PATH AND DIRECTION DISCOVERY IN INDIVIDUAL DYNAMIC MODELS: A REGULARIZED HYBRID UNIFIED STRUCTURAL EQUATION MODELING WITH LATENT VARIABLE

Ai Ye

A Dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctors of Philosophy in the Department of Psychology and Neuroscience in the College of Arts and Sciences.

Chapel Hill
2022

Approved by:

Kenneth A. Bollen

Kathleen M. Gates

Daniel J. Bauer

Vladas Pipiras

Teague R. Henry

ABSTRACT

Ai Ye: Path and Direction Discovery in Individual Dynamic Models: A Regularized Hybrid Unified Structural Equation Modeling with Latent Variable
(Under the direction of Kenneth A. Bollen)

There recently has been growing interest in the study of psychological and neurological processes at an individual level. One goal in such endeavors is to construct person-specific dynamic assessments using time series techniques such as Vector Autoregressive (VAR) models. Within the psychometric field, researchers have developed psychometric modeling frameworks to estimate different variants of VAR models. These modeling frameworks estimate the dynamic relations (e.g., temporal and contemporaneous) unpacked in a multivariate time series data. However, two problems exist with current VAR specifications: 1) VAR models are restricted in that contemporaneous relations are typically modeled either as undirected relations among residuals or directed relations among observed variables, but not both; 2) current estimation frameworks are limited by the reliance on stepwise model building procedures. This study adopts a new modeling approach, i.e., LASSO regularized hybrid unified Structural Equation Model (SEM), for a global search and estimation of a more flexible VAR representation. The present study to our knowledge is the first application of the recently developed regularized SEM technique to the estimation of a type of time series SEM, which points to a promising future for statistical learning in psychometric models.

To my mom and dad, who inspire and love me every day,
who always encourage me to pursue every educational endeavor
and teach me to be humble and kind to all.

ACKNOWLEDGEMENTS

They say it takes a village to raise a child, to this day it takes this child more than three decades to add the three-letter title to her name.

This dissertation could only exist because I made it happen, but I can only make it through with the support, inspiration, encouragement, and company of so many others. With a humble heart, here I would like to express my deepest gratitude to those of you that played a crucial part in this journey. There is only so little space, but so much to thank for.

They say it takes a village to raise a child, the academic village that raised this child as a quantitative psychologist is called the 'L. L. Thurstone Psychometric Lab'. First and foremost, I am deeply indebted to my advisor, Prof. Dr. Kenneth Bollen, whose name and work in quantitative sociology and psychology have influenced and changed the path of my career, and perhaps for many others as well. His 1989 book "Structural Equations with Latent Variables" is still considered the "golden standard" textbook for one of the most fundamental modeling frameworks in psychometrics. I am fortunate to have Ken's guidance and advice that carried me through all the stages of my doctoral career and in writing this dissertation. Besides his academic expertise, Ken is the most responsive and supportive mentor and the most persistent and humble scholar that I could ever imagine. This work and our other collaborative ones would not possibly be accomplished without Ken's effort in providing his guidance and comments in the most timely and approachable fashion. Next, I thank my secondary advisor, Prof. Dr. Katie Gates, who had added profound ideas and instructions to this work. Thanks to our crossed interest in intraindividual research and individual dynamic models using time series data that brought us to work together. Katie is the most collegiate, intelligent, and humorous person to work with. Further, I would like to extend my sincere thanks to all of the other committee members, Profs. Drs. Teague Henry, Dan Bauer, and Vladas Pipiras, for their continuing insightful instructions and comments, to make this a much better work. Special thanks to Teague, an alumnus and a rising star in the field of dynamics and brain network research, for his keen interest in peer mentoring me at an important stage of my doctoral program and also for his prompt inspiration, sharp ideas, and suggestions for me to transform my course project into a publishable paper. I would also like to express my deep gratitude to a retired faculty at the lab, Dr. David Thissen. Even though we have not worked extensively together, I thank Dave for his unconditional spiritual support for me in facing the unique challenges I encountered as a foreign, international student. An understanding from him is one of the reasons I could persist during the most difficult times.

Additionally, it is a genuine pleasure to express my appreciation to all my fellow alumni, peers, cohort colleagues, and friends from my village, the Thurstone lab. Thanks to alumni Drs. Stephanie Lane, Zack Fisher, Michael Giordano, Noah Greifer, Cara Arizmendi, Alexis Georgeson, and Will Belzek, for providing the best examples of how far and wide a psychometric graduate could reach, and to current students Chris Urban, Marco Chen, Chris Strauss, Sandra Williams, Lan Luo, Saifa Pirani, and many junior and incoming students that I have not yet got much chance to interact with particularly during the pandemic time, for being my peers and friends. My thanks also go to alumni and peers from other graduate programs or Departments, e.g., Drs. Yun Chen, Van Liceralde, and Adam Lily, to you and all others above, I thank you for so many small and big moments of sharing ideas and offering suggestions in the hallway of Davie Hall, over a lunch break, or through a campus walk. We all know graduate school is not easy, thank you for being a comrade in this battle!

They say it takes a village to raise a child, one of the key factors to secure growth is professional and financial support from the bigger society. During my Ph.D. years, I was privileged to have interned at two of the most prestigious companies in the triangle area, SAS Institute and RTI International. These two places helped me improve professionally as a psychometrician in the industrial world. Foremost, words cannot fully express my gratitude to my mentor at RTI, Dr. Antonio Morgan-Lopez. Even though the pandemic has limited our entire interaction to be virtual, that did not affect him to be the warmest, most thoughtful, and most caring mentor that I have ever known in this academic field. The virtual interaction during this special time is a source of inspiration and emotional support for me as well. My sincere appreciation also goes to my other mentors and colleagues at both internship institutes, Drs. Ty Ridenour, Stephen Tueller, Lissette Saavedra, Lesia Ruglass (from RTI), and Drs. Jan Chvosta, Xilong Chen, Shiyin Wu, Yin Zhu, Wen Wei, Ye Liu, Laura Castro-Schilo (from SAS), for coaching me and sharing your knowledge, as well as to those with whom I had worked in my other Research Assistant positions, including the

Epidemiology research groups under the leadership of Drs. Yun Li, Alison Aiello, and Chantel Martin. Thank you all for giving me the opportunity for the wonderful research work and training experience.

I would like to take this opportunity to also thank the two other academic institutes that brought me to the current place, the University of California, Los Angeles (UCLA) for my Master's, and the University of Delaware (UD) for two years of research experience. I thank Prof. Dr. Li Cai from UCLA, my first graduate adviser, for setting an example as a pioneering scholar, professor, and mentor with the highest standard in the psychometric world. I thank Prof. Dr. Nancy Jordan and her research team (during 2014-2016) from UD, Drs. Luke Rinne, Nicole Hansen, Ilyse Resnick, and Jessica Rodrigues, as well as the UD Residence Life Housing department, Drs. Tweedys, Kathleen Kerr, Joshua Holmes, Christina King, and Jimmy Howards, for the financial support and the productive collaborative work together. I thank the mentorship from Profs. Drs. Henry May and Elizabeth Farley-Ripple at UD, for your encouragement and hands-on support for me to pursue the doctoral at my dream place and become the type of researcher I wanted to be.

To the broader psychometric society, I would like to thank scholars whose work or profession had inspired me to work on topics that lead to this dissertation, such as Profs. Drs. Hua-hua Chang, Jeff Haring, Sacha Epskamp, Ross Jacobucci, Po-Hsien Huang, and many more. Further gratitude extended to many others that I met during academic conferences, as well as to anonymous reviewers of my papers. The academic interactions with you through conference talks, emails, comments, etc. not only largely improved my work but also provided numerous inspirations for future research to prosperous the field.

They say it takes a village to raise a child, but no one ever tells you where it is and how to get there. In this long endeavor, I am blessed to have met and become friends with so many lovely people whose character and the company encouraged me to carry on. I thank the inspiration and encouragement from my special friend and life coach Dr. Jerry Bell, and my loving and caring godparents Linda and Charles Wooters. Many thanks go to my fellow UCLA cohort and friends Xue Gong, Yi Wei, Shanshan Jiang, Ya Gao, Xiaoxiao Tian, Simon, and Doris, Enping, and Miyeon Song, to conference pals, Drs. Xue Zhang, Manqian Liao, Zi Chen, Yujia Li, and Wei Li, to friends I met from all walks of life, Yijun, Jason, Daniel, Euphy, Christina, Sunhwa, Wenjin, Peihai, Rose, Begona, Denny, Kat, Brandon, Dania, Arne, Laura, Elyse, Jo K, as well as to my tango teachers

John and Rachel, my Ashtanga mentor teacher Lyndsey, my Pole dance instructor Janine. I thank you all for bringing me numerous healthy activities, happiness, and lovely moments throughout this most challenging period, I know I would not be able to survive without them.

They say it takes a village to raise a child, but it takes a lot of solid, stable, loving families to create a village. I am grateful to my lovely family members and childhood friends from my hometown China, Xinchen Jiang, Wenqing Wang, Lian Sun, Han Yang, ShuXiu Dong, and Tingting Wang, for building up the village to let me grow up into the person I am. Most importantly, I am blessed to be the daughter of my mum and dad and to receive their unconditional love and care. I thank them for being the most patient and open-minded parents to let me take the time to figure out what life I would like to live, for always encouraging me to pursue whatever makes me feel confident and happy without worrying about financial burdens, and for the ultimate unconditional support for whatever choice I made. I dedicate all my achievements to my parents.

They say it takes a village to raise a child, and the grownup finds a lifetime partner to join and broaden the village. Finally, I want to thank Dr. Christian Macaro, my friend, my teacher (of life and ski), and my Mr. Right. Thank you for being the motivation to finish my Ph.D. as quickly as I could and at the same time move to a new continent in the middle of the pandemic. Before I have time to digest how I would adapt to European languages and cultures, I found myself already happy with my new families and friends from your hometown and our new routines as global citizens in this magical Bavarian city. You are a true surprise at this finishing line, and a gift of my life to embark on a brand new chapter together. First, you opened my eyes by how you savor everyday moments and massage your (Italian) food in a way that never heard of or seen by me before. Then, you prompted me to rethink and understand what happiness, friendship, and relationship mean to me. Together, we are learning to be the type of romantic partner that encourages and accompanies each other to be the happiest version of ourselves.

They say it takes a village to raise a child, thank you all for being part of my village!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Unifying Themes

With the development of technology to collect intensive longitudinal or time series data (TSD) in a fast and economical pace, recent decades have been witnessing a surge of psychological and neurological research at an individual level. Studies focus on person-specific dynamic assessment emphasize individual differences and the understanding of individual characteristics and development. Along with the shift of research interest is the need for advancing the statistical methods to facilitate person-specific dynamic assessments using TSD. Within the psychometric field, researchers have developed psychometric modeling frameworks to fit traditional time series models, such as the Vector Autoregressive (VAR) models, that have deep roots in statistics and econometrics. These modeling frameworks estimate the dynamic relations (e.g., temporal and contemporaneous) unpacked in a multivariate TSD. For instance, two representative approaches are the unified Structural Equation Model (uSEM), as a time series extension of the SEM, and the graphical Vector Autoregression (gVAR) model, as a time series extension of the network psychometric model. These approaches differ in the variant of VAR representation used in the model (often restricted in one way or another), as well as the estimation framework to select and identify the optimal model. However, there remains much work to do. My dissertation aims to evaluate some of the limitations and to propose methodological improvements. Specifically, I design three studies to tackle each of the following issues and research gaps using simulation studies and empirical example illustrations. Below is a preview of each chapter explaining how I plan to accomplish this goal.

First, despite the vast amount of methodological work, few studies discussed the differences in the interpretations and causal implications carried from the adoption of different VAR variants in these psychometric approaches, even though a unified (often unspoken) goal is to identify the underlying causal mechanisms or key variables of the multivariate dynamic system. One particularly salient aspect regards the modeling of contemporaneous (or instantaneous) relations. Approaches typically fall under two common categories of limited VAR representations: one form is as undirected relations among the errors after conditioning on lagged effects (such as in the gVAR model); the other is as directed relations among observed variables (as is done using the uSEM framework). The interpretations and consequences of each assumption behind these practices have not yet been evaluated. The first paper of my dissertation aims to explicate the interpretative caveats of choosing the options for modeling contemporaneous relations among variables. To do so, I provide small simulation and empirical working examples throughout the paper to demonstrate specific concepts by fitting time series models with both directed and undirected contemporaneous networks. I demonstrate how inferences are altered by the modeling choice that can inform both psychometric and neuroscience research. This focused discussion will fill a timely and vital gap given the increasing application of these dynamic models.

Second, most of the current approaches are limited in at least two ways: 1) VAR models are restricted in that contemporaneous relations are represented either as only undirected or as only directed, but not both. Despite the fact that they carry different casual interpretations that could coexist in a dynamic system. 2) Current estimation frameworks are heavily reliant on pseudo-ML based stepwise model building procedures, and little research has compared the properties of these traditional methods of model selection with those recently adapted from statistical learning (e.g., regularized SEM; Jacobucci et al., 2016). To fill in these gaps, the second paper of my dissertation extends the current practice to the proposal of a novel estimation framework for a more flexible VAR representation. In proposing the new modeling approach, I first extend one of the current frameworks to a hybrid representation, which incorporates both the undirected and directed contemporaneous effects that to be identified by a data-driven algorithm. Second, for a more efficient automatic model search tool, I replace the stepwise fashion with a regularization method implemented within SEM. This new framework allows for a global and continuous search for the optimal sparse model. A simulation study is conducted to investigate whether the proposed approach is superior to alternatives with respect to accurately recovering the presence and directionality of hybrid relations and reliably removing false relations when the data are generated to have two types of contemporaneous relations.

Last but not least, there has been an increasing call to the applications of these methods to study high dimensional TSD (i.e., TSD with large number of variables), with the capacity to account for measurement error of time series variables. When multiple time series indicators are used to measure the same unobserved latent construct, or that theories support the hypothesis of some potential latent variables underlying the manifest variables, factor time series variables could be formed in order to adjust for measurement error and to reduce the dimension of the observed variables. The combination of a factor model with a VAR model leads to the well-known dynamic factor model (DFM), in which dynamic relations are derived within factor series or among factor and observed time series, or both. However, the above limitations in the VAR representation and the model search method also apply to DFM, that is, the VAR model contains either directed or undirected contemporaneous relations but not both, and that regularization has not been implemented or evaluated under DFM. Therefore, the third paper of my dissertation serves to advance and investigate the DFM with the proposed hybrid VAR representations and regularization as the model search method. For the evaluation, I perform a simulation study to compare these methods with respect to 1) the sensitivity of finding the true dynamic relations in the structure model; 2) the specificity of excluding the false relations; and 3) the robustness of parameter estimates in the structural model in response to model structural misspecifications. The current work aims to offer methodological and applied researchers the best practice to conduct model selection and estimation in person-centered dynamic assessments.

CHAPTER 2

STUDY I: On the Interpretation of Directed and Undirected Contemporaneous Relations in Dynamic Psychometric Networks

2.1 Background

The recently introduced network psychometric approaches graphically depict statistical relations among observed variables. These methods initially provided as an alternative to latent variable modeling where observed variables are related to each other directly rather than via a common “unobserved" or “latent" factors (Borsboom & Cramer, 2013; Fried, 2020; Schmittmann et al., 2013). With many attractive properties, network models have gained popularity in psychological science particularly in the study of psychopathological symptoms, mental disorders, and the structure of personality (e.g., Borsboom, 2017; Boschloo et al., 2015; Fried et al., 2017; Rhemtulla et al., 2016). Network representations and metrics are increasingly used to depict idiographic processes using time series data (TSD; Epskamp, Waldorp, et al., 2018; Fisher, 2015). TSD refers to data with many repeated measurements over time, ideally over 60 per person, and has been becoming increasingly available in psychological and neurological science such as ecological momentary assessment (EMA) or neurological measurements obtained from functional neuroimaging (i.e., fMRI). For example, graphical vector autoregression (gVAR) models person-specific intraindividual dynamics (often in EMA studies) and displays results as networks (Epskamp, Waldorp, et al., 2018; Wild et al., 2010). Similarly, the group iterative multiple model estimation approach (GIMME; Gates & Molenaar, 2012) depicts results for multiple subjects from unified Structural Equation Models (uSEM; Gates et al., 2010; Kim et al., 2007), as sparse networks. GIMME has predominantly been used in neuroscience applications for identifying key brain regions that relate to other spatially disparate brain regions across time. In both EMA and brain imaging applications, identifying causal mechanisms or key variables is often the goal (Fisher, 2015), as researchers often wish to identify influential nodes (e.g., emotions, brain regions) that might be targeted to change the system (Fisher & Boswell, 2016).

While much methodological work has focused on advancing these methods, little has been done to evaluate the differing inferences made from the network approaches. This is surprising given 1) the amount of applications that used networks to inform implications, and 2) the topic of causal inferences has been discussed extensively in other literature, such as work in economics and directed acyclic graphs (Pearl, 2009; Ramsey et al., 2011). The two approaches share roots in a common statistical ground within time series literature surrounding the VAR model family. Still, they have differences that must be noted when making inferences. A key aspect of interpreting results is considering the assumptions implicitly made when conducting the analysis. Researchers making inferences from network approaches in both psychometric and neuroscience applications do not always consider the underlying assumptions for the models.

One particularly salient aspect regards the modeling of contemporaneous (or instantaneous) relations. Approaches generally fall under two frameworks for capturing contemporaneous relations: (1) estimate the undirected relations among the residuals after conditioning on lagged effects, as in gVAR or (2) model the directed relations among observed variables with the assumption that nodes of interest are all influenced by the system, as is done using uSEM framework implemented in GIMME. However, the choice of which modeling approach is often difficult, due to a lack of methodological guidance on what interpretations are entailed by the two contrasting practices. Perhaps one reason is that the use of time series analysis in psychology presents a relatively new set of techniques that does not have the same history of methodological evaluation as does the more traditional SEM framework within the field of psychology. Indeed, the existing methodological work has focused on advancing the estimation and evaluation methods or establishing the replicability and stability of these network models (e.g., Costantini et al., 2019; Epskamp, Borsboom, et al., 2018; Forbes et al., 2017). Little has been done to evaluate the implications regarding the causality inference. On the contrary, causality have been discussed extensively in the SEM literature, such as path analysis, classic SEM. One of the important concepts is the directed acyclic graph (DAG) theory. Additionally, dynamic causality concepts are deeply rooted in economics literature. We seek to explicate how insights from the causal inference literature can be applied to the interpretation of both directed and undirected network models.

Goal of Study

The present paper explicates the interpretative caveats of both options for modeling contemporaneous relations among variables. To do so, we provide examples throughout the paper to demonstrate specific concepts using simulated data as well as empirical data gathered from EMA and fMRI scans. By fitting time series models with both directed and undirected contemporaneous networks, we can better demonstrate how inferences are altered by the modeling choice that can inform both psychometric and neuroscience. This focused discussion will fill a timely and vital gap given the increasing application of these network models.

We want to stress that the current work is different from the majority of relevant methodological evaluations which often tend to relate the varied psychometric network modeling (PNM) and latent variable modeling techniques *against* each other, or to arrive at some endorsement on the choice of one modeling framework over the other (e.g., Hallquist et al., 2019). We found that little work has been done in comparing the causal implications in alignment with the assumptions of each modeling approach *as they are currently used*, despite the fact that such methodological evaluation is particularly important to guide the choice of the two approaches that often face applied researchers who analyze time series data and use model interpretation in their policy making. The current study aims to bridge this crucial gap.

The remainder of this article is structured as follows: First, we provide background for the two approaches for modeling time series data, (i.e., gVAR and uSEM/GIMME), their definitions and assumptions. Particularly, we point out the differences in the treatment of the contemporaneous relations. Second, we review important concepts and theories developed in Psychometric and Econometric literatures, relating to casual inference interpretation. Third, we provide two empirical examples to demonstrate the similarities and differences in the interpretation of results obtained from fitting a gVAR versus a uSEM to the same datasets, representing a typical EMA and a fMRI application, respectively. We focus on the model interpretation and causal implication using the causality concepts to provide examples of how the interpretations may be different based on the model used - even if the patterns of relations contain similarities. Finally, we provide recommendations for using these models.

2.2 Two Common Approaches: graphical VAR and unified SEM

Despite often being used under different contexts, gVAR and uSEM shared a common foundation of a multivariate time series technique known as the Vector Autoregressive model (VAR;

Hamilton, 1994; Lütkepohl, 2005; Shumway & Stoffer, 2017). VAR was developed to assess dependencies between multivariate repeated measures. When used in psychology, it typically assess these relations within a person across time. It is thus suitable to construct person-specific dynamic assessments with methods based on VAR. We describe here the basis of VAR from which the modeling frameworks of gVAR and uSEM expand.

2.2.1 The VAR Model

Let $Y_t = [y_{1t}, y_{2t}, ..., y_{pt}]'$ be a vector of a p-variate time series,with $p > 1$, at a given time point t, with $t = 1, ..., T$. Suppose Y_t can be represented by a stationary linear time series (i.e., having constant means and covariance functions), defined by the following VAR model with lag order-l, VAR(l):

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_l Y_{t-l} + \varepsilon_t,\ \varepsilon_t \sim N_p(0, \Theta) \tag{2.1}$$

where the sequence of $p \times p$ dimensioned Φ_k matrices contain the lagged coefficient estimates at order $k, k = 1, ..., l$, with diagonal elements being the AR(k) regression coefficients and the off-diagonal cross-lagged coefficients. In this representation the expected value of all dimensions of Y_t is assumed to be zero and time invariant: $\mu_t = 0$, $\forall$t $\in 1 \ldots T$, and there is no auto or cross-lagged dependency in the error term ε_t, i.e., the temporal covariance at lag u, $\Theta_u = cov[\varepsilon_t, \varepsilon^{'}_{t-u}] = 0\ \ \forall u \in 1 \ldots \infty$. This zero-mean p-variate ε_t may be correlated contemporaneously.

