before fitting the network models. As mentioned earlier, we elected not to collapse any variables because that is best undertaken with expertise in the content domain. But this result suggests what can happen if that step is skipped, namely that associations between each of these nodes and the remainder of the network may be suppressed. For example, the lack of outgoing paths from SE10 to anything else in the network means that, controlling for SE8, SE10 had

no further significant predictive relation to any other variable. If SE8 and SE10 had been collapsed into a single variable, SE8 would not suppress the estimated influence of SE10 [6].

In-prediction and out-prediction indices provide a numerical summary of the influence of individual nodes in the network. We also obtained stability intervals around each of the prediction metrics via bootstrapping. We bootstrapped 800 samples from the original dataset, calculated prediction indices for each of the samples, and then computed the standard deviation of each variable's prediction indices across the bootstrapped samples. The error bars in Figures 14-16 represent $\pm$ one standard deviation around the prediction indices. Figure 14 displays the cross-lagged in-prediction and out-prediction indices for each variable in the CLPN; these indices reveal the extent to which each variable is predicted by (in-prediction) or predicts (out-prediction) all other variables at the previous/subsequent measurement occasion. Figures 15 and 16 display the within-construct and cross-construct versions of these indices, which limit the set of predictors/outcomes to those from within the same construct (within-construct, Figure 15) or to those from the other construct (cross-construct, Figure 16).

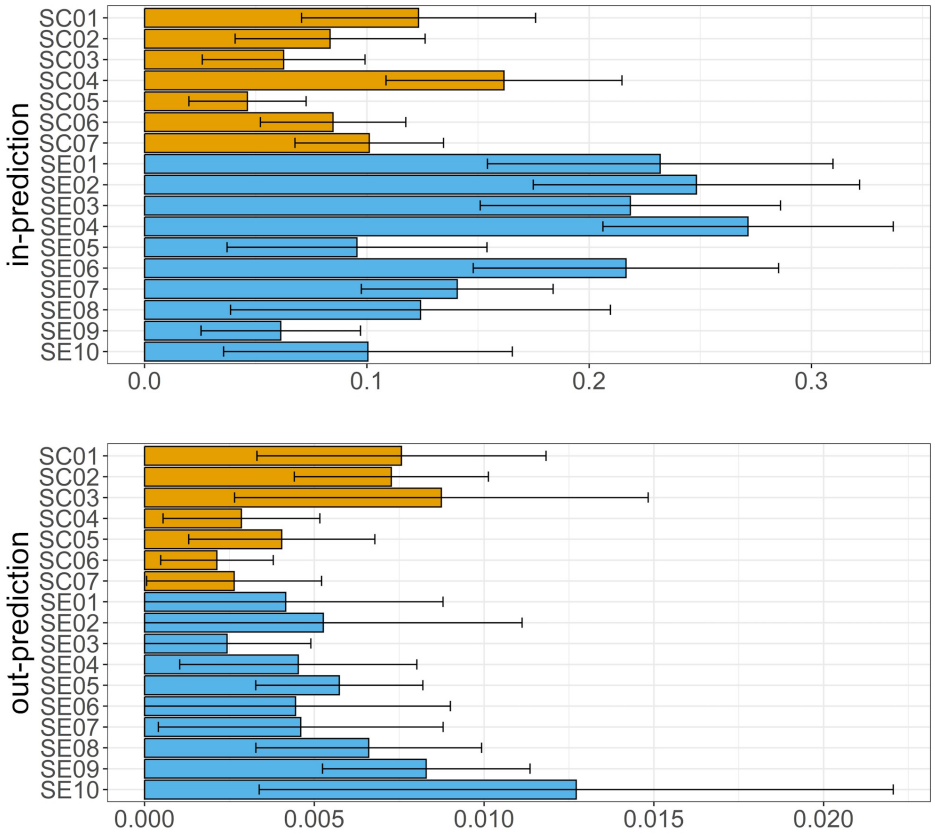
Figure 14 shows that belief in one's abilities (SE2) and competence (SE4) are the most highly predictable beliefs (i.e., they have the highest in-prediction). Overall, school competence items tend to have lower in-prediction than self-esteem items. On the other hand, boredom at school (SC2) and believing oneself to be no good (SE10) are the most useful for predicting other variables in the network (i.e., they have the highest out-prediction), but both of these variables have low stability in their out-prediction indices.