The simplest and mostly commonly used form is the first-order VAR model where l in Equation 2.1 equals 1 and thus reduces to $Y_t = \Phi_1 Y_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, \Theta)$. This reduced form is referred to as lag-1 VAR or VAR(1), which estimates temporal relations of only consecutive measurements. VAR(1) model assumes a “memoryless" (i.e. Markov) process in which all previous information useful to predict the current values of Y_t is in Y_{t-1}. All the gVAR and uSEM models discussed in the present paper are based on VAR(1) processes. The generalization to higher-order VAR models should be intuitive but is beyond the current scope.

2.2.2 The gVAR Model

The gVAR is VAR estimated by the framework of Gaussian Graphical Modeling (GGM; Abegaz & Wit, 2013; Eichler, 2005; Wild et al., 2010). Two levels of relations are estimated by gVAR: on the lag-1 level a weighted, *directed* temporal network is estimated on the temporal ordering dependency exactly as in VAR; at the lag-0 level, a weighted, *undirected* contemporaneous network is drawn on the partial correlations among residuals of the observed variables, following the GGM fashion. GGM is a general network approach that attempts to estimate a sparse set of partial correlations with testable solutions among (usually high-dimensional) variables (or “nodes" in a network) that assumes to follow a multivariate normal distribution (Lauritzen, 1996). The two sets of results produced by gVAR are also termed partial directed correlation (PDC) network and partial contemporaneous correlation (PCC) network, respectively. A sparse pattern of relations among variables at each lag is achieved by applying a regularization method to the matrices containing estimation parameters. The practice of *regularization* or *penalization* refers to a global search algorithm for model selection, in which a penalty term is added to the cost function to control the size and the number of nonzero parameters (Tibshirani, 1996). Estimation using LASSO regularization introduces sparsity and arrives at selection of relations among variables by imposing a penalty term on designated parameters that undergo a continuous shrinking towards zero all at once. This way, only a subset bivariate relations are represented as nonzero edges between nodes while a lack of a relation among contemporaneous relations suggests independence after conditioning on all other variables in the system.

Traditionally, obtaining a sparse PCC and PDC results involves three major steps, after the specification of a VAR model derived by Equation 2.1 (Wild et al., 2010). First, a precision matrix κ is computed by taking the inverse of covariance matrix Θ in Equation 2.1. Next, partial correlations in PCC nework are obtained via a standardization on the elements of precision matrix κ derived by $PCC(Y_{i,t}, Y_{j,t}) = Cov(Y_{i,t}, Y_{j,t} | y_{-(i,j)}) = -\frac{\kappa_{i,j}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}}$, where $\kappa_{i,j}$ denotes an element in , $y_{-(i,j)}$ denotes all variables subtracting variable i and j, hence PCC represents the correlation between i and j at time t, after partialing out the linear effects of all other variables. The temporal relations PDC can be obtained via a similar standardization on ϕ elements, representing the linear lagged relationship between $y_{i,t}$ and $y_{j,t-1}$, net the linear effect of all other variables at $t-1$ that are included in the model. Finally, small partial correlations are first forced to zero by thresholding rules, the remaining parameters are then regularized and shrink towards zero (with some reaching exactly at zero).

The regularized estimates of gVAR has been implemented in the free open-source R package *graphicalVAR* (Epskamp, 2018). This package employs a hybrid variant of LASSO regularization algorithm, specifically the multivariate regression with the covariance estimation (MRCE) first proposed by Rothman et al. (2010), and later extended to time series models by Abegaz and Wit (2013), to jointly estimate the temporal and contemporaneous networks. The joint regularization method involves iteratively optimizing edges of the lag-1 PDC (i.e., the temporal network) by searching across a range of λ_1 using the cyclical-coordinate descent, and the lag-0 PCC (i.e., the contemporaneous network) by λ_2 using the graphic LASSO algorithm (Friedman et al., 2019; Friedman et al., 2008), until model converges with the optimal pair of (λ_1, λ_2) that gives the least mean square error loss indicated by the lowest extended BIC. Finally results are provided for the user both graphically via a network representation and the matrices of coefficient estimates.

We can see that the gVAR model can be considered as a time series extension of GGM by constructing an additional temporal network on temporal dependencies. By combining the two techniques into one, gVAR inherits the features of both: like GGM or other general PNMs, it is powerful in handling high-dimensional data with an efficient estimation and model search algorithm to recover sparse networks of multivariate system (Epskamp, Waldorp, et al., 2018). Like VAR, it can flexibly model essential dynamic processes and can be easily extended to incorporate higher order counterparts or more complicated components such as moving average, heterogeneous error structures, etc. (Shumway & Stoffer, 2017).

2.2.3 The uSEM

Alternatively, TSD can also be modeled under an extended SEM framework fitted to Structural VAR model, the latter being a more general class of VAR models (Chen et al., 2011; Shapiro & Watson, 1988; Sims, 1981). In addition to identifying AR or cross-lagged effects as regression coefficients, SVAR also incorporates the structural paths amongst contemporaneous variables. Therefore, VAR(1) given by Equation 2.1 (with order $l = 1$) can be transformed to an SVAR(1) as demonstrated by Gates et al. (2010): $Y_t = AY_t + \Phi_1 Y_{t-1} + \zeta_t, \zeta_t \sim N(0, \Psi)$, where the $p \times p$ matrix A contains regression coefficients for structural contemporaneous relations, or how variables relate to each other instantaneously. The presence and direction of the relation between two given variables at time t, are estimated by two coefficients a_{ij} and a_{ji}, representing the opposite directions. However, the fully saturated SVAR(1) model in above equation (i.e., with free parameters for all possible structural coefficients for observed variables and covariance coefficients for residuals) is not identifiable. For this reason, the induction of sparsity is required. In this type of SVAR model the residuals (innovations) ζ_t are assumed to be mutually independent white noise processes with all residual covariances in Ψ set to zero (i.e., Ψ becomes diagonal covariance matrix with the off-diagonal all 0's and the diagonal containing the residual variances).

Such structural characteristics makes SEM (Bollen, 1989) a suitable framework to fit the reduced SVAR(1). Previous studies have demonstrated that time series estimates obtained via SEM are similar to those estimated using more traditional approaches such as the Kalman Filter (Bringmann et al., 2017; Fisher et al., 2017) . Conventional SEM has been extended to the uSEM modeling to account for temporal ordering dependency among adjacent measures in addition to the norm structural relations among contemporaneous measures in time series data (Gates et al., 2011; Kim et al., 2007). An advantage of specifying SVAR in uSEM is that it solves the nonunique solution issue arisen from the Choelsky transformation involved in the conventional estimation of the SVAR(1) by enabling a stepwise procedure for obtaining a sparse set of relations (Molenaar, 2019).

To rewrite a SVAR(1) as uSEM using traditional SEM notation (Gates et al., 2017):

$$Y = BY + \zeta,\ \zeta \sim N(0, \Psi), \tag{2.2}$$

where $Y = [Y_{t-1}, Y_t]$. That is, the observations of variables are time-embedded by appending the data at $t-1$ to the data at t. The data are thus expanded to two consecutive time points $t-1$ and t, so time series vector of lagged variables and those of contemporaneous variables are stacked horizontally. This requires that the ζ error vector also be extended, as well as the Ψ matrix (see Gates et al., 2017, for details).

To arrive at an testable and interpretable uSEM model, an exploratory approach (as opposed to confirmatory) is often performed in a step-wise model selection fashion using Lagrange multiplier score tests or modification indices (Jöreskog & Sörbom, 1986). An automatic forward-selection model building algorithm was proposed by Gates and colleagues (2010), where model search at an individual-level begins with a null model, and the path with the highest significant modification index is added iteratively until the model arrives at an acceptable fit as indicated by standard fit indices used in SEM. To account for heterogeneous idiographic models across individuals, this individual-level modeling procedure was later extended to a group-level algorithm GIMME (Gates & Molenaar, 2012), in which shared information across individuals is extracted to constructed first the group-level model and then is used to assist the recovery of the individual-level model. Both the individual-level and the group-level model building procedures have been implemented in the R package *gimme* (Lane, Gates, Fisher, et al., 2019), where the user simply provides the raw time series data, and time-embedding is automatically done within the program. We utilize only the individual-level uSEM approach (i.e., the *indSEM* function within *gimme*) here to be consistent with the use of gVAR and ease in explanation of results.

2.2.4 Similarities and differences in gVAR and uSEM

In one sense, the gVAR is just another representation of relations that can also be captured by simultaneous equation models, or SEMs without latent variables, and both have the traditional VAR at their core. They both model lag-1 relations, whereby the data at the previous time point are included in the search space as potential predictors of the data points at a later time point, and both capture additional contemporaneous relations as well. Both are data-driven, estimated with their own automatic model build algorithm that arrives at an optimal sparse networks.

However, the two frameworks have been serving very different applications. This is largely due to the fact that they differ dramatically in the fundamental hypothesis as to how variables relate. Psychopathology network theory suggests that observed variables correlate because they interact with each other directly, a notion that is similar to that of the simultaneous equation models yet complements the latent variable model, which depicts observed variables as being consequences of some underlying common latent cause (Cramer et al., 2010). In this juxtaposition, it follows that the contemporaneous relations would be bidirectional in nature, in order to accommodate the possibility of being manifest variables representing the same underlying construct such as those in latent variable models. The directed contemporaneous relations seen in uSEM follows directly from the state of the science in fMRI modeling. Here, the data collected (based on blood oxygen levels) are known to be at a far slower temporal resolution than the phenomena of interest, the neuronal activity of the brain (Logothetis, 2008). Thus lagged relations are not always informative since the time from one data point in time to another might be too long to reveal relations. At the same time, the field is interested in arriving at causal models. Much work being done to arrive at reliable data-driven approaches that capture both the presence and direction of relations both lagged and contemporaneously (Ramsey et al., 2011; Smith, 2012; Smith et al., 2011).

Another difference between the two models pertains to statistical assumptions and model identification. A fully saturated directed contemporaneous model (specified as path model) in uSEM is underidentified meaning that no unique solution exists for the set of parameters under the given model constraints. In contrast, without any causality features, a fully saturated undirected contemporaneous model (specified as GGM) in gVAR is just identified (i.e., a unique solution can be reached for the set of parameters under the given model constraints with no additional degree of freedom). This is to say that, sparse exploratory patterns of relations are assumed in gVAR but required in uSEM. The models rely on the built-in automatic model building procedures, regardless of which method is adopted, to arrive at their optimal sparse patterns that carry corresponding model assumptions and optimization constraints. Another difference concerns the modeling approaches. While model assumptions inherent in uSEM correspond to causal hypotheses by nature, those in gVAR are only existence or absence of associations. Therefore, casual models have stronger assumptions and confirmatory tests, and thus is able to offer causal inference and to avoid spurious associations, but is sometimes at the risk of finding false positive structures when the reality is association. In contrast, the downside of GGM is a loss of ability to retrieve the direction of effect or to inform any causal implications. In some cases, this can lead to spurious relations. One benefit of this, however, is that GGM is less likely to commit false causal relations when an edge is actually mere association.

A question facing many idiographic researchers is whether the results and interpretation will differ as a result of choosing one model over the other. The primary difference is in the handling of contemporaneous relations. Turning first to gVAR, a synthesis of GGM and VAR, the direction recovery is handled differently for lagged versus contemporaneous relations: the VAR component of gVAR models the lag-1 autoregressive and cross-lagged structural relations between variables as direct effect, while the GGM models lag-0 contemporaneous relations as (undirected) using residual covariance. The gVAR and the general PNM literature indicate that these can still be indicative of causal relations, or “causal interactions" (Epskamp, Waldorp, et al., 2018; Fried, 2020). As noted by Borsboom and Cramer (2013), these networks rely on the fundamental premise that symptoms “causally" influence one another as part of a complex dynamical system, thereby contributing to study the onset, maintenance, and intervention. On the other hand, one characteristic of the uSEM that contrasts with that of gVAR approach is to treat contemporaneous connections as structural relations by entering the model as direct regression effect, which provides direct causal evidence from the recovered dynamic network. Although a detailed discussion on model choice is not the focus of the current discussion, we caution that a choice between a causal and non-causal model is not purely statistical. Indeed, one needs to have strong theoretical backgrounds to guide through the process.

To help us further unpack the implications of having directed and undirected contemporaneous relations, we revisit a few concepts extensively: 1) directed acyclical graphs; 2) causality from within the economics perspective; and 3) the notion of exogeneity. In each of these sections, we review the fundamentals particularly those relates to the current context with both contemporaneous and temporal relations. We illustrate the implications for gVAR and uSEM using some simple simulated examples throughout.

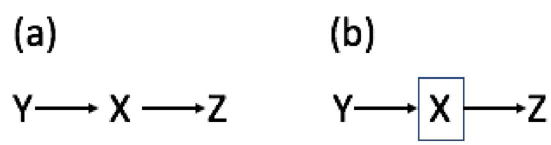
2.3 Causality from a Directed Acyclical Graph Perspective

The directed acyclic graph (DAG), being influenced by path analysis and classic SEM, has played an essential role in causal inference in social sciences. DAG can be understood as a graphical representation of qualitative causal assumptions on the data generating process in the population. DAG graphs encode researchers' hypotheses on how the world works. To do this, DAG researchers establish rules to map their casual assumptions (i.e., statements about probability distributions) onto associations and dependencies within observational data, and provide relation-by-relation confirmatory test on these explicit casual hypotheses (Elwert, 2013; Pearl, 2009). Relatedly, DAGs are nonparametric: they make no statement about the distribution of the variables or the functional form of the direct effects, or the magnitude of the causal effects. That is why some researchers referred to DAG as nonparametric structural equation models (NPSEM; Elwert, 2013). This also makes the DAG framework appropriate for the interpretation of network models where the interest is more so on path recovery (for causal inference) than on parameter estimates (for statistical inference). To facilitate discussion, we briefly introduce the basic concepts and definitions in DAG (readers should refer to Elwert, 2013 and Pearl, 2009 for a complete review on the theory and extended topics).

Basic graphical terminology, three structural sources of observable associations, and three corresponding biases. Every DAG is composed of three elements: variables (nodes, vertices), arrows (edges), and missing arrows. *Arrows* represent an existence of *direct causal effects* between pairs of variables and the order of information flow in time. Relatedly, *missing arrows* indicate a strong assumption, i.e., the so-called "strong null" hypothesis of no effect, or the *complete absence* of a direct causal effect between two variables for every member of the population (keep in mind that for time series applications, the population is all time points for a given person). To derive the testable implications of a causal model, DAG defines three sources of associations to translate between association and causation. All DAGs can be constructed from three elementary causal structures: chains $Y \rightarrow X \rightarrow Z$ (in which X is called the “mediator"), forks or common cause $Y \leftarrow X \rightarrow Z$ (in which X is called the “common cause"), and inverted forks $Y \rightarrow X \leftarrow Z$ (in which X is called the “common outcome" or a “collider"). These three structures correspond exactly to *causation, confounding, and endogenous selection* by the following rules.

First, two variables are said to be marginally associated if one variable directly or indirectly causes the other. For instance, Figure 2.1(a) shows an example where Z and Y are marginally associated because Z is an indirect cause of Y, the effect is mediated by X. In the case where Y and Z are conditionally independent given X (see Figure 2.1(b)), the conditional association does not identify the causal effect of Y on Z. We say that conditioning on X “blocks the path" or “closes the blackdoor" from Z to Y, and refer to this case as *overcontrol bias* (Elwert & Winship, 2014).

Figure 2.1: Basic DAG Structure 1: Association by Causation

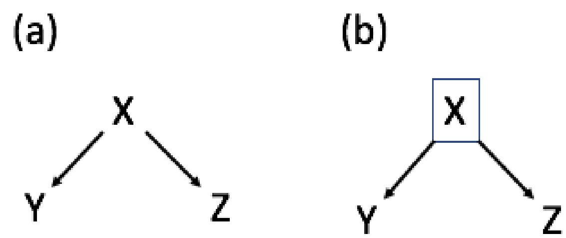


Note: (a) Y and Z are associated by causation. This marginal association identifies the causal effect of Y on Z. (b) Y and Z are conditionally independent given X. This conditional association does not identify the causal effect of Y on Z (overcontrol bias). Figure note adapted from Fig.13.2 in Elwert (2013).

Second, two variables Y and Z can be associated due to a shared common cause X (see Figure 2.2(a)). If these two are correlated only because they both are caused by X, failing to condition on X leads to the familiar situation of *common cause confounding bias.* In this case, the marginal association between Y and Z is spurious, because it does not identify a causal effect of Y on Z or vice versa. Conditioning on the common cause X, i.e., closing the blackdoor, would *eliminate* this spurious association (see Figure 2.2(b)). Hence, an existence of conditional association between Y

and Z given X would identify the causal effect.

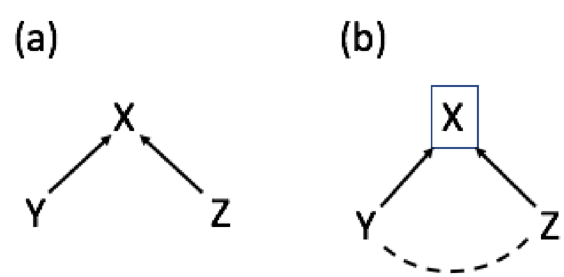
Figure 2.2: Basic DAG Structure 2: Association by Common Cause



Note: (a) Y and Z are associated by common cause. The marginal association does not identify the causal effect of Y on Z (confounding bias). (b) Y and Z are conditionally independent given X. The conditional association does identify the causal effect of Y on$Z$. Figure note adapted from Fig.13.3 in Elwert (2013).

Third, Y and Z in Figure 2.3(a) are marginally independent, i.e., they neither cause each other nor share a common cause. We say that the absence of marginal association between Y and Z identifies a null causal effect. Here if we instead conditioned on the common outcome or collider X (as shown in Figure 2.3(b)), however, a *spurious* nonzero association between Y and Z would have occurred, an issue being referred to as *endogeneous selection bias*. The remedy is *not* to condition on such variables, or we say to let the backdoor open.

Figure 2.3: Basic DAG Structure 3: Independence with a Common Outcome



Note: (a) Y and Z are marginally independent. The marginal association identifies the causal effect of Y on Z. (b) Y and Z are associated due to conditioning on a common outcome (collider). The conditional association between Y and Z given X does not identify the causal effect of Y on Z (endogenous selection bias). Figure note adapted from Fig.13.4 in Elwert (2013).

To recap, *overcontrol bias* results from conditioning on variables in a causal path between, *confounding bias* arises from failure to condition on a common cause, while *endogenous selection bias* results from conditioning on a collider on any path connecting treatment and outcome. This leads to the definition of "D-separation" (Pearl, 2009): two variables is said to be *d-separated* if all the paths between the two are blocked or the backdoors of them are closed, which entails that (1) the path contains a noncollider that has been conditioned on (such as Figure 2.1(b) and Figure 2.2(b)); and (2) the path contains a collider that has not been conditioned on (such as Figure 2.3(a)). Otherwise, a path is said to be *d-connected* if it is not d-separated, i.e., at least one noncollider path is unblocked or a backdoor is open or at least one collider path is blocked or closed. The concept of d-separation distinguishes the three sources of association â causation, confounding, and endogenous selection â into a general graphical rule to determine whether a path transmits association or causation. D-separation makes the distinguish of causation from association remarkably straightforward, and is the essential tool underlying all graphical identification criteria.

DAGs are called "acyclic" as they may not contain directed cycles, that is, paths that can be traced strictly along the direction of the arrows to arrive back at the starting point. Acyclicity assumption preserves the principal that the future cannot cause the past. However, this assumption is largely due to the predominate application of DAG in cross-sectional study, or longitudinal study with static model where temporal dependency is not interested or being accounted for in the model. The concept of temporal order, or the information flow, is embedded in the direction of arrow. That is why the majority of theoretical and practical applications of DAGs in the literature assume that true simultaneity (i.e., Y causing Z and Z causing Y) does not exist. But theory exists for cyclic graphs that incorporate the temporal sequence of events to resolve apparent counterexamples (Greenland et al., 1999). We shall later turn to a detailed discussion of causality under the time series context.

2.3.1 Using DAG to Interpret uSEM and gVAR

It is important to distinguish between models found by gVAR or uSEM. Relations recovered by gVAR or uSEM may entail drastically different causal interpretations. When fitting a uSEM, we can directly apply all the rules in DAG theory to inform the causal mechanism of a dynamic network. For example, if the interest is to recover underlying causal mechanisms behind marginal correlations, following DAG rule on directional and acyclic associations can help distinguish direct cause or mediational effects from spurious ones such as confounding effects or collider effects. The distinction between uSEM and gVAR is that theoretically DAG is applicable to both the lag-1 (temporal) level and the lag-0 (contemporaneous) level relations in uSEM, while only to the lag-1 PDC relations recovered by gVAR. This is because in gVAR only undirected contemporaneous edges are formed, representing conditionally dependent connections between interacting nodes, as previously introduced. If this undirected edge is due to an omitted variable, then this can be part of DAGs. If applications of gVAR attempt to map the causal structure amongst symptoms of traditionally-defined psychological concepts under the DAG rule, it directly applies to the lagged relations. The validation of these causal claims delivered by undirected contemporaneous relations in PCC pertains to two unsolved questions:

1) *To what extent an undirected PCC network can be indicative of true contemporaneous causal relations or distinguish them from spurious ones*;

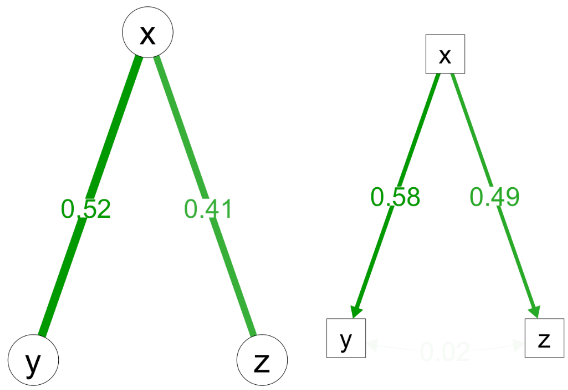
2) *How accurate and reliable it is to use the PDC network in representing all casual relations in a dynamic system.*

The remaining part of this section aims to answer question 1) and we shall return to question 2) after introducing the concept of Granger Causality. To do so, we provide some simulated examples below to demonstrate the consequence of fitting an undirected model to data generated by a causal model. For a complementary comparison, we also provide counter-examples where the consequence of imposing non-existing causal pathways while the ground truth is pure association.