Considering within-construct paths only (including the auto-regressive paths; Figure 15), we find that enjoyment of school (SC1), boredom at school (SC2), belief in one's own abilities (SE2), attitude towards yourself (SE6), and satisfaction with self (SE7) are the variables that are best predicted by other within-construct variables. From the stability intervals, we can also see that most of the self-esteem variables have low stability. We also see that enjoying school (SC1), boredom at school (SC2), school performance (SC3), and perception of one's worth (SE10; low stability) are the best predictors of other within-construct variables.

Figure 16 shows the cross-construct prediction indices: Here, we see that the commitment to school variable that can be most reliably predicted by self-esteem variables is having a sense of belonging at school (SC4), and the self-esteem variables that can be most reliably predicted by commitment to school variables is the belief in one's competence (SE4) and the feeling of having much to be proud of (SE5)—although SC4 and SE5 have low out-prediction stability. In terms of out-prediction, the most reliable predictor of self-esteem is school performance (SC3; low stability). We can see that there are no reliable cross-construct predictors of commitment to school (all self-esteem variables have low

[6] Indeed, we found that, when we collapsed SE08 and SE10 into a single node and re-ran the CLPN analysis, the composite of SE08 and SE10 had more connections than either of the original variables

Figure 14
Cross-Lagged In- and Out-Prediction of Variables in the Estimated CLPN



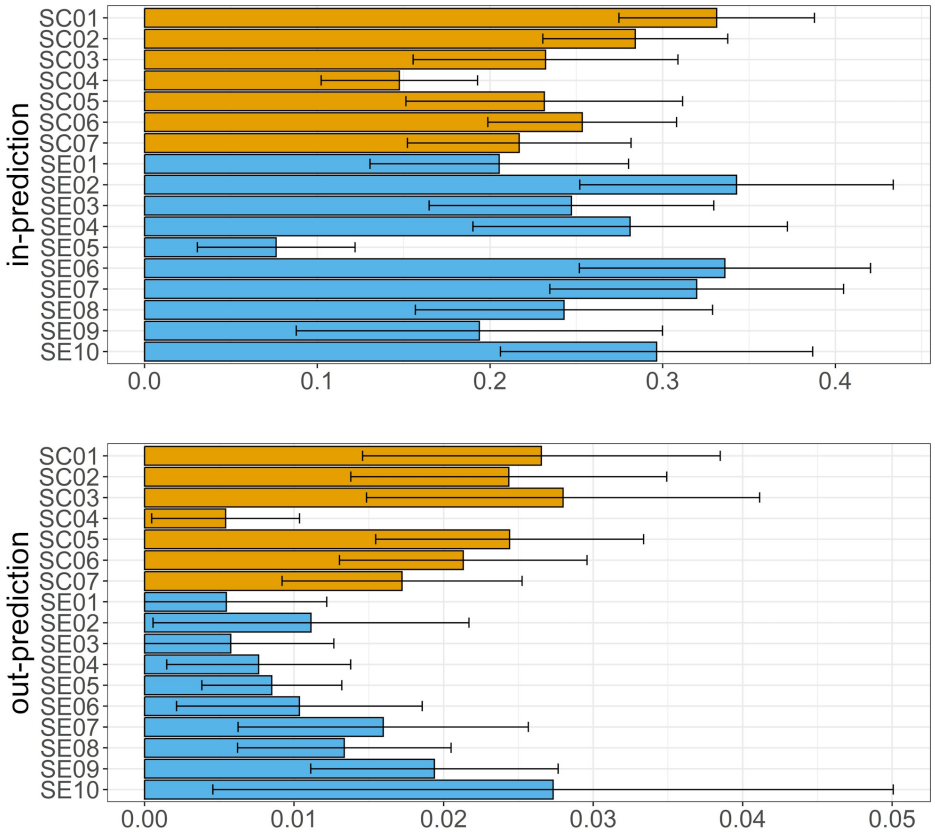
Note. Error bars represent -/+ 1 standard error for each of the prediction indices calculated from the bootstrapped samples.

stability and/or small prediction index values). and the most reliable predictors of commitment to school are feeling like no good (SE10) and feeling like one has much to be proud of (SE5).

6. DISCUSSION

In this paper, we proposed a new model for discovering longitudinal predictive effects of psychological constructs on each other based on data from two or more measurement occasions. The CLPN combines the benefits of latent-variable cross-lagged panel models with the benefits of network representations of psychological constructs to produce a model that can suggest which components of a construct may be involved in predicting change in another construct, and which components of a construct are most susceptible to change from components of other constructs. This work builds on the growing network modeling toolbox in which models are focused on individual variables rather than their aggregates. The resulting model may reveal not only the extent to which constructs forecast change in each other, but also much more fine-grained detail that could shed light on the relations between the components of the constructs.

Figure 15
Within-Construct Cross-Lagged In- and Out-Prediction of Variables in the Estimated CLPN



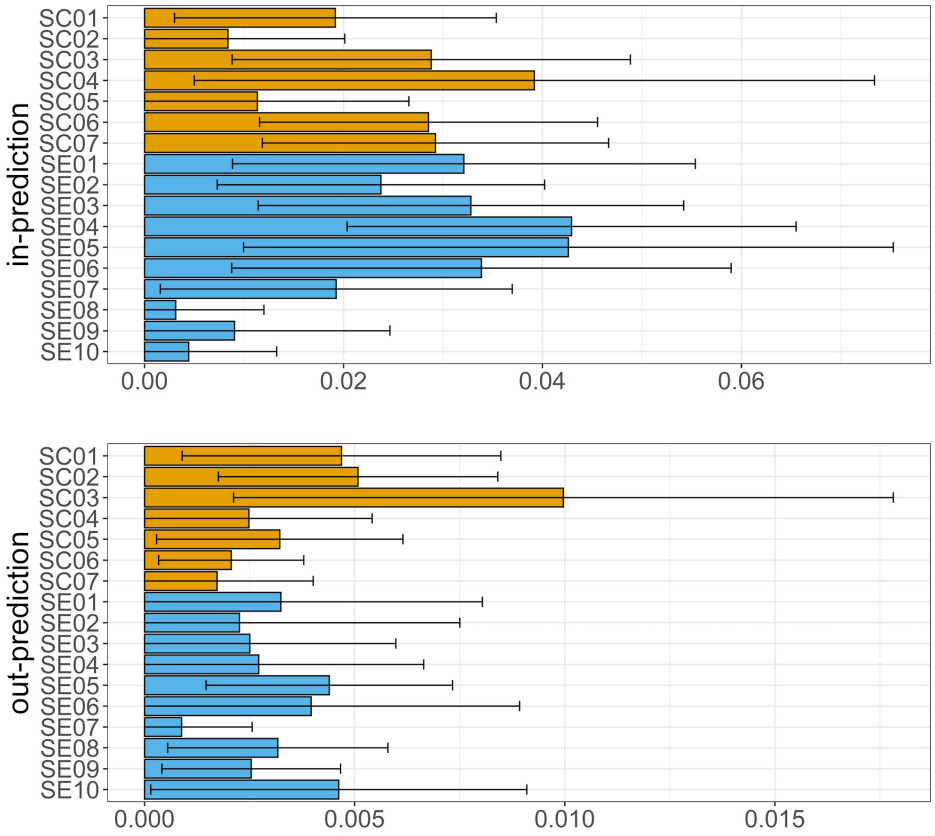
Note. Error bars represent -/+ 1 standard error for each of the prediction indices calculated from the bootstrapped samples.

To be clear, we do not wish to argue that the CLPN is necessarily a better modeling choice than latent-variable CLPM. Both models may provide interesting findings, and one or the other may be more appropriate for a given set of items and for a given *a priori* hypothesis or theory about the nature of the construct under investigation. In the example we used here, the latent variable CLPM returned non-significant cross-lagged paths among the latent variables, whereas the CLPN produced many potentially interesting item-level effects. In a situation where items are truly reflective of a single underlying construct, we would expect the CLPM to have greater power to show cross-lagged effects, and for the CLPN to reveal few differential predictive patterns among the items.

6.1 Methodological Recommendations

In our simulations, SEM and ℓ_1 SEM exhibited the same performance across all simulated conditions, indicating that simply fitting the CLPN in SEM without bothering with the initial ℓ_1 step may be simpler and equally effective. In cases where the dataset is too small to allow the full SEM model to be fit, using the ℓ_1 as a first step may induce sufficient sparsity to allow the SEM to be fit.

Figure 16
Cross-Construct Cross-Lagged In- and Out-Prediction of Variables in the Estimated CLPN



Note. Error bars represent -/+ 1 standard error for each of the prediction indices calculated from the bootstrapped samples.