Let us first turn to the illustration of fitting an undirected model via partial correlation versus an directed structural model to data generated by DAG causal models. This resembles the contemporaneous relations among variables recovered by GGM versus uSEM after considering any lagged relations that may exist in the data. In the first set of examples, data were generated by three distinct DAGs (Figures 2.4-2.6, respectively). Results are shown when data were fitted by a GGM (left panel) and by the corresponding data generating SEM with true causal assumption (right panel). In DAG models as shown in Figure 2.4 and Figure 2.5, variables Y and Z are independent given X in the case where either X is a common cause in a): $Z \leftarrow X \rightarrow Y$, or that X completely mediates the relationship between the two in b): $Z \rightarrow X \rightarrow Y$.

As expected, SEM recovered the structures used in the DGM, while the same GGM was returned for both DGMs. That is, the same edge links $Z - X - Y$ were found in both GGMs when

Figure 2.4: Fit A DAG Model with Common Cause in GGM (left) versus in SEM (right)



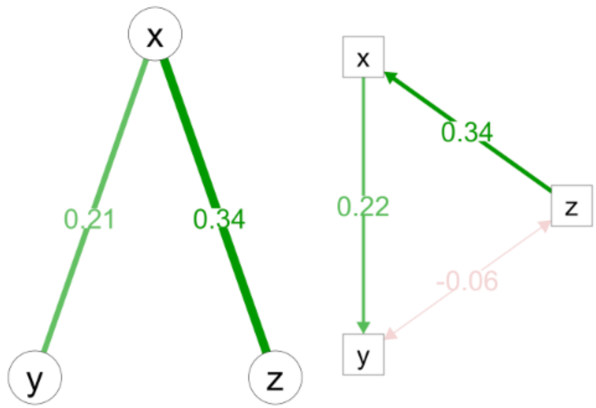
Notes: 1) Plots are consistent with the conventions in respective literature and those provided by the corresponding program. In GGM, variables are represented by node with a circle, and lines without arrow indicate nonzero edges recovered by the model. In SEM, observed variables are squared and latent variable are circled. Lines with arrows are paths recovered by SEM, with one-head arrow representing regression coefficient, and double-headed arrows representing covariance coefficient; 2) Green lines represent positive edges (in GGM) or paths (in SEM), red lines represent negative edges (in GGM) or paths (in SEM); 3) The number along each line is the actual standardized parameter estimate in both models, the width and transparency of the line also represent the relative strength of the relation.

the reality is $Z \leftarrow X \rightarrow Y$ or $Z \rightarrow X \rightarrow Y$. This clearly showed that an undirected GGM cannot distinguish the present of a common cause from an existence of underlying (mediational) causal pathway when a conditional independent relation is observed. By incorporating partial correlation, i.e., conditional correlation that partials out all other variables in the model, GGM automatically rule out spurious relations caused by confounding bias or endogenous bias; however, the downside of avoiding any structural relation is the lose of causal information.

A potentially more problematic situation arises when fitting a GGM to data generated by a DAG model in Figure 2.6 with inverted forks (or a chain with colliders, $Y \rightarrow X \leftarrow Z$. Here, conditioning on the collider or common outcome X introduces a spurious edge, $Z - Y$.

Therefore, this example confirms that if the data generating model (DGM) is a causal model, causal assumptions can only be directly tested by structural models. This is consistent to previous knowledge that DAG is a nonparametric representation of SEM, and each DAG corresponds to an equivalent SEM. However, a DAG cannot be accurately represented by a non-casual models such as GGM. DAGs with distinct assumptions and causal implications might be translated into equivalent

Figure 2.5: Fit A DAG Model with Mediation Path in GGM (left) versus in SEM (right)



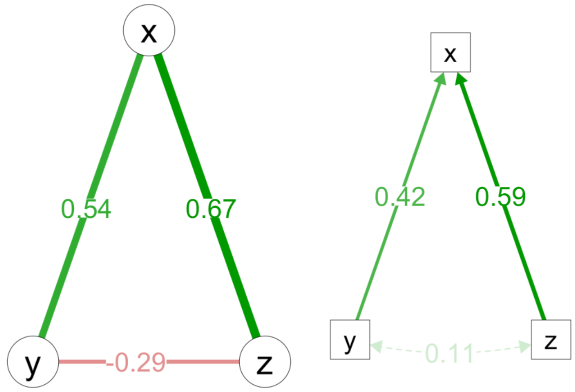
Notes: same with above.

GGMs, at the risk of losing important causal inferences. Going back to question 1) posed earlier, since an undirected PCC matrix can be transformed to *non-unique* DAGs or even one with spurious causal link, it is not an accurate or reliable representation when causal relations are hypothesized.

Nevertheless, this is not to suggest that DAGs should be used in every situation. There are scenarios when theory is scant to support causal hypotheses. As an example, if the DGM contains mere associations instead of any DAG relations, which is essentially a correlation or covariance matrix with no additional causal information, a non-causal model such as GGM might be a more appropriate model choice than a causal model SEM. Because imposing any directional pathways to null-directional relationships is a false causal assumption by itself. Indeed, we found that a GGM can accurately represent a correlation matrix (see Figure 2.7 (a)), while the following three SEM models fit the data equally well (see Figure 2.7 (b)): partial mediational pathway, common cause with correlated residuals, and common outcome of correlated predictors (from left to right, respectively). When the theory to guide a model choice is absent, researcher might arrive arbitrarily at any of the three models with distinct causal interpretations. Hence, when casual assumptions are attempted with a lack of evidence, it could lead to false causal discoveries.

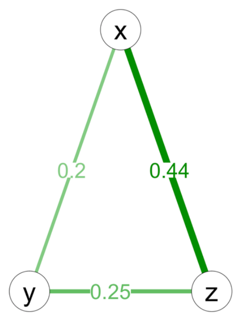
In sum, a causal model should be adopted when there is theory to support a causal hypothesis or evidence to inform directional structures. When not available, we might be on a safer side to choose a non-structural model with no causal assumption, but accordingly the interpretation should *not* contain any causal inference or be used to inform any causal implication.

Figure 2.6: Fit A DAG Model with a Collider in GGM (left) versus in SEM (right)

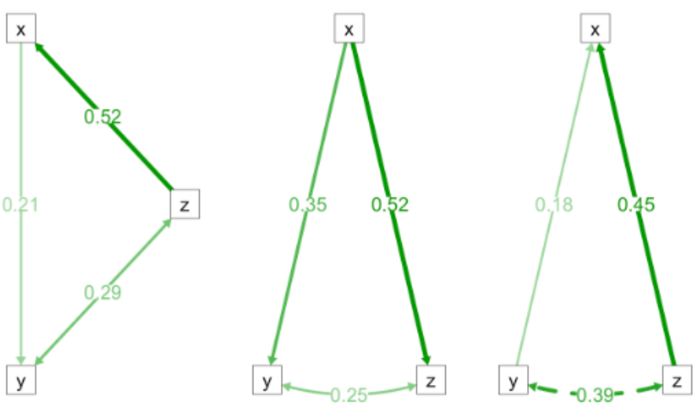


Notes: same with above.

Figure 2.7: Fit A Correlation Matrix in GGM (upper) versus in SEMs (lower)



(a)



(b)

Notes: same with above.

2.4 The Granger Causality Theory

In the previous section, we reviewed the definition and properties of DAGs. We discussed DAG rules and their implications in static models within the cross-sectional framework from which they originally were developed. These interpretations can also be applied in the time series context. But first let us turn to a related concept that was introduced from a time series perspective: i.e., the *Granger Causality* (GC; Geweke, 1982; Granger, 1969). Deeply rooted in Econometrics literature, GC was developed as a statistical tool to test the existence of observed causal relations in dynamic systems. GC is defined together with a list of terms and concepts to facilitate the discussion of dynamic causality, such as dynamic simultaneity and endogeneity, recursive and non-recursive systems, predeterminedness, structural invariance, and exogeneity (e.g., Engle et al., 1983; Koopmans et al., 1995; Strotz & Wold, 1960). We start with the definition of GC and shall turn to some of the others next.

Within a dynamic system, causality is determined by the level of predictability, that is, the current status of one variable is more efficiently predicted if the process of the other variable is taken into account in addition to all other information accumulated as of that moment (Granger, 1969; Lütkepohl, 2005). At the very least, the direction of effect flow can provide evidence of a casual pathway, satisfying the general condition that the cause precedes the effect. In multivariate IAV approaches, a uni-directional GC is said to occur when variable one has a lagged effect on variable two after taking account for the autoregressive process of variable one on itself. The effect could also be bidirectional (or reciprocal). When this occurs we say a cross-correlation effect is observed, indicating a *two-way* GC. In psychology, a two-way GC is not uncommon.

However, what is often less interpreted in applied research outside of Econometrics is the concept of *instantaneous* or *contemporaneous* GC. The definition of GC as first provided by Granger (1969) includes contemporaneous causal relations, i.e., a lag-0 instantaneous predictive effect occurred at the same measurement window. More formally, a given variable X is said to *Granger cause* variable Y if current Y values are better predicted with values of X (either contemporaneously or lagged) than with past values of Y alone (Granger, 1969; Lütkepohl, 2005). In either case the direction of the effect (i.e. $X \xrightarrow{GC} Y$ or $Y \xrightarrow{GC} X$) discloses the information flow, with there being the possibility of what is referred as the “simultaneous causality", “mutual causation", or “bi-causality" (Strotz & Wold, 1960), where causal relations exist in both directions.

In theory, a simultaneous causality or mutual causation can take place, i.e., one exerts immediate effect on each other. The most classic example is perhaps that rolling balls instantly change moving direction as a result of bouncing against each other. Heuristically, if the time lag of an effect is shorter than the measurement scale, a lagged effect might not be observed. Instead, the existing association can be modeled as a contemporaneous connection. In other words, an instantaneous GC is indicative of a causal effect that is failed to be captured by temporal dependency due to inappropriate time scale, which often occurs in studies like fMRI where the underlying neuronal brain activity happens much faster (on the scale of milliseconds) than what is captured via the MRI (typically on the order of seconds). Thus, an instantaneous GC occurs when the time lag of the effect of interest is so short that approaches infinitely to zero. Indeed, mutual causation in a dynamic system can be interpreted as a limit form of the lagged effect schemes where the time lag is reduced towards zero (Strotz & Wold, 1960). Taking these into consideration, the inclusion of instantaneous GC relations is essential to the completeness of causality.

2.4.1 Nonrecursive Relations from a Two-way GC Perspective

The concept of instantaneous GC provides a theoretical foundation to the understanding and interpretation of dynamic causal relations. It helps extend a static DAG with the strict “acyclicity" assumption to a dynamic DAG with the “cyclicity" flexibility. Simply as opposite to “acyclicity", the “cyclicity" assumption allows for “nonrecursive" relations that can be a feedback or reciprocal connection between two variables (e.g., $X \to Y \to X$), or a cyclic chain when more than two variables are involved (e.g., $X \to Y \to C \to X$). The concept of feedback effects or cyclic chains is an ambiguous causal evidence in cross-sectional studies using a static model, due to a lack of repeated measures with a temporal order. But such interpretation is intuitive under the context of longitudinal or time series studies. From the “simultaneous causality" point of view, a nonrecursive relation (or a cycle) reveals variables as instant causes of each other, directly or indirectly (i.e., mediated through other variables). From the practical “approximation of reality" perspective, when repeated observations are measured at an inadequate time scale to capture short-term causal effect, reciprocal relations can be adequately modeled as an approximation to cross-lagged effects unfolding over time. In both scenarios, the presence of reciprocal relations or cycles provides the evidence to existing two-way instantaneous GC relations.

To make this more concrete, let’s say that for one individual fatigue causes a more stressful feeling, which in turn makes one more likely to fatigue. However, when the time scale is limited, i.e., the effect is quick or immediate that it is shorter than the length of the measurement window, cross-lagged effects over time are likely to be “collapsed" into the same measurement and thus be modeled as contemporaneous bidirectional or reciprocal effects. As another example, symptoms might cause each other within a few hours. But if the observation time scale is on the order of weeks or even months, this may look like bidirectional causality. In other words, a reciprocal relation or a cyclic chain at the lag-0 (contemporaneous) level, can be considered as an approximation to a bidirectional cross-lagged effect that happens in reality yet failed to be captured at the lag-1 (temporal) level, which in turn reflects a two-way (instantaneous) GC relations.

One may remember that the “acyclicity" assumption is important to ensure model identification in SEM, an immediate question is how to identify a nonrecursive relation since it is a clear violation of such assumption. A nonrecursive relation can be identified when other predetermined variables are available (Rigdon, 1995), which applies to time series data because more often than not AR or cross-lagged effect present and the lagged variables are treated as predetermined variables in the model.

2.4.2 Using GC to Interpret uSEM and gVAR

Now let us revisit the second question posed in Section 2.3.1, i.e., whether it is adequate to rely on the directed PDC network of the gVAR alone to recover all casual relations in a dynamic system. It is clear that when a one-way instantaneous GC occurs as indicative of instant causality or as resulted from short-term effects missed between repeated measures, gVAR would be very limited to represent such GC relations. Because if instantaneous GC was not captured by temporal dependency between repeated measures, the relations will not be recovered as nonzero edges in the directed lag-1 PDC network. Although the associations might be picked up by residual correlations, which in theory can be recovered as undirected edges in the lag-0 PCC network, we know from previous demonstration that undirected edges are inaccurate and sometimes unreliable in representing causal relations.

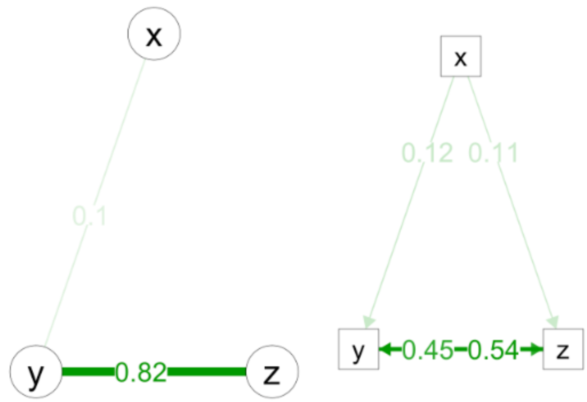
However, there is one scenario we had not covered in our demonstration using the DAG with the “acyclicity" causal assumption in the DGM. What becomes relevant to dynamic models is when the “cyclicity" is allowed, and mutual causality or nonrecursive effects can occur as representatives of two-way instantaneous GC relations. Let’s suppose that “cyclicity" relation exist, i.e., the association between two variables are reciprocal or nonrecursive, the question is to what extent it is meaningful to distinguish a two-way structural relation (as those in uSEM) as opposed to a bidirectional residual covariance between two observed variables such as those in VAR, or as an undirected partial correlation in lag-0 PCC of gVAR? Indeed, one might wonder whether an undirected link from the PCC network in gVAR is equivalent to a reciprocal relation or nonrecursive cycle in the representation of a two-way GC. Put it in another way, would they all be adequate to model an existing two-way instantaneous GC that either represents a simultaneous causality or as an approximation to unmodeled cross-lagged effect between the two time series?

We attempt to answer this question from both an empirical ground and a conceptual ground. From an empirical perspective, we shall use a simulated example to investigate the consequence of using GGM with undirected link to represent a reciprocal causal relation. To do so, we fit GGMs to data generated by two SEMs with variables Y and Z being a mutual cause of each other (see Figure 2.8, note that X is a predetermined variable function as a common predictor for Y and Z in order to identify the model). The two DGM models correspond to cases where the mutual effects have the same sign (in model (a) both are positive) and the opposite signs (in model (b) one is positive and the other negative), respectively.

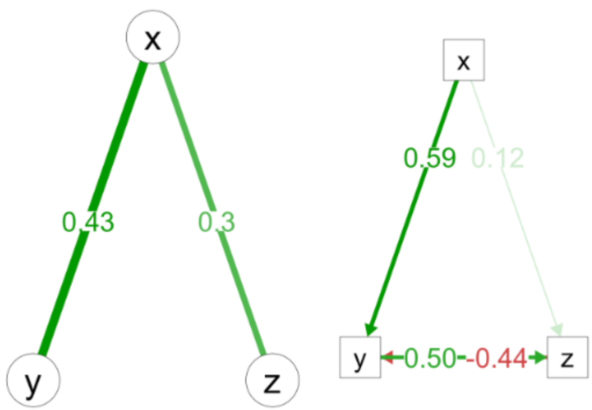
Interestingly, we found that when the signs of the reciprocal relations were consistent (both positive), a strong positive edge $Y - Z$ was produced by GGM as shown in Figure 2.8(a). This indicates that GGM can recover the two individual unidirectional paths as an undirected partial correlation. However, the magnitude of the edge link was much larger than the actual effect of either direction. Besides some analytical proofs, it could be understood intuitively that the two individual effects were *augmented* when estimated by one coefficient. In contrast, when the signs turned opposite, an absent edge was observed in GGM even though the magnitudes of the reciprocal effects were similarly strong (see Figure 2.8(b)). The reason is also intuitive: the underlying reciprocal relations with opposite effect signs could be partially or even completely canceled out. As expected, the SEMs from both cases recovered precisely the corresponding reciprocal relations as two unidirectional effects.

This example shows that an undirected model is not reliable in representing a reciprocal causal relation: GGM would overestimate the magnitudes of individual effects when they have the same sign, yet underestimate them when they have opposite signs. In fact, reciprocal causal relations with opposite signs are not uncommon in practice, particularly in fMRI studies, e.g., an inhibition reciprocal or impluse feedback relation between ROIs. Whenever this is the case, each individual path of the two-way relations needs to be estimated separately with an independent model

Figure 2.8: Mutual Causality Recovered by GGM (left) versus SEM (right)



(a) DGM: SEM with positive reciprocal effects between y and z



(b) DGM: SEM with positive and negative reciprocal effects between y and z

coefficient. The practice of using one coefficient (e.g., a bidirectional covariance or an undirected partial correlation) to estimate two individual paths will lead to biased parameter estimates and misleading inferences. Therefore, by the same token, a two-way instantaneous GC, regardless of mutual causality, cannot be accurately represented by the undirected lag-0 PCC network in gVAR. To conclude from an empirical standpoint, because the undirected lag-0 PCC cannot be accurately or reliably depict varied causal structures, relying on the lag-1 PDC alone is insufficient to recover either a uni-directional causal relation (in representation of one-way GC) or a bi-directional reciprocal causality (in representation of two-way GC in a dynamic system).

2.4.3 Nonrecursive Relations from an Exogeneity Perspective

To understand the theoretical difference in the contrasting ways “acyclicity" or nonrecursive relations are represented, we shall now turn to the concept of “exogeneity" within a dynamic system.

By definition, an endogenous variable refers to an outcome or dependent variable that is being predicted by other variables in the system of variables available, as contrast to an exogeneous variable which enters the system as predetermined and functions solely as a predictor or independent variable. But the level of predictability versus uncertainty in an endogenous variable can vary. Formally, a variable y_t is defined to be *weakly* exogenous if inference of the set of parameters in the model conditional on variable y_t involves no loss of relevant information as compared to using the joint distribution. In other words, knowing the marginal distribution of y_t is irrelevant to the estimation so that it is sufficient to use the the conditional distribution to approximate the joint density (Egler et al., 1983). The definition of a *weakly exogenous* variable pertains to efficient estimation. Further, a variable y_t is said to be *strongly exogenous* if it is *not caused* in the sense of GC by any of the *endogenous variables* in the model, in addition to being *weakly exogenous* (Egler et al., 1983). Strong exogeneity pertains to the predictability of the variable from the system, where higher levels of exogeneity (or a system composed of many strongly exogeneous variables) involves fewer GC relations (Egler et al., 1983).

This is to say that a model having many weak and/or strong exogenous variables can be estimated fairly efficiently, by using a factorization of conditional distributions instead of joint distribution, and is not imposing or testing GC relations (Egler et al., 1983). But to the point view of the authors, this is achieved upon strong (and sometimes unrealistic) exogeneity assumption, as well as at the sacrifice of an arbitrary structure. Because information carried by variables with strong exogeneity can go *in to* the system, but only a few or minimum can *come out* of it. A system composed of primary strong exogeneous variables is dependent on external world (where exogeneous variables come from). On the contrary, a system with fewer weak or strong exogenous variable provide a more confirmatory, self-independent structure, and is less dependent on external variations.

2.4.4 Using Exogeneity to Interpret gVAR and uSEM

If we apply these concepts to current investigation, we find the two representations of VAR model extreme cases with respect to the level of exogeneity allowed in a dynamic system. On the one hand, gVAR first adopts a *weak exogeneity* assumption by estimating the conditional correlation rather than the joint distribution. Furthermore, a *strong exogeneity* assumption is imposed in that associations are modeled as undirected partial correlations among the residuals and no direct regression coefficients amongst contemporaneous themselves are allowed (the latter would represent instantaneous GC effects from other endogeneous variables). Contemporaneous variables are endogeneous relative to the lagged variables because directed autoregressive and/or cross-lagged effects are incorporated and often observed. This will not break the strong exogeneity distinction for the contemporaneosu varibles because lagged variables are treated as strong exogeneous.

Therefore, the contemporaneous variables in gVAR model meet the two criteria and thus are endogenous variables with strong exogeneity. The variances of contemporaneous variables are only explained by other exogenous (the lagged) variables, leaving the remaining associations entirely correlated residuals. Such specification entails that the source of contemporaneous associations is completely exogenous to the network system, preventing it from representing simultaneous causality or instantaneous GC relations. Such a dynamic model imposes very strong exogeneity assumptions to achieve estimation efficiency, yet at the cost of reaching at a conservative, potentially less-informative description of the system. This is part of the reason why researchers have consistently found undirected psychometric network models vulnerable to unmodelded measurement error or omitted variable bias (Hallquist et al., 2019; Henry & Ye, 2019). Indeed, a system with a strong exogeneity assumption achieves high estimation efficiency at the sacrifice of being subject to instability and non-robustness to model misspecification, as has already been pointed out by evaluation studies on PNM in general (e.g., Forbes et al., 2017; Hallquist et al., 2019).