A harder question is whether to chose the CLPN approach or the panel VAR approach implemented in *psychonetrics*. A strength of the CLPN is that it allows for relations to change across time, a feature that may be intrinsically of interest in a panel design where measurement occasions are separated by long periods of time and/or in cases where the data collection period spans multiple developmental periods. But this strength is countered by the CLPN assumption that there are no stable individual differences in the modeled variables. An empirical approach to this decision could be to begin with the CLPN and test the stationarity assumption by testing whether constraining the parameters to be the same across time (not only the lagged relations but also means and covariances within time points) leads to significant decrease in fit, and move to the panel VAR model if these constraints appear to be tenable. Another approach would be to use model fit indices (available in the SEM framework and also in psychonetrics) to assess the fit of either model. We caution, however, that no research has yet investigated the sensitivity of fit indices to violations of either stationarity (in the case of panel VAR) or to the presence of stable individual differences (in the case of the CLPN). Future research should also investigate what kind of violations of stationarity (e.g., changes in lagged parameters or changes in means or covariances within timepoints) warrant the use of CLPN and in

what cases it is better to switch to the panel VAR model.

6.2 Challenges and Extensions

Both the traditional cross-lagged panel model and cross-sectional network models are known to have several limitations that also affect CLPNs. We describe some of the major issues here and consider how they might be dealt with.

6.2.1 Time Lag

Estimated autoregressive and cross-lagged regression paths describe effects at a particular time lag, which may not be appropriate for some or all of the effects under investigation (Cole & Maxwell, 2003; Deboeck & Preacher, 2016; Gollub & Reichardt, 1987). For example, suppose a researcher wants to study the reciprocal effects of school motivation and school achievement. Suppose further that the motivation $\rightarrow$achievement pathway takes days or weeks (e.g., motivation is tied to particular assignments, topics, tests), whereas the achievement $\rightarrow$motivation pathway happens on the scale of semesters (e.g., school grades from the previous semester affect motivation in the current semester). If waves of data collection are separated by weeks, the researcher might conclude that motivation affects achievement but not vice versa; if instead waves are separated by semesters, the researcher could conclude the opposite. It is thus essential to choose the time-lag thoughtfully, and to interpret the effects in light of the time-lag represented in the data. For research questions pertaining to smaller timescales than those at which panel data tend to be collected, it is advised to move toward intensive longitudinal designs, for which many sophisticated methods are available.

6.2.2 Causation and Mediation

Cross-lagged panel models are frequently used to assess hypotheses about longitudinal mediation effects, but there are two typically untenable assumptions that are necessary to estimate the degree of causal mediation from an indirect path ($X \rightarrow M \rightarrow Y$; Bullock et al., 2010; Fiedler et al., 2011; Imai et al., 2010; Judd & Kenny, 1981; Pearl, 2012). First, when levels of X are not randomly assigned to participants, it is necessary to evaluate the assumption that, given any observed covariates, X is independent of potential values on M and Y (Imai et al., 2010); which is another way of saying that X is independent of the residuals of M and Y (Bullock et al., 2010). This assumption is untenable whenever any unaccounted-for third variable simultaneously affects levels of X and M, or X and Y.

Second, unbiased estimation of a causal mediation effect requires the assumption that the mediator, M, is independent of potential values of the dependent variable, Y, given the observed values of X and any observed covariates. In other words, M is uncorrelated with any other variable that may influence Y. Unless both X and M are experimentally manipulated and randomly assigned, these assumptions

are untestable and typically untenable (Bullock et al., 2010; Judd & Kenny, 1981).

In addition to these assumptions, inferring causality from predictive paths also requires that the measurement time-lag corresponds to the timing of the causal effect (see previous section) and that all the variables are measured without error (see "Measurement Error", below). As these assumptions, taken together, are typically impossible to evaluate, it is rarely if ever legitimate to infer causality from cross-lagged coefficients.

While experimental manipulation of X and M is the best way to ensure that causal mediation conclusions are warranted, many psychological constructs do not allow for experimental manipulation (e.g., neither commitment to school nor self-esteem can be independently manipulated). In such cases, it is recommended to acknowledge these assumptions and to measure as many omitted variables as possible to control for confounders. In addition, Imai et al. (2010) propose a sensitivity analysis that indicates how large the correlation between error terms would have to be for a mediation effect to disappear.

6.2.3 Ergodicity and Within- vs. Between-Subjects Effects

The CLPN model, like any model based on group-level data, relies on the assumption that participants all come from a single population that is described by a single set of dynamic processes. To the extent that the data contain subgroups that are described by different causal patterns, observed effects will reflect a mixture of processes that may describe no actual individuals (Molenaar, 2004). Put another way, the CLPN model relies on between-subjects covariance to infer predictive processes. For these processes to hold at the level of individuals, all individuals must be assumed to be exchangeable draws from a homogeneous population.

A related issue is that CLPN models conflate between-subject and within-subject variance, which can lead to biased model results when the constructs under investigation contain stable individual differences (Berry & Willoughby, 2017; Hamaker et al., 2015). Hamaker et al. (2015) showed that when stable individual differences exist in the constructs being measured, cross-lagged paths conflate these individual differences (between-subjects effects) with the within-subjects effects that researchers are typically interested in. For example, if some people tend to have higher levels of Self-Esteem than others in general (i.e., if people who have higher than average Self-Esteem at $T1$ also have higher than average Self-Esteem at $T2$) and if this stable Self-Esteem trait is correlated with a stable Commitment to School trait, then the CLPM or CLPN will produce positive cross-lagged paths among these correlated variables, even if there is no causal effect within individuals.

Several models have been proposed to disentangle these effects (Grimm et al., 2017; Hamaker et al., 2015), all of which require at least 3 measurement occasions. Hamaker et al. (2015) random-intercept cross-lagged panel model (RI-CLPM) involves fitting a latent factor to the repeated observations,

which has the effect of removing the stable individual differences in levels of the trait. The latent factor accounts for between-subjects effects, leaving the within-subjects effects to be revealed in the auto-regressive and cross-lagged paths. Extending the RI-CLPM model to the CLPN would involve fitting a latent factor to each variable in the network simultaneously to modeling the cross-time paths. The panel VAR method that we examined in our simulation is a version of the RI-CLPM, but because it comes from the intensive longitudinal data tradition, it assumes a stationary longitudinal process.

6.2.4 Measurement Error

A common criticism of network models is that, unlike common factor models that assume each observed item is an unreliable indicator of an error-free latent variable, by modeling effects at the level of individual observed items, networks do not correct for measurement error. Proponents of network models tend to consider this as a feature, rather than a bug: Unlike latent variable models, which equate all unique item variance (reliable and unreliable) with measurement error, network models allow for the reliable unique variance of each item to be a central component of the causal system. As such, networks use the entire variance of each variable to model a construct, rather than just the part that is shared with all other items. However, it is undeniably true that random measurement error also exists, and to the extent that observed scores are unreliable, estimated regression coefficients will be biased. One way to improve this situation would be to include multiple indicators (e.g., self- and partner-report) of each behavior, attitude, emotion, or symptom of interest. Each set of items would then be collapsed (i.e., averaged or summed) into a single score before conducting the CLPN analysis. Such an approach would be expected to result in improved reliability of the observed scores while maintaining the network approach's focus on single behaviors, emotions, attitudes, and symptoms as possessing causal power.

7. CONCLUSION

Many researchers have access to longitudinal panel data with two or a few fixed measurement occasions on a few key constructs of interest. To address questions of how these constructs affect each other across the time lag of the study, researchers may consider using the CLPN in addition to, or instead of, the traditional CLPM. The CLPN relies on the insight that many psychological constructs do not function as a single causal entity but, instead, their components may have unique causal force. As such, the effects of one construct on another may be due to particular components of each construct. The CLPN can illuminate these specific relations to reveal nuanced and specific effects, and, as such, generate specific, testable causal hypotheses.