On the other end of the spectrum, a minimum exogeneity assumption is imposed in the uSEM where estimation is performed on the joint distribution of the contemporaneous variables (given only exogeneous lagged variables), and that all contemporaneous associations in the uSEM are modeled only as directed relations amongst themselves. As a result, endogenous contemporaneous variables become part of the GC sources of variability to each other. This entails that the variance of a given variable y_t can be explained by other contemporaneous variable(s) at time t conditioning on predictions from its previous measurement(s) or other lagged variables, assuming a conditionally independent white noise process. That is, the multivariate white noise processes are mutually independent at each time point. This is to say that any contemporaneous variable involved in the directional contemporaneous correlation can contribute to a more efficient prediction in the sense of instantaneous GC for another contemporaneous variable. The downside is the loss in estimation efficiency because only the minimum predetermined (strongly exogenous) variables (lagged variables) are assumed, although it is complemented by the added value to model simultaneous causality or instantaneous GC. If lagged effects are not captured by the data due to the temporal resolution of the measurements (i.e., the data are collected at a longer time scale than the process being studied), they still have the chance to be represented as instantaneous directed effects. By avoiding the strong exogeneity assumption, uSEM aims at a more confirmatory and informative model.

Nevertheless, we imagine some middlepoint between the two ends is perhaps more realistic in practical settings. The gain of having some weakly exogenous variables is estimation efficient, but if the level of exogeneity goes too far (i.e., variables meet the criteria to be strongly exogenous), the loss is GC inference. Endogenous variable y Granger causes z and therefore z is not strongly exogenous or strictly exogenous. However, the important criterion for efficient estimation is weak exogeneity, not strong exogeneity. The two concepts serve different purposes: weak exogeneity validates conducting inference conditional on z, while GC validates forecasting y conditional on the future z's. As is known, the condition that y does not Granger cause z is neither necessary nor sufficient for the weak exogeneity of z. Obviously, if estimation is required before conditional predictions are made, then strong exogeneity which covers both GC and weak exogeneity becomes the relevant concept. One promising future direction to reach a trade-off between estimation efficiency and GC inference is to adopt only weakly exogenous contemporaneous variables (estimating conditional distribution) but not strong ones (allow GC on endogenous variables).

2.5 Empirical Examples

In this section, we illustrate the different interpretations by using two empirical time series data. We first fit a gVAR and uSEM model to publicly available daily dairy data from a study conducted by Fisher and colleagues (2017), in which they study idiographic dynamics of mood and anxiety symptomatology. Next, we repeat this procedure on another data set obtained from the Autism Brain Image Exchange (ABIDE; Craddock et al., n.d.; Di Martino et al., 2014) repository of fMRI data. In both empirical examples, we compare the models recovered by the two approaches, and evaluate to what extent the interpretations using the causality theories discussed previously differ as a result of adopting directed versus undirected contemporaneous networks. By making a comparison on the two approaches in analyzing the most common types of time series data currently seen in psychological studies, the goal here is to interpret the differential interpretations based on their own model assumptions when *the ground truth is unknown.* Lastly, we summarized the characteristics of the dynamic model produced by either approach. To inform applications, we also offered some precautions regarding potential issues and caveats arising from these model assumptions that researchers should be aware of when adopting one over the other. For the analytic procedure, data estimated by uSEM model was fit using the R package *gimme* (Lane, Gates, Fisher, et al., 2019), and the gVAR model fit via the R package *graphicalVAR* (Epskamp, 2018).

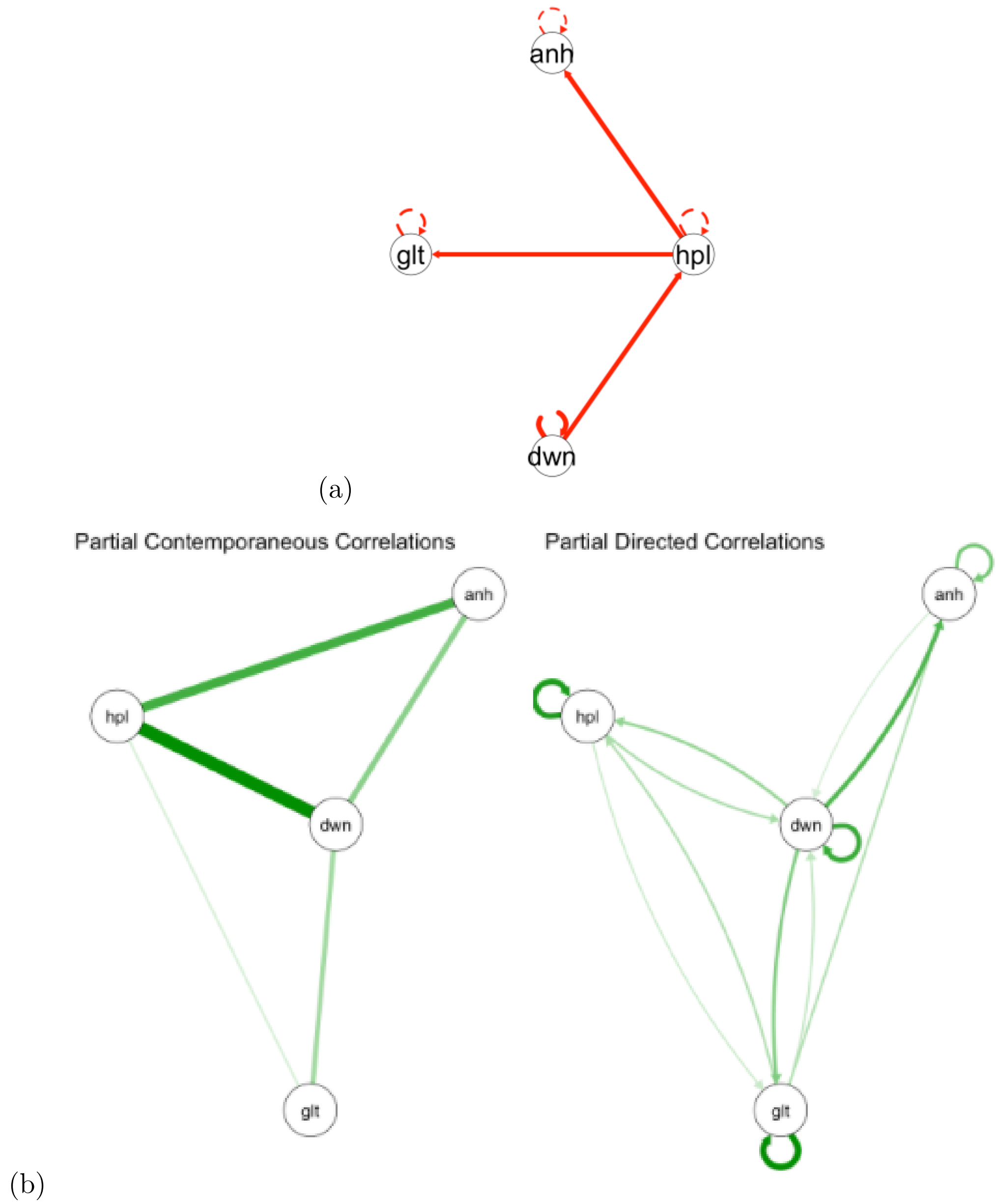
2.5.1 Empirical example I: fit gVAR and uSEM to EMA data

In the first empirical example, we used EMA time series data collected to study idiographic pattern of mood and anxiety symptomatology. These data have been explored previously using the uSEM search procedure described here (Fisher & Boswell, 2016) as well as a variate of graphical VAR model that is similar to but slightly different from the gVAR approach described here (Fisher et al., 2017). Data were collected on 40 individuals with generalized anxiety disorder, major depressive disorder, or comorbid diagnoses. Participants answered questions about 21 descriptors of mood and anxiety symptomatology 4 times a day over a period of approximately 30 days. To illustrate, we used data from one exemplar participant whose results were reported in the papers (i.e., P25). For simplicity, we fit the uSEM and gVAR model to a subset data of four variables measured on this participant: *down*, *hopeless*, *guilty* and *anhedonia*.

Exampler Result from the EMA data Figure 2.9 displays the network models for the exampler from uSEM and gVAR. Graph (a) presents result from the uSEM model, with solid lines for lag-0 paths and dashed lines for lag-1 paths. In following heat map conventions, red indicates a positive coefficient estimate and blue indicates negative. Graph (b) presents the gVAR results for the lag-0 PCC network (left) and lag-1 PDC network of the gVAR model (right), respectively, with green indicating a positive coefficient estimate here and red for negative, as recovered by the hyrbid-lasso regularized approach implemented in the package.

Relations recovered by the uSEM were primarily at the contemporaneous level. The contemporaneous network from uSEM revealed a mediational pathway, where feeling *down* directly relates to predicted increases in the feeling of *hopeless*, which in turn relates to more *guilty* and *anhedonia*. The pathway from *down* to *guilty* and *anhedonia* is d-separated by *hopeless*, while *guilty* and *anhedonia* are independent experiences after conditioning on the common cause *hopeless*. All symptoms exhibit an autoregressive (AR-1) process, yet no cross-lagged effect was found after conditioning on the AR effects, suggesting that the mood symptomatology and depression disorders

Figure 2.9: EMA Exampler Result from uSEM (a) versus gVAR (b)



Note: 1) hlp = hopeless, dwn = down, glt = guilty, anh = anhedonia; 2) Plots are consistent with those provided by the respective packages. For the uSEM results, red indicates positive values and blue for negative ones; for gVAR, green indicates positive values and red indicates negative. Dashed lines indicate lag-1 relations in uSEM. Line width correspond with coefficient estimate size for both.

were likely co-occurring based on the measurement time scale. The experience in one symptom at one time is unlikely to (or has not yet) transmit to variation in other symptoms over time, even to the next measurement.

In terms of interpretations from a GC perspective, *down* has the strongest exogeneity as it is only predicted by the exogenous variable of itself at the prior time point. *Hopeless* is weakly exogenous (only Granger Caused by the exogenous variable *down*), while *guilty* and *anhedonia* have the weakest exogeneity as they are both Granger caused by the weakly exogenous variable *hopeless* and thus are considered endogenous variables since they can be predicted by other contemporaneous variables in the system. Putting it together, the recovered dynamic model by uSEM traced back to the source of variation and informed us the potential direction of the information flow.

The networks recovered by gVAR represented a somewhat different story. First, the relations are bidirectional (or undirected) in contemporaneous variables, suggesting a dynamic with nonrecursive interactions rather than some information flow in a direction. Specifically, in the lag-0 PCC network, all pairwise partial correlations were observed except for *guilt* and *anhedonia*, making *down* and *hopeless* equally central to the contemporaneous network. Compared to the model recovered by uSEM, the relation from *down* to *guilty* or to *anhedonia* was no longer blocked (or "d-connected"), leading to additional relations. Consistent with the uSEM though, *guilt* and *anhedonia* were independent given the other two observed variables, although uSEM only suggests conditioning on *hopeless* while gVAR does *hopeless* and *down*.

A greater number of differences of the two approaches were observed with respect to the temporal models: the lag-1 PDC network was much denser than that of the uSEM. Specifically, *down* was the most central node that showed reciprocal relations with all other nodes, followed by *hopeless* and *guilty*. We see that gVAR result emphasized the co-existence of both co-occurring (instant) as well as delayed (lagged) bidirectional effects. The network recovered by the gVAR model suggested that mood symptomatology and depression disorders are characterized by their mutual influences on each other, although it did not further unpack the sign and strength of the influence from one to the other. The nature of the relations recovered by gVAR were that symptoms are somewhat mutually related to each other, and they are equally strong exogenous variables to the dynamic system instead of any one being more a source than an outcome.

2.5.2 Empirical Example II: fit gVAR and uSEM to fMRI data

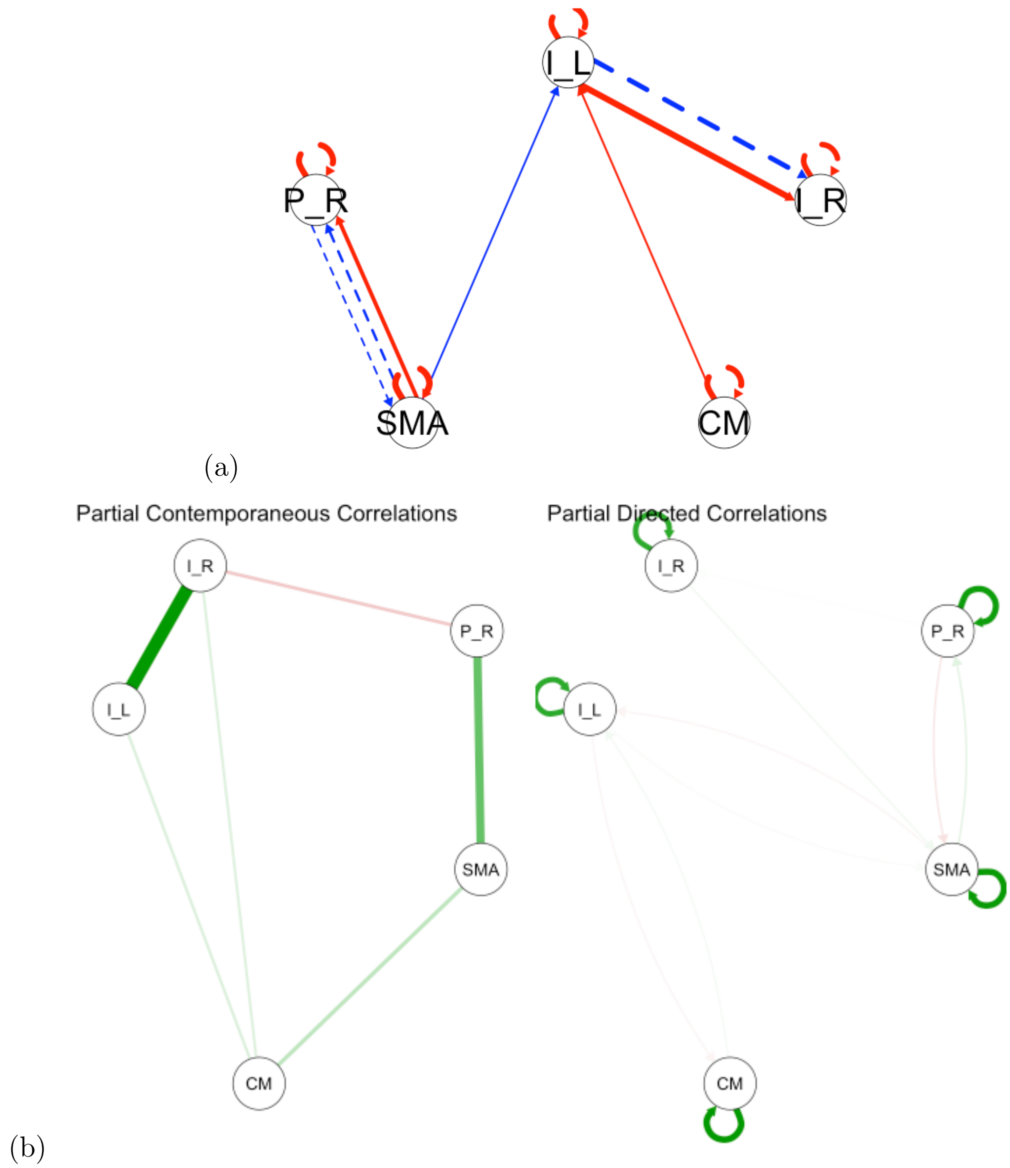
Processed brain data were obtained from the Autism Brain Imaging Data Exchange (ABIDE; Di Martino et al., 2014). We selected one individual from one of the sites that provide data to the ABIDE project. The number of observations for each individual was $T = 246$. The data used here were collected while the participant was in a resting state, meaning that they were not involved in a task and were instructed to stare at a cross-hair. We utilized data that underwent a standard preprocessing pipeline called the Configurable Pipeline for the Analysis of Connectomes which was made publicly available by Craddock et al., n.d. As a result of the preprocessing steps the data are normally distributed and weakly stationary (i.e., the mean and variance is constant) across time with cyclical trends removed. We selected time series for five brain regions: the left and right Insula (Insula-L and Insula-R), right median cingulate (CM-R), right precentral gyrus (P-R), and the left supplementary motor area (SMA-L).

We selected the left and right Insula because bilateral brain regions typically evidence contemporaneous relations across time (e.g., Beltz & Molenaar, 2016; Gates et al., 2017) and we wanted to see how these two methods may differ in the recovery and interpretation of paths that likely will exist. The CM-R is an integral part of the limbic system, and thus is likely to relate to at least one of the Insula regions as they both have been related to aspects of consciousness and higher level brain processing such as emotion processing. We would also expect that the two motor regions, the P-R and SMA-L, would relate to each other.

Exampler Result from the fMRI data Figure 2.10 shows the network models for the ABIDE study exampler from uSEM (a) and gVAR (b). There were a number of relations observed on both the contemporaneous and lagged orders in the uSEM results. To start, expected relations were found between Insula-R and Insula-L, P-R and SMA-L, and CM-R and Insula-R. On the contemporaneous level, SMA-L was a common cause for P-R (positive effect) and Insula-L (negative effect). Insula-L was also caused by CM-R as expected. As the common outcome of SMA-L and CM-R, Insula-L d-connected the backdoor between the two. That is, no relation exists between SMA-L and CM-R after considering their relations with Insula-L. Insula-L also served as a mediator, causing d-separated the paths from SMA-L to Insula-R as well as CM-R to Insula-R. In interpreting the role each brain region plays in the overall system, SMA-L and CM-R are considered exogenous variables since they are not predicted by any other brain region. P-R and Insula-L are endogeneous variables with strong exogeneity as they are caused by exogeneous variables, and Insula-R is an endogenous variable with weak exogeneity as it is caused by an endogenous variable.

At a lag order of 1, after controlling for AR-1 relations SMA-L and P-R showed bidirectional

Figure 2.10: fMRI Exampler Result from uSEM (a) versus gVAR (b)



Note: 1) IL = left Insula (Insula-L), IR = right Insula (Insula-R), CMR = right median cingulate (CM-R), PR = right precentral gyrus (P-R), SMA = left supplementary motor area (SMA-L); 2) Plots are consistent with those provided by the respective packages. For the uSEM results, red indicates positive values and blue for negative ones; for gVAR, green indicates positive values and red indicates negative. Dashed lines indicate lag-1 relations in uSEM. Dashed lines indicate lag-1 relations in uSEM. Line width correspond with coefficient estimate size for both.

and negative cross-lagged effects on each other, although the contemporaneous path was unidirectional and positive. While the direction of the lagged effect from Insula-L to Insula-R was consistent to the contemporaneous level, the sign of which had switched. In both cases the opposite sign seen at the lag-1 level from the lag-0 level suggested a dampening effect. Both of these revealed distinct structures between the two temporal orders on fMRI data. In summary, CM-R surfaced uniquely as the variable with the strongest exogeneity, suggesting it influences the system. By contrast, Insula-L, Insula-R, and P-R were endogenous and Granger Caused by other brain region activity, indicating that to some degree the variability seen in these brain regions can be explained using other brain regions included here.

The gVAR approach similarly found expected paths between Insula-R and Insula-L, and P-R and SMA-L, and the Insulae and CM-R. Another similarity is that there were bidirectional relations at the lag-1 level between P-R and SMA-L. However, here one was positive and the other was negative. The difference with uSEM result might relate to the undirected relation (thus non-distinguishable direction and sign) at the lag-0 or contemporaneous level. Despite having some similarities in the presence of relations among these nodes, the interpretation differs based on the modeling approach here. Interesting differences from the uSEM contemporaneous model were observed in the patterns of relations. For instance, the uSEM results suggested that Insula-L might play a critical role as two brain regions appeared to instantaneously Granger cause it, which in turn Granger caused another region simultaneously. From the gVAR results, this region did not seem to be very central or critical to the system in terms of processing. Additionally, while in uSEM CM-R and SMA-L were d-connected by their common outcome Insula-L, here a direct edge within the pair was produced. This could possibly be a spurious link due to the incorrect conditioning on Insula-L if it was indeed a collider. The gVAR model also returned a sparse temporal network with very weak lagged relations, this may largely be due to the fast-speed nature of brain activities and limitation from neuroimaging data. This further points to the importance of including directed relations on the lag-0 level to account for structural functional connectivity that are commonly seen (Friston et al., 2013; Smith, 2012).

2.5.3 Additional Notes

Finally, since we do not know which model is closer to the DGM, we remind the readers that our interpretation from the recovered model is based on the underlying assumption embedded behind each model. The degree of similarities and differences to which the networks disclosed by the two modeling approaches may vary from case to case. There are certain caveats in each modeling approach as a cost for the assumptions. Here we discuss some possibilities. The adoption of a undirected, non-casual contemporaneous network with strong exogeneity assumptions in gVAR models renders the risk of false positive paths in lag-1 PDC network if some contemporaneous effects were recovered as lagged relations, which has been seen when the DGM is known (Gates et al., 2010). The missing direction from lag-0 PCC has the potential to impact the recovery of both levels.

Take the EMA data example, if in reality *hopeless* is a common outcome of other variables in the model, then the associations between *down* and *guilty* or *down* and *anhedonia* we observed might very likely be a result of “endogenous selection bias" due to the conditioning on the “collider" variable *hopeless* (i.e., the bidirectional associations open up all the paths between *hopeless* and other variables). Indeed, by recovering edges encoded by partial correlations, gVAR is not robust to “endogenous selection bias" or “overcontrol bias". This is because both the following two conditions are met: 1) nodes that process undirected or bidirectional connections can function as a common cause and outcome simultaneously to each other, and 2) partial correlation entails that all nodes are being conditioned on, including those that should not (e.g., nodes on a collider path that should have remained open). We see potential evidence of this here given the large number of contemporaneous bidirectional relations.

However, it is also noted that the unidirectional assumption in uSEM poses a danger to false causal inference as well. This is analogous to regression with reverse order of a predictor and an outcome when no clear clue of temporal order is available, or a false prediction hypothesis when the true relation is pure correlation. Scenarios of relations among more than two variables could be more complicated. In the above examples, if the PCC from gVAR is more accurate, i.e. some contemporaneous associations are bidirectional rather than causal, then an arbitrary assignment of the direction associated with the connection will lead to misleading structural paths. Alternatively, issues such as “overcontrol bias" would have emerged from conditioning on variables along true causal pathway, e.g., controlling *hopeless* committed an “overcontrol bias" if *down* directly causes *guilty* and *anhedonia*. Our goal here is to offer readers precautions when applying these models to their research, no matter which model is chosen.

2.6 Discussion

Recent methodological evaluation on methods stemming from network psychometric literature has focused extensively on the statistical properties or the contrasting philosophies of the cause of relations as compared to those of traditional latent variable models. Given that very little work has discussed how interpretations differ in alignment with model specification and casual assumptions, the current work filled in a timely gap that particularly pertains to the application of the PNM to idiographic research. We investigated model inferences from fitting VAR-based model under the PNM framework (i.e., gVAR) as compared to that of the SEM framework (i.e., uSEM), in depicting time series data as networks. Specifically, we focused on how networks recovered by the two modeling approaches differ with respect to causal implications as a result of the contrasting way of representing contemporaneous relations.

We need to emphasize some major conceptual points in interpreting the directed and undirected contemporaneous networks. First, valid causal inferences can be made only from structural relations that carry casual hypotheses, such as the directed paths in lag-1 networks from gVAR or uSEM; whereas partial correlations or marginal correlations without causal assumptions do not correspond to a unique causal interpretation. It means that the undirected edges recovered from the lag-0 PCC network in gVAR lack the ability to accurately distinguish different causal relations or to reliably represent marginal independence. Second, not all GC relations can be depicted by the temporal dependencies between repeated observations and estimated as lagged relations. Particularly, the lag-1 PDC is insufficient to represent instantaneous GC relations that occur when short-term effects are missing from the temporal ordering information in the data. While the lag-0 PCC network can recover some remaining information as undirected partial correlations, we know that the information of the direction and hence the inherited causal implication is lost. Further, this conclusion extends to the recovery of two-way instantaneous GC relations, at the presence of simultaneous causality or as approximation to bidirectional cross-lagged effect unfolding over time. The “acyclicity" assumption from the DAG model needs to be removed to extend static causation to dynamic causal representations, allowing for nonrecursive structures to accommodate potential two-way instantaneous GC relations. However, when nonrecursive relations are modeled as bidirectional covariances or as undirected partial correlations, biased parameter estimates and misleading inferences might occur.

Last but not least, the adoption of an undirected contemporaneous network without instantaneous GC assumptions encodes a dynamic system with strong exogeneity. While it is less likely to commit spurious causal links, the network represents an ambiguous, uninformative, and externally dependent system. On the contrary, a fully directed contemporaneous network recovers a weakly exogeneous system that is confirmatory and informative. The model can reach a full capacity of deriving GC implications, yet is at a higher risk of committing false positive causation. When it comes to model choice, a model with causal assumptions should be employed under the theoretical guidance and supported by empirical (causal) evidence.

The two empirical examples further illustrate and complement to these points under application settings. From both the psychopathological and the neuroimaging studies, we observed similarities and differences in the networks and inferences disclosed by the two approaches. Specifically, differences from the gVAR networks and the uSEM networks were seen at the variable level regarding the relative importance of each variable, or at the network level with respect to the sparsity, the presence and the direction of relations. Particularly, the dynamic networks revealed by uSEM trace back to the sources of variation (e.g., *down* in the EMA network, or SMA-L and CM-R in the fMRI network), and point to the end (e.g., *guilty* and *anhedonia*, or P-R and I-R, respectively). It provides us with the direction of information flows as indicative of underlying causal pathways. The importance of each construct involved is described by where it stands on the pathway: *down* “causes" *hopeless*, *hopeless* “passes" on the effect, *guilty* and *anhedonia* “receives" it and “closes" it. The varying importance amongst variables is associated with differential levels of exogeneity, i.e., closer to the source of information corresponds to higher exogeneity, with the cause itself possessing the strongest exogeneity (e.g., *down*, or SMA-L and CM-R).

The gVAR dynamic networks, by contrast, are characterized by mutual interactions among nodes rather than the information flow in a direction. The centrality of nodes is revealed by the quantity of related existing links (often measured by some defined statistics) instead of the relative position along a flow. For instance, *down* or *hopeless* seem to be more central than *anhedonia* because the former two are also connected to *guilty* and both mutually relate to the latter. In fact, pathways do not exist in undirected networks, and it is meaningless to call a node of an edge the “source" or the “outcome" because they are not distinguishable. Relatedly, constructs represented by nodes are all strong exogenous variables to the dynamic networks, i.e., they relate to each other to varying degrees but are not caused by any source information from within the system. For example, *down* and *hopeless*, or P-R and SMA-L, are roughly the same central to the system. They are both mutually related to each other and to other nodes, but we cannot tell if they are the causes or the outcomes.

In terms of the degree of agreement in the interpretation pertained to the two areas of applications, it seems that inferences from the uSEM and the gVAR models share more similarities for the EMA data than for the fMRI data. The observation that gVAR is more capable to recover structural relations in the EMA study might relate to the fact that the time scale of measurements used in the EMA matches appropriately with the duration of symptomatologic effects. However, results are more divergent when using the fMRI data: while temporal structures in the gVAR model are sparse, uSEM is able to unpack more heterogeneous patterns (i.e., direction, recursivenesss, sign) within each level, together with between-level hybrid relations among the same pair of brain regions (e.g., P-R and SMA-L). It seems that gVAR model is conservative to disclose the short-term functional connectivity mapped under fast brain activities. Idiographic researchers from different domains need to pay close attention to the characteristics of their time series data when making model choice.

The current paper, to our knowledge, represent the first evaluation of a popular time series network using psychometric method on functional MRI data. In sum, we reviewed traditional causality concepts in the interpretation of individual-level directed and undirected contemporaneous models, and illustrated how person-specific inferences are altered by the model choice that can inform both psychometric and neuroscience research. We note that our work does not target at a comprehensive methodological evaluation on the underpinning statistical properties of the modeling approaches. It is out of our scope to provide a systematic model choice procedure. Instead, this pedagogical piece highlights the extent to which inferences of an idiographic dynamic network obtained by one approach differs from that of the other, without making claims as to which is absolutely superior to recover the data generating model. It implies that there is an increasing demand of an active methodological research area, e.g., the development of some types of hybrid approaches that encompass the current ones. In addition, due to limited space and time, we do not aim for providing an exhausting list of data examples in this paper, but future work can extend the evaluation to diverse contexts or perhaps with more complex data structures. Another arising and promising direction is to investigate the comparison of modeling approaches that account for heterogeneity in idiographic patterns with time series data from multiple subjects.

CHAPTER 3

STUDY II: Path and Directionality Discovery in Individual Dynamic Models: A Regularized Unified Structural Equation Modeling Approach for Hybrid Vector Autoregression

3.1 Introduction

There recently has been growing interest in the study of psychological processes at the individual level, partly due to the increasing availability of time series data in psychological and neurological science. Time series data here refers to data with many measurements over time, ideally over 60 per person. Examples include ecological momentary assessment and experience sampling methods such as daily dairy (e.g., Myin-Germeys et al., 2009; Wright et al., 2015), or neurological measurements obtained from functional neuroimaging (Price et al., 2017). One common goal is to construct person-specific dynamic assessments based on intraindividual variability (within-person variability; e.g., Bringmann et al., 2013; Bringmann et al., 2015; Ram & Gerstorf, 2009; Wigman et al., 2015). Arriving at person-specific time series models and depicting them as networks can aid clinicians in developing treatment plans for patients (Fisher & Boswell, 2016) and also help in making valid inferences in research studies. For example, identifying individual-level nuances can aid in the understanding of heterogeneity in symptom representation among clinical patients (Wright et al., 2015) and differences in cognitive processes (Nichols et al., 2014). Person-specific intraindividual variability research has been applied to widespread psychological and neurological contexts such as language learning (Yang et al., 2015), the interplay of personality disorder symptoms (Wright et al., 2015), and traumatic brain injury (Hillary et al., 2015), to name a few. In this paper, we present a novel data-driven modeling approach to circumvent the current limitations seen in some popular methods. We first extended the current unified SEM (uSEM) framework, a widely used structural VAR model, to a hybrid representation (i.e., “huSEM") that includes both undirected and directed contemporaneous effects. Next, we apply a LASSO-type regularization for a global search to arrive at the optimal sparse model. Importantly, our approach expands the analytic capabilities of intraindividual variability research by allowing for a larger search space than currently available methods. By doing so, we allow for more reliable interpretation of the resulting models.

Intraindividual variability is commonly investigated using multivariate time series methods. One framework is the Vector Autoregressive (VAR) model that assesses the dependency between multivariate repeated measures across time using linear lagged relations. However, intraindividual variability researchers face a conundrum when using the VAR models in practice. The issue is the manner with which contemporaneous effects, or relations among variables that appear to occur *instantaneously*, are considered. Such effects are common among psychological variables and often quantified after taking into account any lagged relations. Widely-used VAR specifications are restricted in that contemporaneous relations are modeled either as structural (directed) types of associations among observed variables or as residual covariances (bidirectional or undirected), but not both. However, the way in which the relations are modeled carry substantive interpretations that differ theoretically. In an effort to ensure that the resulting model will be estimable (i.e. identified), modeling approaches typically require that the researcher choose to model all contemporaneous relations as either entirely directed or undirected. It is likely that both types of contemporaneous relations are present for a given individual's dynamic process. For example, we might expect that daily stress has a direct (perhaps causal) influence on negative affect on that same day, whereas any relation between fatigue and energy levels may be caused by something external to the system of variables acquired and thus exhibited as a covariance among residuals. An appropriate model should have the flexibility to disclose mixed types of contemporaneous relations should they exist. Only allowing for one type of specification has the potential to result in either missing relations that do exist or introducing false positives when the wrong expression was chosen or in reality both exist.

This limitation has been discussed previously, with the general expression that allows for both types of contemporaneous relations provided by Lütkepohl (2005, who termed it the full VAR) and by Molenaar (2019, 2016, who termed it the hybrid VAR). The hybrid VAR approach for individual-level analyses has not been fully evaluated with simulation studies or used in any empirical study in psychology. One reason why this has not been explored is the challenge of arriving at patterns of relations that are identifiable and can be estimated. Without inducing sparsity in some way the a hybrid VAR would be saturated. Therefore, when fitting a VAR model to data in practice, one needs to arrive at a sparse, identifiable model, that is, one with a unique solution for the remaining nonzero parameters. Ambiguity exists because *a priori* theory to guide researchers through the discovery of true sparseness is often scant. As Lütkepohl (2005) describes, it is exceedingly difficult to define rules for identifiability. The question becomes: How to determine the identifiable model that best corresponds to the underlying causal structure of the data? One available option involves a step-wise searching schema whereby relations are added in a forward-search, which we will discuss in detail in later section. A drawback of this approach is that the final model is heavily influenced by relations that are added early in the search procedure.

Inspired by recent developments of statistical learning technique in psychometric models (Huang et al., 2017; Jacobucci et al., 2016; Lane, 2017), this article presents a novel modeling framework for a flexible VAR approach that estimates a sparse set of both undirected and directed contemporaneous effects. We build from the unified SEM (Gates et al., 2010; Kim et al., 2007) approach which estimates VAR models from within an SEM framework. The uSEM specification only allows for directed relations among observed variables. We expand this to include bidirectional relations among residuals in the search space (Molenaar, 2017). The approach introduced here, called regularized hybrid unified SEM (or regularized huSEM, or simply, Reg-huSEM), simultaneously performs a global search and estimation for the optimal sparse hybrid VAR model using regularization approaches. We evaluate currently popular VAR specification approaches using simulation studies - specifically, graphical VAR (Epskamp, Waldorp, et al., 2018) which allows solely for partial correlation among relations and unified SEM (Gates et al., 2010), which allows solely for directed relations among observed variables - as well a the huSEM. Certainly, our ultimate interest is to evaluate the performance of regularization under the hybrid VAR model that allows for both types of contemporaneous relations. Of greatest interest is the ability to reach an optimal model that recovers precisely the presence and directionality of only the true relations.

The remainder of the paper is structured as follows. We first provide the technical background of major VAR specifications. Next, we review and compare model discovery and estimation techniques under each VAR model. We then describe the rationale and implementation of the new model discovery and estimation framework into the current VAR specifications. This is followed by a simulation study to compare the performance of the existing and newly proposed modeling framework under conditions that mimic empirical settings. To close , we discuss interpretations of hybrid relations in VAR models, as well as the contribution of our proposal to the development of exploratory SEM research.

3.2 Technical Background

We begin with some definitions necessary for understanding VAR models. In some cases a variable measured at a given time can explain variability in a future observation of itself; this relation is a lagged relation often referred to as an “autoregressive (AR) effect". In a multivariate case, a lagged effect from one variable to another at a later measurement is referred to as a “cross-lagged effect". They are the source for observed temporal dependency among repeated measures across time. When a cross-lagged effect is significant after taking into account the AR effect it is said that Granger causality (Granger, 1969) takes place. Granger causality is defined as follows: variable X is said to *Granger cause* variable Y if Y values are better predicted with values of X (either contemporaneously or lagged) than with past values of Y alone (Geweke, 1982; Granger, 1969; Lütkepohl, 2005). In either case the direction of the effect (i.e. $X \xrightarrow{GC} Y$ or $Y \xrightarrow{GC} X$) discloses the information flow, with the possibility of both variables Granger causing each other.

In multivariate intraindividual variability approaches, Granger causal relations occurring at the same measurement is referred to as *contemporaneous* or instantananeous effects. Contemporaneous relations exist in empirical data due to subsampling. This occurs in functional MRI since data is collected at a rate far slower than the neuronal processes of interest, and similarly in daily diary studies where the data collection is at a slower rate than the speed in which constructs of interest (e.g., emotions) change. Contemporaneous relations are similar to cross-sectional (or interpersonal variability analysis in interpretation. However, contemporaneous relations in interindividual variability are usually modeled upon the assumption that observations are independent, while in intra-individual variability the relation is conditional upon the AR and cross-lagged relations. This is a key difference between cross-sectional and time series modeling â ordering of observations matters in the latter approach. Just as in inter-individual variability, contemporaneous relations in intra-individual variability can be undirected or directed. With intra-individual variability, an instantaneous directed effect is likely to be observed when a causal effect unfolds within a shorter span than the size of lag interval. We will return to the interpretation of the directed contemporaneous relations with further detail in the discussion.

3.2.1 Traditional VAR

The VAR model (Hamilton, 1994; Shumway & Stoffer, 2017) is a traditional time series technique for modeling person-specific processes using time series data. Let $Y_t = [y_{1t}, y_{2t}, ..., y_{pt}]'$ be a vector of a p-variate time series, $p > 1$, at a given time point t, with $t = 1, ..., T$. Suppose Y_t can be represented by a stationary linear time series (i.e., having a constant means and covariance function), defined by the following VAR model with lag order-m, VAR(m):

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_m Y_{t-m} + \varepsilon_t, \varepsilon_t \sim N_p(0, \Theta) \tag{3.1}$$

where the sequence of (p, p) dimensioned Φ_k matrices are the lagged coefficient matrices at order $k, k = 1, ..., m$, with diagonal elements the AR(k) regression coefficients and the off-diagonal cross-lagged coefficients. For simplicity, the intercept is omitted here. The coefficients are assumed to be time invariant. Further, there is no auto or cross-lagged dependency in the error term ε_t. i.e., the temporal covariance at lag u, $\Theta_u = cov[\varepsilon_t, \varepsilon_{t-u}'] = 0\ \forall u \in 1 \ldots \infty$. This zero-mean p-variate ε_t may be correlated contemporaneously.

The simplest form is the first-order standard VAR model, or lag-1 VAR, or simply VAR(1). VAR(1) estimates temporal relations of only consecutive measurements:

$$Y_t = \Phi_1 Y_{t-1} + \varepsilon_t, \varepsilon_t \sim N_p(0, \Theta) \quad . \tag{3.2}$$

VAR(1) model assumes a "memoryless" (i.e. Markov) process in which all previous information useful to predict the current values of Y_t is contained in Y_{t-1}. For an illustration of the new modeling framework, all the VAR models discussed in the current study are VAR(1) processes. This also follows the use of such personalized models in practice (e.g., Fisher, 2015; Wright et al., 2015). The approaches discussed here generalize to higher-order VAR models but is outside the scope of the present paper.

3.2.2 Structural VAR

As pointed out by Econometricians (e.g., Shapiro & Watson, 1988; Sims, 1981), VAR is a "reduced form" of a more general class of models referred to as the Structural VAR (SVAR; Chen et al., 2011). Instead of only identifying AR and cross-lagged coefficients, SVAR incorporates directed contemporaneous associations. A fully saturated SVAR model includes free parameters for all possible direct effects among observed variables and among residuals. However, the fully saturated model is not identifiable. In practice, restriction is imposed to reduce it to a testable model, typically with only one form of contemporaneous relations allowed and the lower diagonal of the corresponding matrix estimated (Lütkepohl, 2005). The uSEM approach represents a reformulation of a type of SVAR model where the directed contemporaneous relations occur among the observed variables. The uSEM is written as (Gates et al., 2010):

$$Y_t = AY_t + \Phi_1{}^{*}Y_{t-1} + \zeta_t, \zeta_t \sim N(0, \Psi). \tag{3.3}$$

The $p \times p$ A matrix contains the contemporaneous estimates for relations among the component series of Y_t, with independent white noise ζ_t (or "innovations") having a diagonal covariance matrix Ψ. Two coefficients, a_{ij} and a_{ji}, captures the presence and directionality of the relation between two given variables at time y_{ti} and y_{tj}, with a_{ij} representing the opposite directionality of a_{ji}. Note that estimates in the lagged relation matrix now differ: $\Phi \neq \Phi^*$

3.3 A Synthesis: the Hybrid VAR

VAR-based models estimate the directed lagged relations first and then assess the (contemporaneous) covariance among residuals as a second step. The lagged relations thus get priority as these are discovered and estimated without any possibility of conditioning on potential contemporaneous relations. Should they exist, failure to include the contemporaneous directed relations will influence the estimates obtained in VAR. Additionally, contemporaneous directed relations will often surface as lagged relations when modeled as a VAR (Gates et al., 2010), which can lead to inaccurate inferences. From an interpretation standpoint, one can consider a bi-directional instantaneous connections as a contemporaneous *non-causal mutual dependence*. In this case, the relation among two given variables is thought to occur due to the influence of something outside the system.

On the other end of the spectrum, only uni-directional contemporaneous relations are considered in the uSEM approaches. One variable can contemporaneously explain the variability in another alongside the lagged relations. In this case, the errors are assumed to have no contemporaneous relations among them. The interpretation here is that all variables in the model are endogenous and as such can be explained by other variables within the system. This might be too strong an assumption in some cases, such as when two variables are related through an latent or observed variable that are not included in the model. The interpretation of these directed contemporaneous relations is that if we have knowledge of the independent variable, we can explain a significant degree of variability in the target variable contemporaneously. We refer to such contemporaneous relations as instantaneous *causal structural path*.

In practice, researchers tend to either conduct the VAR model with bidirectional or undirected relations among residuals (e.g., graphical VAR, Epskamp & Fried, 2016) or the uSEM with directed contemporaneous relations among observed variables. The decision of which to use carries with it assumptions that relate to how the data at hand are thought to be generated and how to interpret results. As just outlined, the interpretation of contemporaneous relations in VAR and uSEM models differs. However, it is possible that the underlying relations in empirical settings are mixed, with some instantaneous associations being best captured as non-causal dependencies and other relations in the data aligning with causal structural paths. When this is the case, neither uSEM nor VAR with contemporaneous residual relations would be appropriate. Molenaar (2017, 2019) promoted the use of a *hybrid VAR* approach which combines uSEM and traditional VAR into one model. The model search space of the hybrid VAR incorporates a contemporaneous relation as either a structural relation among observed variables (as in uSEM) or a covariance among residuals (as in VAR), allowing for the flexibility in the types of contemporaneous relations. This hybrid VAR is a special case of the full SVAR. To formally introduce the hybrid VAR, we turn to the details of the modeling framework used here, which is based within the SEM framework.

3.4 Modeling Frameworks

3.4.1 Estimating VAR under the extended Gaussian Graphical Modeling

A variant of the VAR model is often performed by use of a graphical representation under the framework of Gaussian Graphical Modeling (GGM; Abegaz & Wit, 2013; Eichler, 2005; Wild et al., 2010). At its core, the GGM attempts to model the relations between variables (also called "nodes") using a sparse set of partial correlations, which is achieved by applying regularization to a given inverse covariance matrix under the assumption that the data is distributed as a multivariate normal distribution (Lauritzen, 1996). Meaningful bivariate relations are represented as nonzero edges between nodes while a lack of edge suggests independence after conditioning on all other variables in the system. Since conditional independence among psychological variables is more common than is marginal independence (Meehl, 1990), it is more likely to attain a parsimonious model by turning the estimation of marginal relations to conditional relations (Wild et al., 2010). GGM has been extended to what is called graphical VAR (gVAR; Epskamp, Waldorp, et al., 2018; Wild et al., 2010): at the lag-0 level, a weighted, undirected contemporaneous network, called *partial contemporaneous correlations* (PCC), is drawn following the general GGM fashion on the observed variables. In the meantime, the approach incorporates the temporal ordering dependency by generating a weighted, directed temporal network, termed *partial directed correlation* (PDC). This is formed at the lag-1 level of gVAR (Eichler, 2005; Wild et al., 2010).

Obtaining a sparse PCC and PDC results involves three major steps. First, the temporal relations in PDC are obtained via a standardization on ϕ elements, representing the linear lagged relationship between $y_{i,t}$ and $y_{j,t-1}$, net the linear effect of all other variables at $t-1$ that are included in the model. This comprises the VAR portion of the model. Next, a precision matrix is computed by taking the inverse of covariance matrix Θ in Equation 3.2. Partial correlations in PCC nework are obtained via a standardization on the elements of precision matrix κ derived by $PCC(Y_{i,t}, Y_{j,t}) = Cov(Y_{i,t}, Y_{j,t}|y_{-(i,j)}) = -\frac{\kappa_{i,j}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}}$, where $\kappa_{i,j}$ denotes an element in , $y_{-(i,j)}$ denotes all variables subtracting variable i and j, hence PCC represents the correlation between i and j at time t, after partialing out the linear effects of all other variables and any lagged relations. Finally, small partial correlations at both levels are first forced to zero by thresholding rules The remaining parameters are then regularized and shrink towards zero (more regularization background will be discussed in later sections).