REFERENCES

- Aalbers, G., McNally, R. J., Heeren, A., Wit, S. de, & Fried, E. I. (2019). Social media and depression symptoms: A network perspective. *Journal of Experimental Psychology: General*, *148*(8), 1454–1462. https://doi.org/10.1037/xge0000528
- Abplanalp, S. J., Braff, D. L., Light, G. A., Nuechterlein, K. H., Green, M. F., Gur, R. C., Gur, R. E., Stone, W. S., Greenwood, T. A., Lazzeroni, L. C., et al. (2022). Understanding connections and boundaries between positive symptoms, negative symptoms, and role functioning among individuals with schizophrenia: A network psychometric approach. *JAMA psychiatry*, *79*(10), 1014–1022. https://doi.org/10.1001/jamapsychiatry.2022.2386
- Armour, C., Greene, T., Contractor, A. A., Weiss, N., Dixon-Gordon, K., & Ross, J. (2020). Posttraumatic stress disorder symptoms and reckless behaviors: A network analysis approach. *Journal of Traumatic Stress*, *33*(1), 29–40. https://doi.org/10.1002/jts.22487
- Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Physical Review Letters*, *103*(23), 238701. https://doi.org/10.1103/PhysRevLett.103.238701
- Beck, E. D., & Jackson, J. J. (2020). Consistency and change in idiographic personality: A longitudinal esm network study. *Journal of Personality and Social Psychology*, *118*(5), 1080–1100. https://doi.org/10.1037/pspp0000249
- Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*(1), 78–117. https://doi.org/10.1177/0049124187016001004
- Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child development*, *88*(4), 1186–1206. https://doi.org/10.1111/cdev.12660
- Bien, J., & Tibshirani, R. J. (2011). Sparse estimation of a covariance matrix. *Biometrika*, *98*(4), 807–820. https://doi.org/10.1093/biomet/asr054
- Borsboom, D. (2008). Latent variable theory. *Measurement*, *6*(1-2), 25–53. https://doi.org/10.1080/15366360802035497
- Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608
- Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., Borkulo, C. D. van, Bork, R. van, & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Prmers*, *1*(1). https://doi.org/10.1038/s43586-021-00055-w
- Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PloS one*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism?(don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. https://doi.org/10.1037/a0018933
- Cervin, M., Perrin, S., Olsson, E., Aspvall, K., Geller, D. A., Wilhelm, S., McGuire, J., Lázaro, L., Martínez-González, A. E., Barcaccia, B., et al. (2020). The centrality of doubting and checking in the network structure of obsessive-compulsive symptom dimensions in youth. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(7), 880–889. https://doi.org/10.1016/j.jaac.2019.06.018
- Christensen, A. P., Garrido, L. E., & Golino, H. (2023). Unique variable analysis: A network psychometrics method to detect local dependence. *Multivariate Behavioral Research*, *58*(6), 1165–1182. https://doi.org/10.1080/00273171.2023.2194606
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. https://doi.org/10.1037/0021-843X.112.4.558
- Conger, R. D., Lasley, P., Lorenz, F. O., Simons, R., Whitbeck, L. B., Elder, G. H., Jr, & Norem, R. (2011, November). Iowa youth and families project, 1989-1992. https://doi.org/10.3886/icpsr26721
- Conlin, W. E., Hoffman, M., Steinley, D., & Sher, K. J. (2022). Cross-sectional and longitudinal aud symptom networks: They tell different stories. *Addictive Behaviors*, *131*, 107333. https://doi.org/10.1016/j.addbeh.2022.107333
- Costantini, G., Saraulli, D., & Perugini, M. (2020). Uncovering the motivational core of traits: The case of conscientiousness. *European Journal of Personality*, *34*(6), 1073–1094. https://doi.org/10.1002/per.2237
- Cramer, A. O. J., Borkulo, C. D. van, Giltay, E. J., Maas, H. L. J. van der, Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS One*, *11*(12), e0167490. https://doi.org/10.1371/journal.pone.0167490
- Cramer, A. O. J., Van Der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*(4), 414–431. https://doi.org/10.1002/per.1866
- Cramer, A. O. J., Waldorp, L. J., Maas, H. L. J. van der, & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–50, discussion 150–93. https://doi.org/10.1017/S0140525X09991567
- Deboeck, P. R., & Preacher, K. J. (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling*, *23*(1), 61–75. https://doi.org/10.1080/10705511.2014.973960
- Deserno, M. K., Borsboom, D., Begeer, S., & Geurts, H. M. (2017). Multicausal systems ask for multicausal approaches: A network perspective on subjective well-being in individuals with autism spectrum disorder. *Autism*, *21*(8), 960–971. https://doi.org/10.1177/1362361316660309

- Epskamp, S. (2020a). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3
- Epskamp, S. (2020b). Psychonetrics: Structural equation modeling and confirmatory network analysis. R Package Version 0.13. https://doi.org/10.32614/cran.package.psychonetrics
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4). https://doi.org/10.18637/jss.v048.i04
- Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167
- Epskamp, S., Hoekstra, R. H., Burger, J., & Waldorp, L. J. (2022). Longitudinal design choices: Relating data to analysis. In *Network psychometrics with r* (pp. 157–168). Routledge.
- Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The gaussian graphical model in cross-sectional and time-series data. *Multivariate behavioral research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823
- Fan, J., Liao, Y., & Liu, H. (2016). An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal*, *19*(1), C1–C32. https://doi.org/10.1111/ectj.12061
- Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, *47*(6), 1231–1236. https://doi.org/10.1016/j.jesp.2011.05.007
- Finkel, S. E. (2008). Linear panel analysis. *Handbook of longitudinal research: Design, measurement, and analysis*, 475–504.
- Freichel, R., Pfirrmann, J., Jong, P. J. de, Cousijn, J., Franken, I. H., Oldehinkel, A. J., Veer, I. M., & Wiers, R. W. (2024). Executive functioning, internalizing and externalizing symptoms: Understanding developmental dynamics through panel network approaches. *JAACAP Open*, *2*(1), 66–77. https://doi.org/10.1016/j.jaacop.2023.11.001
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*(1), 1. https://doi.org/10.18637/jss.v033.i01
- Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and experience*, *21*(11), 1129–1164. https://doi.org/10.1002/spe.4380211102
- Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026
- Godfrey-Smith, P. (2010, January). *Causal pluralism* (H. Beebee, C. Hitchcock, & P. Menzies, Eds.). Oxford University Press.
- Gollub, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*, 80–92. https://doi.org/10.2307/1130293
- Grimm, K., O'Rourke, H., & Helm, J. (2017). Return of the bivariate growth model: Separating within- and between-person effects in

longitudinal panel data. *Paper presented at the Meeting of the Society for Research in Child Development*.

- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889
- Haslbeck, J. M. B., & Waldorp, L. J. (2020). Mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software*, *93*(8), 1–46. https://doi.org/10.18637/jss.v093.i08
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological methods*, *15*(4), 309. https://doi.org/10.1037/a0020761
- Isvoranu, A.-M., Borkulo, C. D. van, Boyette, L.-L., Wigman, J. T., Vinkers, C. H., Borsboom, D., & Investigators, G. (2017). A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms. *Schizophrenia bulletin*, *43*(1), 187–196. https://doi.org/10.1093/schbul/sbw055
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the n: Q hypothesis. *Structural equation modeling*, *10*(1), 128–141. https://doi.org/10.1207/S15328007SEM1001_6
- Joe, H. (2006). Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis*, *97*(10), 2177–2189. https://doi.org/10.1016/j.jmva.2005.05.010
- Johnson, M. D., Galambos, N. L., Finn, C., Neyer, F. J., & Horne, R. M. (2017). Pathways between self-esteem and depression in couples. *Developmental psychology*, *53*(4), 787. https://doi.org/10.1037/dev0000276
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*(5), 602–619. https://doi.org/10.1177/0193841X8100500502
- Kossakowski, J. J., Epskamp, S., Kieffer, J. M., Borkulo, C. D. van, Rhemtulla, M., & Borsboom, D. (2016). The application of a network approach to Health-Related quality of life (HRQoL): Introducing a new method for assessing HRQoL in healthy adults and cancer patients. *Quality of Life Research*, *25*(4), 781–792. https://doi.org/10.1007/s11136-015-1127-z
- Kuismin, M., & Sillanpää, M. J. (2016). Use of wishart prior and simple extensions for sparse precision matrix estimation. *PloS one*, *11*(2), e0148171. https://doi.org/10.1371/journal.pone.0148171
- Little, T. D. (2013). *The oxford handbook of quantitative methods, volume 1: Foundations*. Oxford University Press.
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, *31*(4), 357–365. https://doi.org/10.1177/0165025407077757
- Lucas, R. E. (2023). Why the cross-lagged panel model is almost never the right choice. *Advances in Methods and Practices in Psychological Science*, *6*(1), 25152459231158378. https://doi.org/10.1177/25152459231158378
- Mackinnon, S., Curtis, R., & O'Connor, R. (2022). A tutorial in longitudinal measurement invariance and cross-lagged panel models using lavaan. *Meta-Psychology*, *6*. https://doi.org/10.15626/MP.2020.2595