3.4.2 Estimating SVAR under Unified Structural Equation Modeling

Structural equation modeling (SEM; Bollen, 1989) is a suitable estimation framework for SVAR models since time series estimates obtained via SEM are similar to estimates obtained using more traditional approaches such as the Kalman Filter (**Chow2010**; Hamaker et al., 2002). Conventional SEM has been extended to unified SEM modeling (uSEM; Gates et al., 2011; Kim et al., 2007) for time series data. As the name explains, uSEM "unifies" temporal ordering dependency and contemporaneous associations among observed variables. One advantage of specifying SVAR in uSEM is that it solves the issue of "nonunique solution" by using a stepwise procedure for selecting paths Molenaar (2019).

We can write Model 3.3 using conventional SEM notation (Gates et al., 2017):

$$Y = BY + \zeta, \zeta \sim N(0, \Psi) \tag{3.4}$$

where $Y = [Y_{t-1}, Y_t]$ is a $2pxT$ matrix. That is, the observations of variables are time-embedded by appending the data at $t-1$ to the data at t. The data are thus expanded to two consecutive time points $t-1$ and t, so time series vector of lagged (exogenous) variables and those of contemporaneous (endogenous) variables are appended horizontally. This requires that the ζ error vector also be extended, as well as each of the corresponding matrices. The contemporaneous and lagged coefficients collapse into a single $2p \times 2p$ regression coefficient matrix:

$$= \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots & \vdots & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots & \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \phi^*_{11} & \cdots & \cdots & \cdots & \phi^*_{1p} & 0 & a_{12} & \cdots & \cdots & a_{1p} \\ \vdots & \ddots & & & \vdots & a_{21} & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & a_{(p-1)p} \\ \phi^*_{p1} & \cdots & \cdots & \cdots & \phi^*_{pp} & a_{p1} & \cdots & \cdots & a_{p(p-1)} & 0 \end{pmatrix}_{2p\times 2p}$$

in which the upper left $p \times p$ and the upper right $p \times p$ matrix block coefficients are set to zero for lagged (exogenous) variables. That is, no variable observed at time directly predicts a variable at time-1, and variables at time-1 cannot predict each other. The lower left and right blocks are Φ^* and A matrices in Model 3.3.

The covariance matrix of the residuals in uSEM is provided in the Ψ matrix:

$$\Psi = \begin{pmatrix} \psi_{11} & & & & & & & & & \\ \psi_{21} & \ddots & & & & & & & & \\ \vdots & \ddots & \ddots & & & & & & & \\ \vdots & & \ddots & \ddots & & & & & & \\ \psi_{p1} & \cdots & \cdots & \psi_{p(p-1)} & \ddots & & & & & \\ 0 & \cdots & \cdots & \cdots & 0 & \ddots & & & & \\ \vdots & \ddots & & & \vdots & 0 & \ddots & & & \\ \vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & & \\ \vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \psi_{(2p)(2p)} \end{pmatrix}_{2p \times 2p}$$

where the upper left triangle contains freely estimated variance and covariances among the lagged (exogeneous) variables, and the lower right triangle is the variance/covariance matrix of residuals as described in Model 3.3. Note that the covariances among residuals are set to zero in the traditional uSEM framework. These matrices provide the foundation from which the expansion to hybrid-uSEM is developed.

3.4.3 From uSEM to Hybrid uSEM

uSEM is extended into hybrid uSEM (huSEM) to represent a hybrid VAR by altering the residual covariance matrix Ψ. Specifically, we relax the conditional independence assumption on the contemporaneous residuals of uSEM, and allow the errors be correlated:

$$\zeta^* \sim N(0, \Psi^*)$$

$$
\Psi^* = \begin{pmatrix}
\psi_{11} & & & & & & & & & \\
\psi_{21} & \ddots & & & & & & & & \\
\vdots & \ddots & \ddots & & & & & & & \\
\vdots & & \ddots & \ddots & & & & & & \\
\psi_{p1} & \cdots & \cdots & \psi_{p(p-1)} & \ddots & & & & & \\
0 & \cdots & \cdots & \cdots & 0 & \ddots & & & & \\
\vdots & \ddots & & & \vdots & \psi_{(p+1)(p)} & & & & \\
\vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & & \\
\vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & \\
0 & \cdots & \cdots & \cdots & 0 & \psi_{(2p)(p)} & \cdots & \cdots & \psi_{(2p)(2p-1)} & \psi_{(2p)(2p)}
\end{pmatrix}_{2p \times 2p}
$$

Note that the lower-right matrix block of Ψ^* is now a symmetric matrix with contemporaneous variances $\psi_i^*, i = p + 12p$ of the residuals on the diagonal as seen in Ψ. By substituting the off-diagonal 0's in Ψ with parameters of contemporaneous residual variance $\psi_{ij}^*, i, j = p + 12p, i \neq j$, Ψ is turned to Ψ^*. Each element in the lower right hand corner are now candidates in a model search procedure; we will not estimate all of them but provide them here for illustrative purposes. Additionally, the regression coefficient matrix B is turned to B^* as the relations and estimates may change from what they were previously, and the model-implied covariance matrix is now Σ^* derived by:

$$
\Sigma^* = (-^*)^{-1}\Psi^*(-^*)^{-1\prime} \tag{3.5}
$$

In this hybrid VAR, the relation between two given variables y_{ti} and y_{tj} has the potential to be estimated by three parameters, $_{ij}$, a_{ji}^*, and ψ_{ij}^*, with the first two representing a structure-type, directional relation, and the latter a covariance-type, bidirectional relation. In most cases they cannot exist simultaneously for identification purposes. If either $a_{ij} \neq 0$ or $a_{ji} \neq 0$, or both are true, then $\psi_{ij} = 0$ should hold.

3.4.4 A note on likelihood equivalence

Likelihood equivalent models occur when models with two different structure and/or estimates result in the same likelihood value. There are two primary sources for models to be likelihood equivalent in the context of the time series models described herein. We describe both of them here as well as rationale for how model search approaches might still be able to identify the data generating models. The first source of model equivalence would occur when we wish to identify a relation between two variables. Let's consider variables Y_i and Y_j in this thought example. If the data were generated such that $Y_i \rightarrow Y_j$, one would have a hard time recovering this relation since $Y_j \rightarrow Y_i$ would provide an equivalent likelihood - and when standardized, and the same coefficient estimated. This standardized coefficient estimates would equal the correlation between the two variables.

In the present modeling framework, this source of model equivalence is removed by the use of AR relations. Once a covariate is introduced this equivalence no longer holds. Hence, Y_i predicted by Y_j as well as itself at a prior time point will not be likelihood-equivalent to Y_j predicted by Y_i and itself at the prior time point. In equation form, we can see that estimates would be the same for each equation if the variables are standardized:

$$y_{t,i} = a_{ij} y_{t,j} + \zeta_{t,i} \tag{3.6a}$$

$$y_{t,j} = a_{ji} y_{t,i} + \zeta_{t,j}, \tag{3.6b}$$

with $a_{ij} = a_{ji}$. Once the AR relations are included the coefficient estimates would differ:

$$y_{t,i} = a_{ij1} y_{t,j} + \phi_{ii} y_{t-1,i} + \zeta_{t,i} \tag{3.7a}$$

$$y_{t,j} = a_{ji1} y_{t,i} + \phi_{jj} y_{t-1,j} + \zeta_{t,j}. \tag{3.7b}$$

Here, $a_{ij1} \neq a_{ji1}$. A similar finding results when examining the relations among residuals in that the presence of the AR relations enables for identification if whether coefficient a_{ji1}, a_{ij1}, or if the covariance of the residuals of y_i and y_j (after regressing out the AR effects) best portrays the contemporaneous relation among the two variables. They will no longer be likelihood equivalent. It is clear that the addition of the AR relations resolves this potential issue and underscores the importance of the Granger causality perspective.

The second source of equivalence is specific to the VAR, uSEM, and hybrid-uSEM models and requires more consideration. It has been shown previously that a VAR model with correlated residuals can be transformed to an equivalent uSEM model (Gates et al., 2010; Lütkepohl, 2005). It follows that hybrid-uSEMs similarly can also be transformed (Molenaar, 2017). While the transformations may be likelihood-equivalent, the coefficient estimates will not be identical. That is, the values obtained for, say, the correlation among residuals will not be the same value obtained for a standardized solution involving a contemporaneous value in the A matrix for the same two variables. We provide an illustration of this in the Appendix A for interested readers. The reader is also referred to (Gates et al., 2010; Molenaar, 2017) for more extensive coverage of this problem.

The take home is that the coefficients will be altered if a VAR model is fit to data that is generated from a huSEM or uSEM model, and similarly, for a uSEM fit to either of the other models. When the coefficient values are higher in the original data generating model than in the transform, then search procedures should be able to detect the higher (true) path. Although, the data generating model may not always provide the higher coefficient - in that case, what may look like spurious results may arise, such as a contemporaneous directed path in the generative model being found in the covariance among residuals and/or the lagged Φ matrix when a VAR is fit. It is also possible that data generated to have a covariance among residuals may surface as a directed relation in the contemporaneous A matrix.

Given the equivalent (and thus accurate) models that emerge via transformation, when evaluating results from the VAR and uSEM based approaches in our simulation study we focus mainly on detection of the presence of any contemporaneous relation between two variables. While we provide results on correct directionality recovery (i.e., directed contemporaneous, lagged, or bidirectional among residuals), this will be less of a focus due to the impossible task of true direction correct recovery for these models when data are generated under a huSEM framework. When evaluating huSEM results we consider the relation to be correctly recovered via presence for the same definition as VAR and uSEM and focus much more attention on the degree to which the approach recovered the relation in the data generating location.

3.4.5 Model Searching: Forward-selection vs. Regularization

To identify the hybrid model and estimate the covariance matrix defined by Equation 4.5, some parameters need to be constrained as a saturated model (where B^* and Ψ^* are unconstrained) will be under-identified. Exploratory procedures in search for an identifiable model, often referred to as specification search (Kaplan, 1988; MacCallum, 1986) or modeling modification (Chou & Huh, 2012), are used either when the theory to inform assumptions is lacking, or when identification or convergence issues exist due to any reason possible with time series data (e.g., insufficient data, high-dimensional variables, misspecifications) .

The exploratory search in SEM is often performed through a step-wise fashion using Lagrange multiplier score tests (Jöreskog & Sörbom, 1986, modification indices MI). Built upon this, the current most widely-adopted specification search by uSEM is a forward-selection modeling building procedure proposed by Gates et al. (2010). This search begins with a null model, and the path with the highest and significant MI is added iteratively until the model arrives at an acceptable fit as indicated by two of four standard fit indices used in SEM: the NNFI (Bentler & Bonett, 1980), CFI (Bentler, 1990), RMSEA (Steiger, 1990), and SRMR (Jöreskog & Sörbom, 1981). Readers should refer to (Gates et al., 2010) for a detailed description or Gates and Molenaar (2012) for a large-scale simulation study identifying issues with this approach, namely the lower rates correctly detecting the true directionality of the contemporaneous relations in individual-level analyses. This individual-level model building procedure is available in the free open-source R package `gimme` (Lane, Gates, Fisher, et al., 2019), where the user simply provides the raw time series data and utilize the *indSEM* function.

In considering the optimal approach for arriving at sparse huSEM models, we opted against step-wise procedures for a variety of reasons. Extant literature has suggested that step-wise model build procedures in general are subjects to problems such as finding local solutions as a final model (thus missing an optimal model) and not being robust to misspecifications (Chou & Bentler, 1990; MacCallum et al., 1992). These issues have been highlighted in the present context when conducting individual-level analysis since results may be driven by misspecifications made early in the search procedure (e.g., Gates & Molenaar, 2012). Due to the discrete searching nature, results from forward selection depend heavily on the starting model, intermediate modification or stopping criteria, and are subject to local maxima. A more effective search for the optimal sparse model should be performed on a continuum (Huang et al., 2017). It is not to say that regularization can always distinguish likelihood equivalent models, but a global, continuous model search has a better chance to achieve the optimal model (closer to truth) over other possible likelihood-equivalent alternatives which include wrong parameters and misleading interpretations. By contrast, SEM estimated via maximum likelihood is limited in the number of variables it can handle, so a backward deletion approach that starts with a full model is not an option here.

Regularization, on the other hand, will allow such a global search with all parameters undergo a continuous shrinking all at once. The term “regularization" (also referred to as “shrinkage" or “penalization") refers to the practice of imposing a penalty term to the cost functions controlling the size or the number of nonzero parameters. With a reasonable trade off of unbiasedness for estimation efficiency, regularization overcomes many modeling issues (e.g. overfitting, model complexity, under-identification), and have been widely used as variable selection and model estimation in contexts ranging from regression (Tibshirani, 1996), high-dimension reduction techniques (Zou, 2006), to graphical modeling (Friedman et al., 2008; for a broad overview, see Zou, 2006). Although historically less well received in SEM or psychometric modeling in general, recent developments of regularization method within the conventional SEM context (Huang, n.d.; Huang et al., 2017; Jacobucci et al., 2019; Jacobucci et al., 2016) has opened up the possibility for adopting regularization to uSEM (Lane, 2017) and as an extension, huSEM.

Inspired by the advantage of using regularization in reaching a sparse graphical VAR model, the purpose of the present study is to develop a similarly flexible modeling and estimation approach for hybrid uSEM (huSEM). This is necessary because for one, traditional VAR cannot offer a holistic picture of possible contemporaneous relations no matter how efficient the model search algorithm is since it does not include the possibility of directed contemporaneous relations. Two, while uSEM models allow for structural type of contemporaneous relations among observed variables, it does not allow for contemporaneous covariances among the errors and typically model searches are conducted without regularization. Below we first demonstrate the extension of uSEM to regularized uSEM (i.e., Reg-uSEM) and hybrid VAR to regularized huSEM (i.e., reg-huSEM), following the framework of regularized SEM (Jacobucci et al., 2016). In our simulation study, we evaluate if the proposed global-searching Reg-uSEM and Reg-huSEM are indeed superior to the modification index-based stepwise uSEM and huSEM search procedures, respectively. Our prediction is that Reg-huSEM will surface as the optimal model in terms of recovering the data generating relations. Our intention is to implement it under the context of huSEM not only serves to facilitate the applicability of more flexible VAR models for intraindividual variability research in practice, we also view this a new step forward for regularized SEM under time series context.

3.4.6 Model Expression and Estimation for A Sparse hybrid VAR: Reg-huSEM

The Reg-SEM approach uses regularized maximum likelihood estimation (MLE) to derive an implied covariance matrix imported to the statistical program. When fitting a Reg-huSEM , a penalty function is added to the MLE cost function of uSEM, $F_{ML}(\theta)$:

$$F_{Reg}(\theta) = F_{ML}(\theta) + \lambda P(\theta^*) \tag{3.8}$$

Where

$$F_{ML}(\theta) = \ln|\Sigma^*| + tr[\Sigma^{*-1}] - \ln|| - p \tag{3.9}$$

Where the model implied covariance Σ^* is given by Equation 4.5. Note that substitute Σ^* by Σ will reduce it to a Reg-uSEM. S is the observed covariance matrix, p is the number of model parameters. The set θ includes all the parameters estimated in the model, while θ^* is the subset containing user-specified parameters under penalization. The penalty function $P(\theta^*)$ is the general function (see below for details) to sum up the values of all the regularized parameters by a user-defined penalty method, with the level of regularization or sparsity controlled by a tuning parameter λ. λ controls the amount of shrinkage such that when λ equals zero, Equation 4.6 reduces to MLE; as λ increases to infinity, the penalized parameters are shrunken towards zero.

To estimate a Reg-uSEM or Reg-huSEM using the `regsem` function, parameters in $P(\theta^*)$ include all the regression coefficients, i.e., Φ^* and A matrices of the B^* matrix, as well as covariance of contemporaneous variables, $\psi_{ij}, i, j = p+1, , 2p, i \neq j$ elements in Ψ^* for huSEM, with the level of penalty being controlled by λ. The model does not penalize the variance parameters or covariance of lagged variables $\psi_{ij} i, j = 1, p, i \neq j$. Our particular interest is the shrinkage of the three parameters: a) a_{ij}; b) a_{ji}; and c) ψ_{ij}, that estimate contemporaneous relation between variables Y_i and Y_j. Ideally, the optimal of λ (with the lowest BIC) penalizes all unnecessary parameter(s) to zero and estimate the remaining parameter(s), unraveling the true type of relation from a), b), or c). In comparison, only regression weight parameters in B will be penalized when fitting Reg-uSEM, as there are not any free parameters for the covariances of contemporaneous variables for the contemporaneous relations between variables Y_i and Y_j. If neither, only one, or both of a_{ij} and a_{ji} remain nonzero under optimal λ, they present null, a unidirectional, or a reciprocal contemporaneous relation between y_i and y_j, respectively.

The two most common basic regularization methods are the $L2$-norm (i.e., ridge; Hoerl & Kennard, 1970), which penalizes the sum of the squared values of the parameters, and the $L1$-norm (aka. the least absolute shrinkage and selection operator, LASSO; Tibshirani, 1996), which penalizes the sum of the absolute values of the parameters. There are also various alternative forms and adaption of the two has been developing vastly, for instance, elastic net, SCAD, and MCP are widely used in practice (Hastie et al., 2015; Zou & Hastie, n.d.). Penalized parameters are shrunken as λ increases, but they can reach exactly zero only under $L1$-penalty. Hence, LASSO penalty is often used in favor of variable selection, while ridge penalty is chosen when minimizing predictor error is a priority. LASSO estimation introduces sparsity by imposing the penalty, the level of which is gauged by searching across a pre-specified range of λ values until the optimal λ (hence the sparsity level) is reached such that the model has the least mean square error (i.e., the lowest BIC).

Because of our goal to reach a sparse model in order for huSEM (and in the process, uSEM) to have a unique solution, we examine the use of two LASSO-types of penalty: the LASSO and adaptive LASSO (aLASSO; Zou, 2006). That is, $L1$-norm is the functional form for $P(\theta)$. The idea is that when one parameter is shrunk to zero, the model gains one degree of freedom. We need LASSO-penalty to identify the true zero parameters for a sparser model. It has been shown that the number of nonzero coefficients is an unbiased estimate of the degrees of freedom for the model (Zou, 2006). As the penalty increases and more selected parameters are set to zero, the degree of freedom increases and hence goodness of fit. Although the gained degree of freedom can not ensure a unique solution for the remaining (true nonzero) parameters, it eliminates the possibility of under-identification resulting from over-parameterization, and thus increases the chance of model identification and model fit. To balance the introduced bias, the aLASSO was developed so as to adjust for bias in large coefficients by imposing parameter-specific weighted $L1$-penalty. In aLASSO, each regularized parameter is scaled by the corresponding free MLE estimate such that large coefficients are less penalized. Both penalty methods are implemented in the current version of R package `regsem` (Jacobucci et al., 2016). An added benefit of `regsem` program is that it inherits model specification and syntax language from the most popular R program for SEM, `lavaan` (Rosseel, 2012).

A popular optimization method for penalized likelihood estimation is proximal gradient descent, a two-step procedure in each iteration, where step-one involves a method for calculating the step size and in step-two a soft-thresholding operator is used to overcome non-differentiability of the LASSO penalty at the origin (for more technical details on optimization, see Hastie et al., 2015;

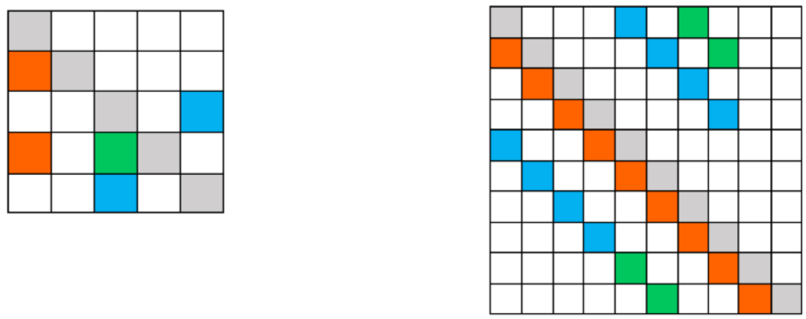
Jacobucci, 2017). Similarly, the package `graphicalVAR` employs a hybrid LASSO type of regularization to jointly estimate the temporal and contemporaneous network structures (Abegaz & Wit, 2013), extending upon the multivariate regression with the covariance estimation (MRCE) algorithm (Rothman et al., 2010). The joint regularization method involves iteratively optimizing edges of the lag-1 PDC (i.e., the temporal network) by searching across a range of λ_1), using cyclical-coordinate descent, and the lag-0 PCC (i.e., the contemporaneous network) by λ_2), using the graphic LASSO algorithm (Friedman et al., 2019), until model converges with the optimal pair of (λ_1, λ_2) that gives the least mean square error loss indicated by the lowest extended BIC.

3.5 The Simulation Study

The simulation study aims to serve two major goals: 1) to validate whether the huSEM can correctly recover the data generating relations when directed and undirected contemporaneous relations are both presented in addition to AR and cross-lagged effects among multivariate time series, regardless of the model search procedures. At the same time, we will assess to what extent the other VAR variations (i.e., uSEM and gVAR) can recover the presence of relations even though the data generation does not match the modeling approach; and 2) to evaluate whether regularization performs better than does forward selection for the uSEM and huSEM models. We test if model recovery performance differs between the adaptive LASSO and the LASSO penalty when applied to uSEM or to huSEM, respectively. Regarding the stepwise model search, emerging work suggests the stepwise approach performs well when conducted within the GIMME algorithm (Luo et al., Under Review). However, it remains unknown how it will perform within the individual-level search described above. In proposing the utilization of regularization framework, we thus additionally evaluate huSEM using the stepwise MI approach, which we term here MI-huSEM to distinguish it from the Reg-huSEM (similarlty for MI-uSEM and Reg-uSEM). Related, we are interested in assessing the consequence of model misspecification for contemporaneous relations in different limited VAR models (i.e., undirected contemporaneous relations in uSEM models and directed in gVAR models). We evaluate these modeling framework with the consideration of the impact from conditions including the amount of variables, strength of relations, as well as number of timepoints.

Data were generated using the hybrid VAR representation that is flexible enough to allow for both types of contemporaneous relations. Following a transformation of Equation 3 (Gates et al., 2010, see Appendix A, supplemental code, and), directional contemporaneous and lagged relations

Figure 3.1: Data Generating Model (the 5-variable and the 10-variable)



Notes: Orange - contemporaneous regression coefficients where $a_{ij} \neq 0$ or $a_{ji} \neq 0$, and $\psi_{ij} = 0$; Green - Cross-lagged regression coefficients where $\phi_{ij} \neq 0$ or $\phi_{ji} \neq 0$, and $\psi_{ij} = 0$; Blue - contemporaneous covariance coefficients where $\psi_{ij} \neq 0$, $a_{ij} = 0$ and $a_{ji} = 0$.

(i.e., autoregressive and cross-lagged) were incorporated as regression coefficients in B^*, while bidirectional contemporaneous connections were incorporated as marginal residual covariances in Ψ^*. We prespecified a stationary model that included a mix of different relations and chains of connections (see Figure 3.1). For example, in the 10-variable model, each variable has a directed contemporaneous relation with a neighbor variable; variable 7 has a lagged effect on variable 1, and variables 1 and 5 have contemporaneous covariances.

3.5.1 Simulation Design Conditions

We designed the simulation to evaluate the models of interest by varying the number of timepoints, number of nodes, and path strength for each of the types of conditions. Our simulations were fully crossed. The choice of these design factors was decided such that they represent data structure and characteristics of time series data in psychological and psychophysiological research.

Number of timepoints We varied the number of time points across three levels: T = 60, 200, and 1000, representing a range of potential time lengths encountered in practice. A previous simulation study had validated the use of MI-uSEM with as few as T = 60 time points (Lane, Gates, Pike, et al., 2019), which is also close to the sample size common in daily diary studies (e.g., Wright et al., 2015). Time series data derived from neuroimaging or psychophysiology studies (e.g. fMRI, EEG) typically range from 200 to 1200 time points (Smith et al., 2011). We chose 200 and 1000 to represent this range.

Number of variables We varied the number of variables in two levels: a 5-variable model and a 10-variable model. This choice corresponds to previous simulation studies emulating time series data obtained from brain studies (Smith et al., 2011). We were interested in finding to what extent the increased complexity of a larger model (more connections per variable) influences the performance of path recovery and/or the elimination of false paths. For simplicity, we obtained a moderately dense model with roughly 10% path density for both the smaller and larger models. This leads to roughly one relation per variable (as the outcome) for the 5-variable model and on average two relation per variable for the 10-variable model. The choice of a relatively moderate number of variables was made for the sake of computational feasibility. Because current regularization algorithm is limited both in the computation speed and the number of regularized variables in the model.

Strength of relationship We evaluated the extent to which the path and direction recoveries are a function of the strength of the relation. Recall that the focus of our investigation is the identification of true connections that are a mix of cross-lag, directed contemporaneous, and bidirectional contemporaneous relations. To avoid confounding effect between path type and path strength, we altered path strength by path type in the way such that in each condition paths falling to one of the two strength categories: moderate and strong [1]. For instance, in one condition all of the directed contemporaneous relations are considered strong with a uniform magnitude while those of the other two types of relations were relatively moderate (also with a uniform magnitude). As for the magnitude of path strength, previous simulation had used a mean of .40 and a standard deviation of .1, with a range limited to (.20, .60). To keep it close to these prior studies and simple, we used .7 for strong relations, and .25 for the moderate relations in the 5-variable model. Since each variable has more relations in the 10-variable condition, the strengths had to be decreased to prevent instability of the data. Here, .45 for strong relations, and .25 for moderate ones for the 10-variable model. This helped to avoid inflated variances in the endogenous variables for the larger sized model. In addition, all variables in the DGMs for either number of variables were generated to have what we consider a moderate autoregressive effect (.25). This was to mimic a typical empirical scenario

[1] Note that these are arbitrary and relative terminology used for ease in conveying the results.

where some lagged relations are expected. Note that this is a weaker AR effect size than is often used in simulation studies (Epskamp, Waldorp, et al., 2018; Gates et al., 2017).We chose this value as it may impact the accurate recovery of directions in the step-wise approach (Lane, Gates, Pike, et al., 2019; Weigard et al., Under Review) yet it was unclear the impact on regularized methods.

3.5.2 Analytic Procedures

All datasets are generated and analyzed in R (R Core Team, 2019). The R codes are available in the supplemental material. For the goals described before, each dataset was fitted by a combination of a model representation (i.e., uSEM, hybrid-uSEM, or VAR) mixed with a modeling framework (i.e., MI, regularization under SEM, or regularization under GGM). Specifically, uSEM and huSEM are evaluated by MI and regularized SEM with two penalty methods (LASSO and aLASSO). VAR (with covariances allowed among residuals) is modeled by GGM with the regularization method described above. This leads to seven models under investigation: MI-uSEM, Reg-uSEM with LASSO, Reg-uSEM with aLASSO, MI-huSEM, Reg-huSEM with LASSO, Reg-huSEM with aLASSO, and gVAR with MRCE regularization (see Table 4.1). Data fitted by MI-uSEM and MI-huSEM used the *indSEM* function under (Lane, Gates, Fisher, et al., 2019) package. The latter used the recently introduced huSEM stepwise approach (Luo et al., Under Review) available as an option in the `gimme` package (Lane, Gates, Fisher, et al., 2019). We regularized gVAR using `graphicalVAR` (Epskamp, 2018). We adopted a fully-crossed simulation design with the aforementioned factors, resulting in 2x3x3 =18 data generating conditions, each with 1,000 datasets.

Reg-uSEMs and Reg-huSEMs were conducted using the *cv_ regsem* function in the `regsem` package (Jacobucci et al., 2019). The following major three-step procedure is needed in fitting a Reg-uSEM or Reg-huSEM: 1) specify and estimate a uSEM or huSEM model in *lavaan*, where the model implied covariance Σ is imported to a lavaan function, and an (unpenalized) MLE cost function $F_{ML}(\theta)$ is optimized; 2) the lavvan object from step 1) is input to regsem function and the model structure matrices are transformed to the Reticular Action Model notation (RAM; McArdle & McDonald, 1984, McArdle, 2005); 3) perform regularized estimation on the cost function $F_{regsem}(\theta)$, with user-specified penalty method and selected parameters in $P(\theta^*)$.

For example, in our DGM where $p = 5$ or $p = 10$,

$$P(\theta^*) = ||\phi_{ij}, \beta_{kl}, \psi_{mn}||_{l1} \tag{3.10}$$

Table 3.1: Models and Estimation Approaches

Models	Model 1: gVAR	Model 2: MI-uSEM	Model 3: RL-uSEM	Model 4: RA-uSEM	Model 5: MI-huSEM	Model 6: RL-huSEM	Model 7: RA-huSEM
lag-0 (Directed/Undirected)	U	D	D	D	D/U	D/U	D/U
Est.Method (Regularization/pseudo MLE)	Reg	pMLE	Reg	Reg	pMLE	Reg	Reg
R Package	graphicalVAR	gimme	lavaan, regsem	lavaan, regsem	gimme	lavaan, regsem	lavaan, regsem
Search Algorithm	MRCE	Forward	LASSO	aLASSO	Forward	LASSO	aLASSO
Missing Data Allowed	No	Yes	No	No	Yes	No	No

Notes: aLASSO = adaptive LASSO. Model 2-4 or Model 5-7 represent the same VAR models, they only differ in model search and estimation approaches, respectively. Model names in red indicate the data generating model.

in which

$$i, j = 1, ..., p, i \neq j;$$

$$k, l = 1, ..., p, k \neq l;$$

$$m, n = p + 1, ..., 2p, m \neq n$$

For models that converged, results of the optimal model with the smallest BIC are returned. Note that in the circumstance of an under-identified model, results from the optimization of $F_{ML}(\theta)$ might be untrustworthy, but *lavaan* (used within the package *gimme*) will return MLE estimates regardless. This is also the MLE weights in aLASSO, which is a downside of using parameter-specific penalty method when the unpenalized model fails to achieve identification.

3.5.3 Outcome Measures

Because our primary goal is to compare model performance in terms of reliably recovering relations among variables with a specific direction, we used sensitivity and specificity measures for the evaluation. Sensitivity and specificity are both popular outcome measures for evaluating capability of recovery of connections in network research (e.g., Abegaz & Wit, 2013; Epskamp & Fried, 2016). Sensitivity is calculated by the ratio of the true positive count discovered in the search over the sum of all true relations in the data-generating model (i.e., true positives and false negatives). Sensitivity represents the power to detect true relationships. Specificity, in comparison, is calculated by the ratio of true negative count over the sum of negatives in the DGM (i.e., the sum of true negative count and false positive count). This represents the percentage of non-existing paths in the data generating model that the search procedure accurately omitted in the final model.

These measures allow for a global evaluation of a model's ability to detect true recovery and to reject false ones. In both sensitivity and specificity measures, higher values indicate better performance in terms of the selection of true data-generating relations. We do not assess bias or variance in parameter estimation because models different from data generating model are fitted to the data. Due to misspecification in some part of the fitted models (except the hybrid-uSEM), models have different sets of possible parameters from the DGM. The values of which are thus not comparable and the concept of bias is useful here.

Importantly, we examined sensitivity and specificity for lagged, regression and covariance contemporaneous relationships separately. Additionally, towards the special interest for the direction of contemporaneous relation, we distinguish *presence sensitivity* from *direction sensitivity*. Presence sensitivity refers to the identification for the existence of a relation between two variables regardless of the direction associated with such relation (including $X \rightarrow Y$, $X \leftarrow Y$, $X \leftrightarrow Y$, or $X - Y$). Direction sensitivity refers to the more precise recovery of not only the existence but also the correct direction ($X \rightarrow Y$ would be distinguished from $X \leftarrow Y$ for directional contemporaneous, bidirectional can be covariance in the uSEM model $X \leftrightarrow Y$ or partial correlation in the gVAR model $X - Y$).

Since it has been shown above that uSEMs can be transformed to equivalent VAR models with correlated residuals, it is critical to note that our method of assessing presence separately from direction accommodates this. We also note that presence sensitivity or direction sensitivity might suggest different forms of representation across models, due to the fact that the models differ in how contemporaneous relations are specified. For instance, since contemporaneous effects are only specified as undirected (i.e., partial correlation) in the gVAR modeling framework, the presence recovery of a true directed contemporaneous relation (e.g., $V1 \rightarrow V2$ in the DGM Figure 3.2) refers to the recovery of undirected contemporaneous relation (i.e., $V1 - V2$ in the PCC of Figure 3.3). Epskamp, Waldorp, et al. (2018) and Gates et al. (2010) both illustrated that the directed contemporaneous relations surface as an undirected contemporaneous edge in PCC plus an additional lagged edge in PDC, as correspond to the directed contemporaneous structure in the SVAR model (e.g., $V1 \rightarrow V4$ in the PDC of Figure 3.3). For this reason, we would count these lagged relations as correct presence and direction recovery (for the true path $V1 \rightarrow V4$ in Figure 3.2). Similarly, for uSEM models any covariance among the residuals (e.g., $V3 \leftrightarrow V5$ in Figure 3.2

Figure 3.2: Data Generating Model: the 5-variable huSEM

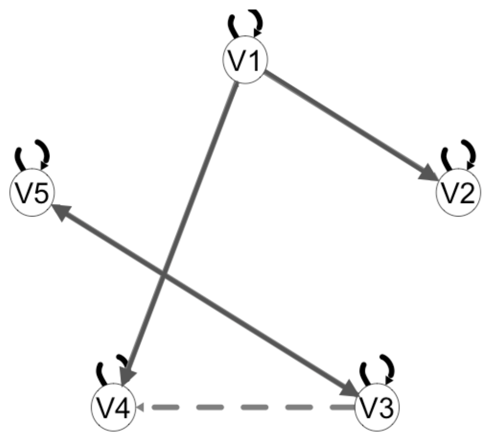


Figure 3.3: The 5-variable huSEM recovered by a gVAR model

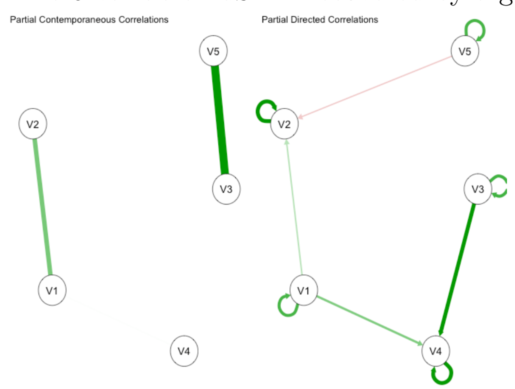
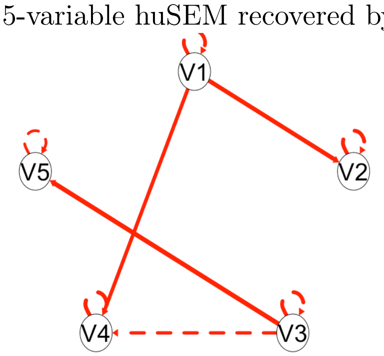


Figure 3.4: The 5-variable huSEM recovered by a uSEM model



Note: Plots are consistent with those provided by the respective packages. For the uSEM results from *gimme*, red indicates positive estimates and blue for negative. Solid lines represent lag-0 paths and dashed lines indicate lag-1 paths. For gVAR results from *graphicalVAR*, green lines indicating positive edge and red for negative. Line width correspond with coefficient estimate size for both.

will be forced into contemporaneous directed paths with an arbitrary direction (e.g., $V3 \rightarrow V5$ in Figure 3.4), and the correct "direction" of bidirectional would not have been found (Gates et al., 2010). Contemporaneous covariance recovered as directed paths are thus considered as presence (but not direction) recovery.

However, in uSEMs the recovery of the presence of a contemporaneous path includes regression coefficients of either direction (i.e., $V1 \rightarrow V2$, or $V1 \leftarrow V2$), and in huSEMs the regression coefficients and residual covariances (i.e., $V1 \rightarrow V2$, $V1 \leftarrow V2$, or $V1 \leftrightarrow V2$). The direction recovery is the rate of the true path only, i.e., $V1 \rightarrow V2$. Similarly, for the recovery of a data generating bi-directional marginal contemporaneous correlation ($V3 \leftrightarrow V5$), a relation not incorporated in uSEMs and as partial correlation in gVAR, the presence recovery include contemporaneous $V3 \rightarrow V5$ or $V3 \leftarrow V5$ for the uSEM and $V3 - V5$ for the gVAR. The corresponding direction recovery rate is zero if only one of the two directed path was recovered in uSEM, while it equals to that of the presence recovery in gVAR.

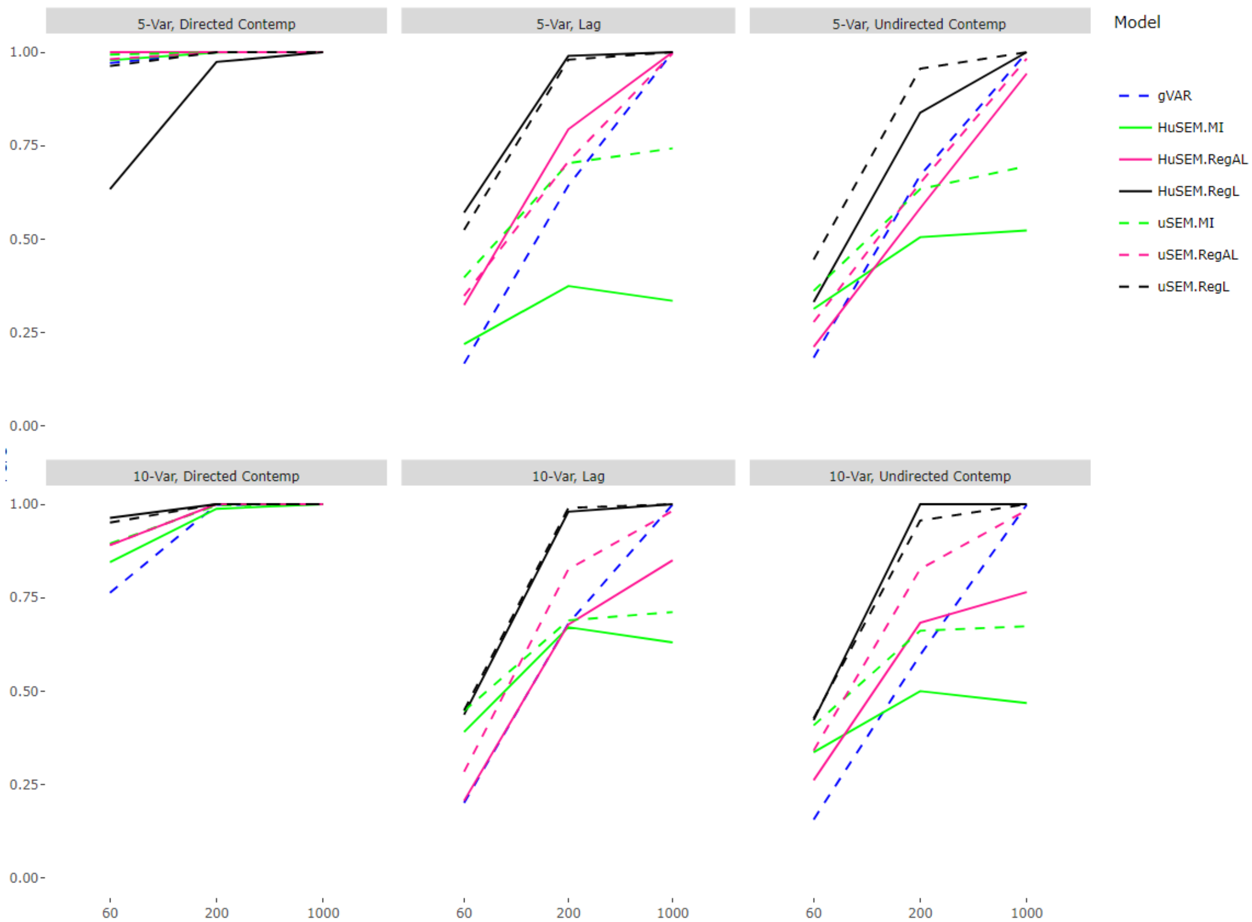
3.6 Results

As a reminder, the huSEMs discovered with MI, LASSO, or aLASSO regularization were the only models where all forms of paths were allowed in the model search space. Hence these are the only onees we expect to perform well both in terms of recovering the presence of a relation as well as the direction type. The MI-uSEM and reg-uSEM models have the same model structure as each other; they differ only in model estimation method. The gVAR model also has a different structure from huSEMs, and is estimated under GGM framework using regularization. Per our interest in the recovery of different types of relations, below we present path recovery and direction recovery for each form across the modeling approaches, respectively. Note that there are some cases where the models in *cv_ regsem* did not converge, especially for the reg-uSEM when fitted to data with 10 variables and a small to moderate sample size (see Appendix B, Table **??** for the details). For the sensitivity and specificity results, the counts were calculated using all the converged models. Model nonconvergence is a complex issue, which we will return to in the limitation section.