- Mayer, L. S. (1986). On cross-lagged panel models with serially correlated errors. *Journal of Business & Economic Statistics*, *4*(3), 347–357. https://doi.org/10.1080/07350015.1986.10509531
- McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems. *Clinical Psychological Science*, *3*(6), 836–849. https://doi.org/10.1177/2167702614553230
- McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1
- Pearl, J. (2012). The causal mediation formula—a guide to the assessment of pathways and mechanisms. *Prevention science*, *13*, 426–436. https://doi.org/10.1007/s11121-011-0270-1
- Qiu, W., & Joe., H. (2023). *Clustergeneration: Random cluster generation (with specified degree of separation)* [R package version 1.3.8]. https://CRAN.R-project.org/package=clusterGeneration
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/
- Revelle, W. (2024). The seductive beauty of latent variable models: Or why i don't believe in the easter bunny. *Personality and Individual Differences*, *221*, 112552. https://doi.org/10.1016/j.paid.2024.112552
- Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of abnormal psychology*, *125*(6), 747. https://doi.org/10.1037/abn0000181
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling. *Journal of statistical software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02
- Rutten, R. J., Broekman, T. G., Schippers, G. M., & Schellekens, A. F. (2021). Symptom networks in patients with substance use disorders. *Drug and Alcohol Dependence*, *229*, 109080. https://doi.org/10.1016/j.drugalcdep.2021.109080
- Savalei, V., Brace, J. C., & Fouladi, R. T. (2023). We need to change how we compute rmsea for nested model comparisons in structural equation modeling. *Psychological Methods*. https://doi.org/10.1037/met0000537
- Savalei, V., Brace, J. C., & Fouladi, R. T. (2024). We need to change how we compute RMSEA for nested model comparisons in structural equation modeling. *Psychological Methods*, *29*(3), 480–493. https://doi.org/10.1037/met0000537
- Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality mark-

ers for more accurate predictions. *European Journal of Personality*, *32*(3), 186–201. https://doi.org/10.1002/per.2147

- Sher, K. J., Wood, M. D., Wood, P. K., & Raskin, G. (1996). Alcohol outcome expectancies and alcohol use: A latent variable cross-lagged panel study. *Journal of abnormal psychology*, *105*(4), 561. https://doi.org/10.1037/0021-843X.105.4.561
- Stewart, R. D., Mõttus, R., Seeboth, A., Soto, C. J., & Johnson, W. (2022). The finer details? the predictability of life outcomes from big five domains, facets, and nuances. *Journal of personality*, *90*(2), 167–182. https://doi.org/10.1111/jopy.12660
- Tanaka, J. S. (1987). "how big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child development*, 134–146. https://doi.org//10.2307/1130296
- Taylor, S., Fong, A., & Asmundson, G. J. (2021). Predicting the severity of symptoms of the covid stress syndrome from personality traits: A prospective network analysis. *Frontiers in Psychology*, *12*, 632227. https://doi.org/10.3389/fpsyg.2021.632227
- Van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, *4*(1), 5918. https://doi.org/10.1038/srep05918
- van der Maas, H. L. J., Savi, A. O., Hofman, A., Kan, K.-J., & Marsman, M. (2019). The network approach to general intelligence. In D. J. McFarland (Ed.), *General and specific mental abilities* (pp. 108–131). Cambridge Scholars Publishing.
- van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological review*, *113*(4), 842. https://doi.org/10.1037/0033-295X.113.4.842
- Williams, D. R. (2021). Ggmnonreg: Non-regularized gaussian graphical models in r. *Journal of Open Source Software*, *6*(67), 3308. https://doi.org/10.21105/joss.03308
- Williams, D. R., & Mulder, J. (2020). Bggm: Bayesian gaussian graphical models in r. *The Journal of Open Source Software*, *5*(51), 2111. https://doi.org/10.21105/joss.02111
- Williamson, J. (2006). Causal pluralism versus epistemic causality. *Philosophica*, *77*(1). https://doi.org/10.21825/philosophica.82198
- Wulff, D. U., & Mata, R. (2025). Semantic embeddings reveal and address taxonomic incommensurability in psychological measurement. *Nature Human Behaviour*, *9*, 944–954. https://doi.org/10.1038/s41562-024-02089-y
- Wysocki, A. C., & Rhemtulla, M. (2021). On penalty parameter selection for estimating network models. *Multivariate Behavioral Research*, *56*(2), 288–302. https://doi.org/10.1080/00273171.2019.1672516
- Zavlis, O., Butter, S., Bennett, K., Hartman, T. K., Hyland, P., Mason, L., McBride, O., Murphy, J., Gibson-Miller, J., Levita, L., et al. (2022). How does the covid-19 pandemic impact on population mental health? a network analysis of covid influences on depression, anxiety and traumatic stress in the uk population. *Psychological Medicine*, *52*(16), 3825–3833. https://doi.org/10.1017/S0033291721000635