3.6.1 Sensitivity for Directed Contemporaneous Relations

gVAR Recovery rates for gVAR model are indicated by the blue dashed lines in Figures 3.5 to 3.9. Since directed contemporaneous relations cannot be specified according to the DGM, the presence recovery rate of contemporaneous associations was calculated as the percent of undirected edges in

Figure 3.5: Path Presence Sensitivity (Y-axis) for Directed Contemporaneous Relation by Timepoint (X-axis), Strength (column), and Dimension (row)

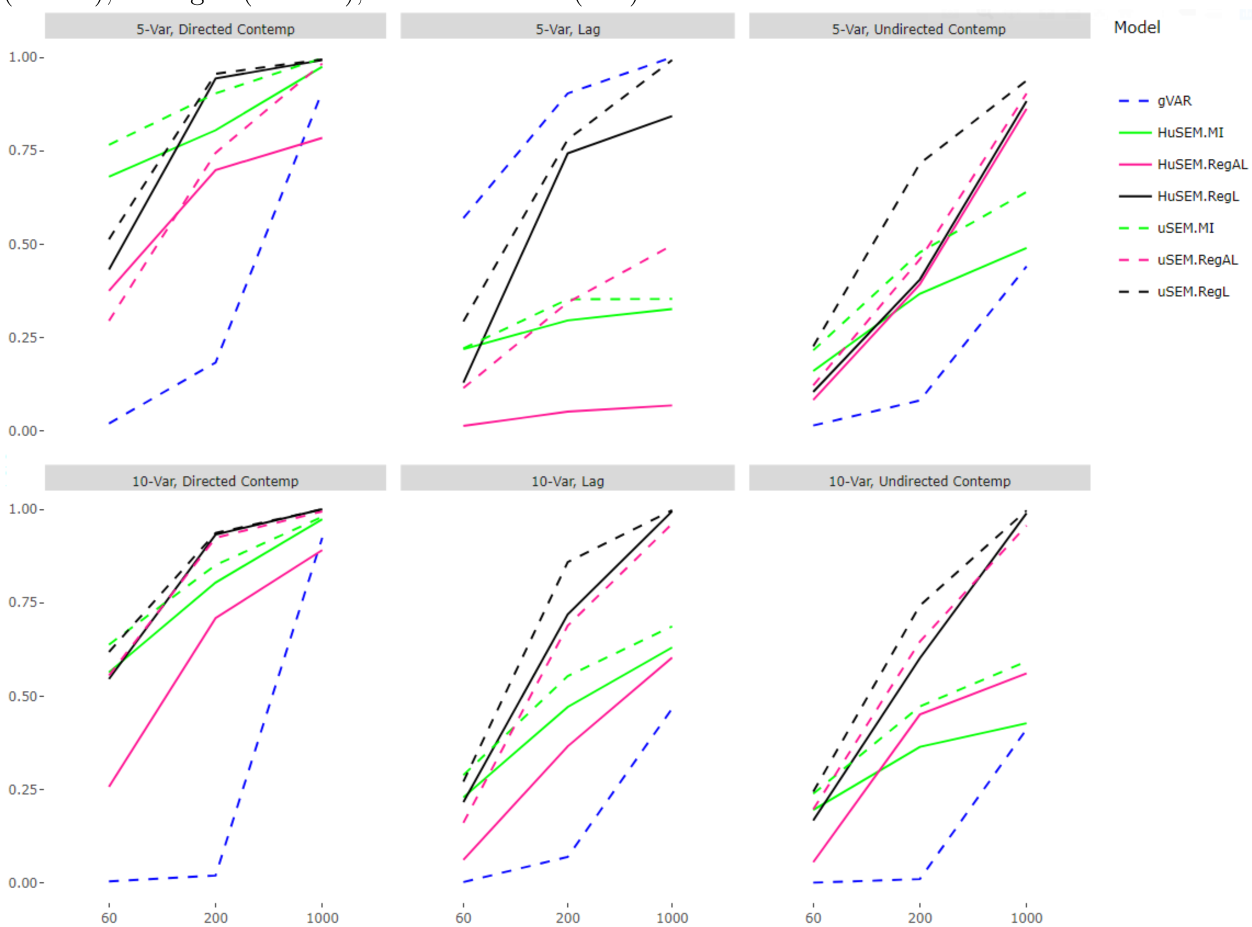


Note: 5-Var = 5-variable model, 10-Var = 10-variable model; “V5-Var, Directed Contemp" means the DGM is 5-variable model and the strongest type of relation is the directed contemporanous; x-axis represents number of timepoint (from the lowest on the left to the highest on the right); y-axis represent the % of the recovery; models are line coded in such a way that solid lines represent correctly specified models while dashed lines for alternative representations, and model search procedures are color coded.

the lag-0 PCC contemporaneous network and directed edges in the PDC lagged network across all true paths (Figure 5). As expected from the transformation of hu-SEM to VAR, the presence of directional contemporaneous relations introduced a directed edge between corresponding nodes in the lag-1 PDC network, this recovery was consistently found when such structural path was strong and the sample size was large. This confirmed that the presence and direction of a directed contemporaneous relation were recovered separately by two edges in gVAR, resulting in a less parsimonious model. While the undirected edge in lag-0 PCC only reveals the presence, the direction information was disclosed by a temporal sequential effect in lag-1 PDC. However, the performance of the direction recovery (Figure 6) was less consistent, varying across sample sizes and the relative strength of the relations. For a strong directed contemporaneous relation, the presence was nearly always recovered when N = 1,000 (left column in Figure 3.5), and direction was correctly disclosed by the lag-1 relation (see left column in Figure 3.6). When such relations were moderate, the presence (middle and right columns in Figure 3.5) was still recovered given a sufficient sample size (100% for N = 1,000, 60-70% for N = 200, and 16-20% for N = 60), but the direction (middle and right columns in Figure 3.6) was no longer recovered consistently as cross-lagged relation (50% with N = 1000, and 10% for N = 60 or 200). The tendency of gVAR to recover one directed contemporaneous connection by two edges (i.e., one being undirected and one being lagged) might very likely decrease the power to recover other true connections in the entire model and increase the chances of recovering false positive edges that provide more parsimonious models.

MI-uSEM and Reg-uSEM (with LASSO or aLASSO penalty) We similarly found that all uSEM approaches showed excellent presence sensitivity for strong directed contemporaneous relations among converged models, with recovery being highest for the MI approach for the 5 variable condition across the time length conditions. When the directed contemporaneous paths were moderate, the Reg-uSEM models (black and pink dashed lines) outperformed MI-uSEM (green dashed line) in recovering relations with moderate strength (middle and right columns) but did not perform as well as gVAR when the lagged relations were strong. The presence recovery of a strong path was excellent for MI-uSEM, and the direction recovery was near perfect (100% for V = 5 and 98% for V = 10) when T = 1000, with these values slightly decreasing for presence recovery as T decreased. Recovery of the true directionality was barely influenced by the decreasing number of

Figure 3.6: Path Direction Sensitivity (Y-axis) for Directed Contemporaneous Relation by Timepoint (X-axis), Strength (column), and Dimension (row)



time points for the strong condition. In comparison, the presence recovery of a moderate contemporaneous path was not satisfactory across conditions. In terms of Reg-uSEMs, when $T \geq 200$ both Reg-uSEM estimation approaches exhibited perfect presence sensitivity for the directed contemporaneous relation when that relation was strong and $T = 1000$. Among the two penalty methods, LASSO (black dashed line) outperformed aLASSO (pink dashed line) in both types of recovery. For example, with a medium sample size, the presence recovery rate were 96-99% for LASSO-uSEM, compared with 65-83% for aLASSO-uSEM. It was found that direction recovery depended largely on sample size, despite the strength of the relation, with poor recovery occurring for the $T = 60$ condition. For example, for a small-size small-sample model, nearly half of the presence recovery by LASSO was associated with an opposite direction, this was even worse for aLASSO. The false direction issue was immediately improved with an increasing sample size: for LASSO, it dropped to about one fourth with a medium sample size and almost eliminated when sample size hit 1000. It is surprising that even with a medium to large sample size, about 30%-50% of the presence recovery by aLASSO was associated with a wrong direction. Interestingly, direction recoveries were slightly higher overall in the 10-variable model than that of the 5-variable model overall (e.g., in the 10-variable model, false direction only emerged when sample size was small: 30-40% for T = 60, and 13-22% for T = 200). In summary, when the relation is strong, MI-uSEM has higher detection of contemporaneous relations and specificity of the direction. For smaller coefficient values, Reg-uSEM with LASSO penalty was the best choice.

MI-huSEM and Reg-huSEM (with LASSO or aLASSO penalty) Results of Reg-huSEM using LASSO penalty (black solid line) was similar to that of Reg-uSEM using LASSO penalty (black dashed line), while the other two models actually underperformed their uSEM counterparts (note that directed contemporaneous relations are correctly specified in all uSEM models). The recovery rate of MI-huSEM (green solid line) for moderate directed relations was particular low, even worse than those of MI-uSEM (green dashed line). For Reg-huSEM with aLASSO penalty (pink solid line), although the presence recovery was similar to those of its uSEM counterpart (pink dashed line), the direction recovery dropped substantially. This resulted from even more cases of presence recovery with a wrong direction for a structural contemporaneous connection. Reg-huSEM with aLASSO penalty was at the highest risk of committing this false positive, and it was more likely to occur with small-to-medium sample sizes. While these only counted for a small portion, this explained part of the bigger gap between the presence and direction sensitivity among huSEMs compared to that of uSEMs.

Taken together, we found that while all models can reliably recover the presence of a strong structural contemporaneous relation, only Reg-uSEM and Reg-huSEM with LASSO penalty can accurately recover the presence of a moderate relation along with the direction of such relation. We note that This might be an overly optimistic result for Reg-uSEM though - recall that this result must be considered in light of the fact that the result is based on the converged models while in fact a certain portion of the 10-variable Reg-uSEM models in some condition failed to reach convergence. Interestingly, the aLASSO penalty method has the tendency to introduce directed paths with a false direction, no matter in Reg-uSEM or Reg-huSEM. In comparison, the stepwise approach typically only performed well for paths that had strong coefficient sizes. MI-uSEM is not sensitive to a moderate relation. This is likely due to the algorithm favoring parsimony, and stopping the search after the strong relations were recovered and variability in the variables adequately explained. The gVAR performed well in terms of presence recovery for the moderate contemporaneous relations and did not experience convergence issues for the 5-variable condition. The gVAR approach typically did not recover directed contemporaneous relations very frequently when $T < 1000$ and in the 10 variable condition. When sample size is sufficiently large ($T = 1,000$), presence of directed contemporaneous paths were estimated as undirected edges in PCC (i.e., partial correlation), comparable to the bidirectional covariance in the uSEM or huSEM models. The direction is recovered as the corresponding lagged effect only under favorable condition and when this relation is sufficiently strong.

3.6.2 Sensitivity for Undirected or Bi-directional Contemporaneous Relation

Note: In the upper left graph, the black horizontal line represents six lines stacking on top of each other. This is because all models except the HuSEM.RegAL reached a 100 % recovery across the three timepoint conditions.

gVAR Bidirectional contemporaneous associations were (correctly) modeled as lag-0 undirected edges (Figure 3.7). The sensitivity performance of gVAR seemed to depend largely on the strength of the connection and the sample size. When this connection was the strongest of the three types of relations (left column), gVAR showed 100% sensitivity for such relation in all small-size models, and

Figure 3.7: Path Presence Sensitivity (Y-axis) for Undirected Contemporaneous Relation by Timepoint (X-axis), Strength (column), and Dimension (row)

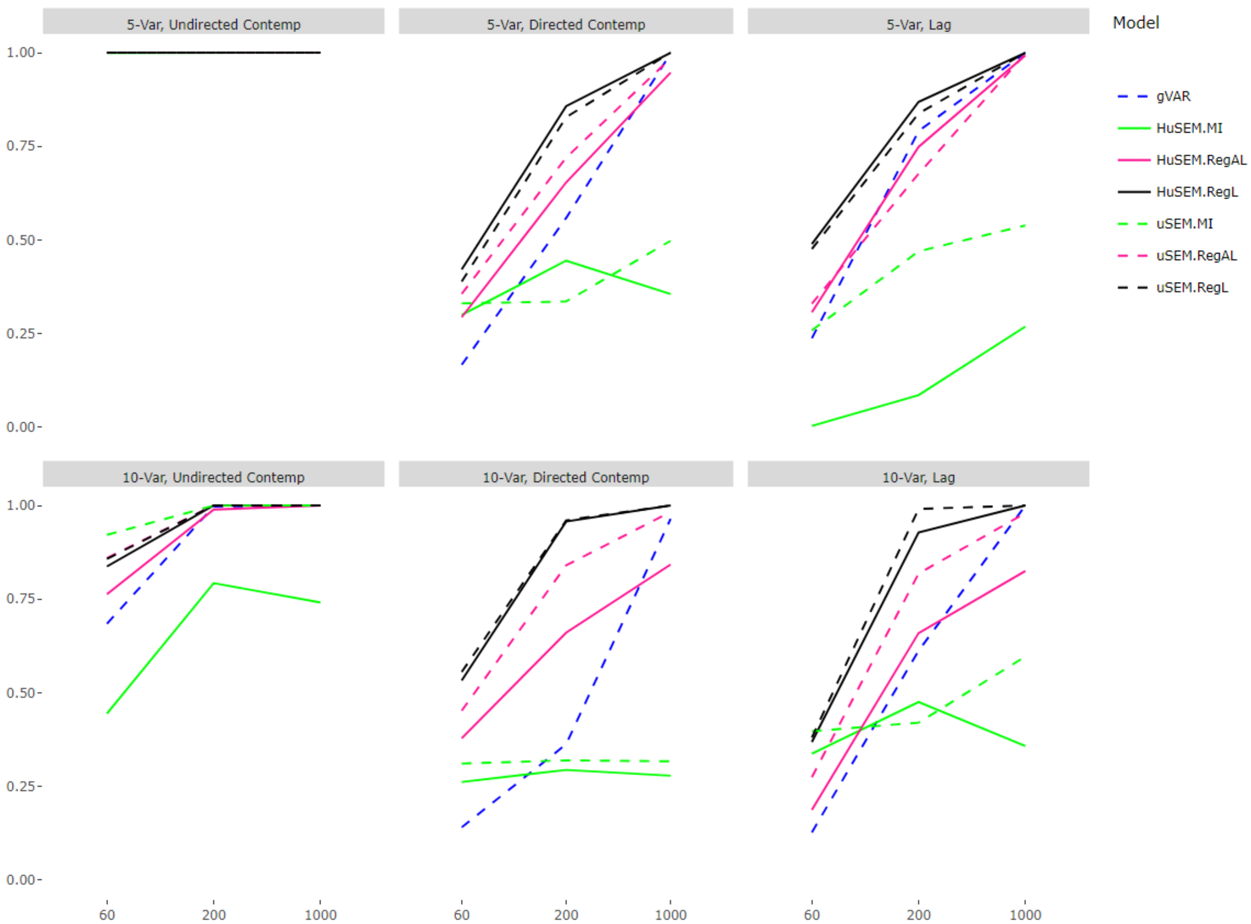
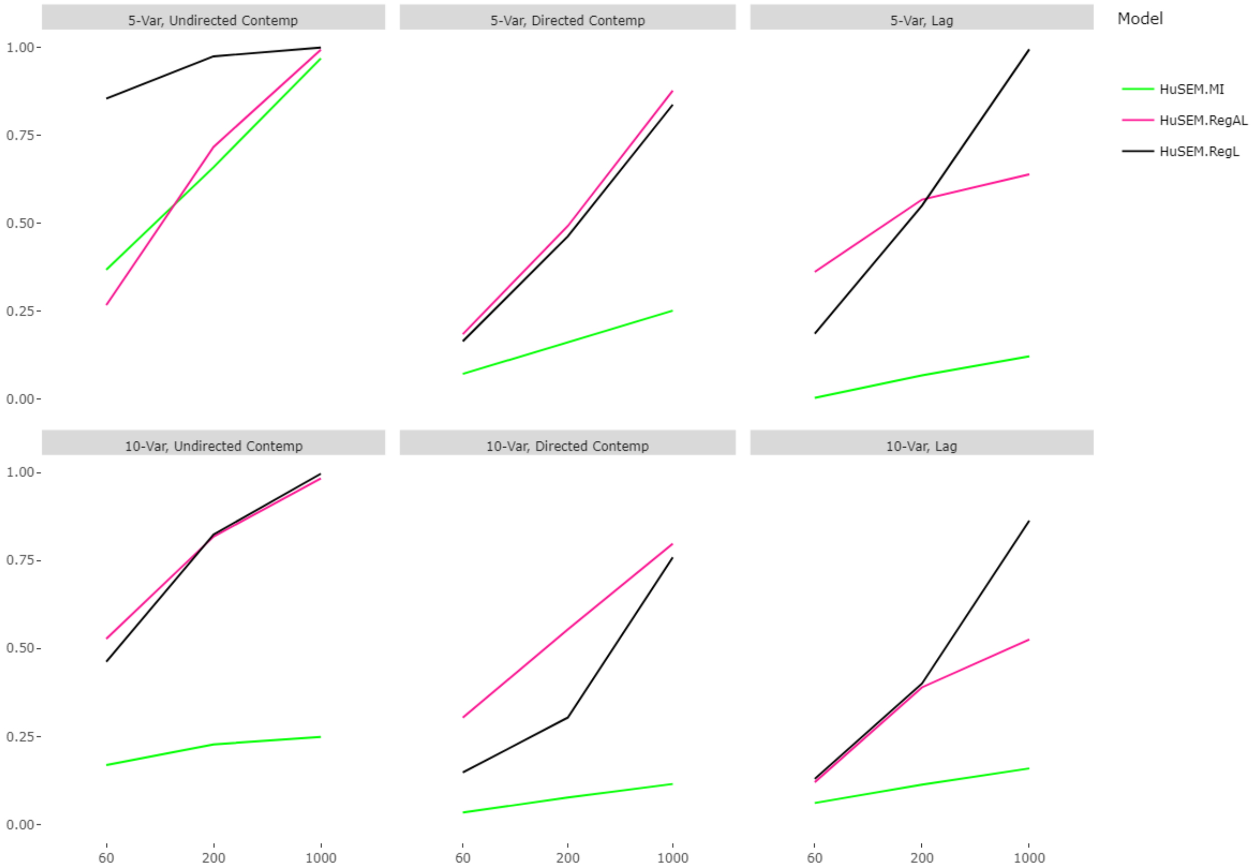


Figure 3.8: Path Direction Sensitivity (Y-axis) for Undirected Contemporaneous Relation by Timepoint (X-axis), Strength (column), and Dimension (row)



in models with 10 variables and $T = 200, 1000$ (68.5% for V = 10, T = 60). gVAR with a large sample size can also guarantee a near perfect sensitivity even when the connection is weaker than directed relations (left and middle columns), but such recovery rate was below satisfactory with a smaller sample sizes, ranging from 36.3% to 79.1%. A small sample size was not sufficient to recover consistently a moderate bidirectional contemporaneous relation (ranged from 12.8% to 23.8%). These observations suggested that a marginal covariance in VAR can be represented as a partial correlation in gVAR given enough statistical power.

Note: Results are shown for the four out of the seven models that contain undirected contemporaneous relations; in contrast, the three uSEMs only allow directed contemporaneous paths have zero direction sensitivity for undirected contemporaneous relations.

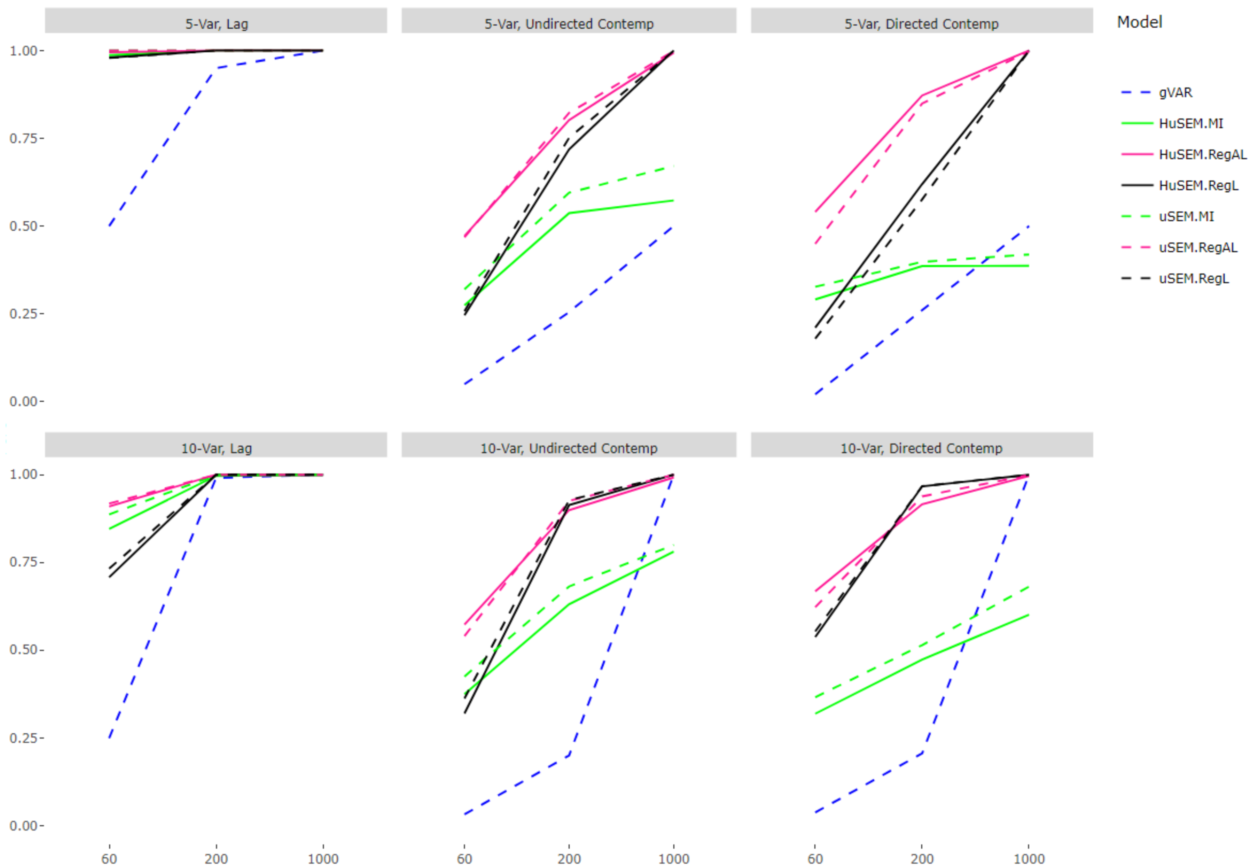
MI-uSEM and Reg-uSEM (with LASSO or aLASSO penalty) Presence recovery of undirected contemporaneous relations was counted if it was recovered as a directed contemporaneous path with either direction (nonzero beta-weight). As a result, however, the direction recovery would be zero as the chance of a false positive direction was 100%. The merit is that the presence of a strong relationship was not completely omitted, yet at the cost of a spurious causal evidence, i.e., unavoidable random direction imposed on the relation. The presence recovery for strong covariance contemporaneous relations was overall excellent in MI-uSEM, but sensitivity for moderate covariance relations dropped substantially. In comparison, Reg-uSEM with either penalty type exhibited a higher presence sensitivity for any undirected relations. LASSO slgihtly outperformed aLASSO.

MI-huSEM and Reg-huSEM (with LASSO or aLASSO penalty) Much like the previous models, reg-huSEM conducted with LASSO showed perfect presence recovery for a strong undirected contemporaneous relation. This contrasts reg-huSEM with adaptive LASSO, which greatly underperformed when compared to all other approaches. When the relation was detected with adaptive LASSO, it was likely to be the correct direction (i.e., not a directed relation but rather an bidirected one). MI-huSEM was not sensitive to moderate relations and no slight improvement over MI-uSEM was observed. In addition, although covariance contemporaneous relations were incorporated in huSEM, MI-huSEM and aLASSO-huSEM still showed a high tendency to recover them as directed contemporaneous paths when such relations were strong, just as did in the constrained uSEM models. Indeed, Reg-huSEM with LASSO penalty was superior to MI-huSEM in recovering the presence of moderate undirected relations. Taken together, Reg-huSEM with LASSO penalty is uniformly the best in the recovery of the presence of covariance contemporaneous relations. When considering directionality (i.e., was it considered a undirected/bidirected relation among residuals or a directed relation among observed variables), gVAR performed better at $T = 1000$ and generally matched the performance of Reg-huSEM. This provides evidence in favor of regularization approaches over stepwise procedures.

3.6.3 Sensitivity for Cross-lagged Relation

Cross-lagged effects are the only relations correctly specified in all the candidate models. As expected, the recovery was very satisfactory across models (Figure 3.9). Here are some observations.

Figure 3.9: Path Sensitivity (Y-axis) for Cross-lagged Relation by Timepoint (X-axis), Strength (column), and Dimension (row)



First, sample size was again the deciding factor for the sensitivity of a cross-lagged effect in gVAR with perfect recovery for strong cross-lagged effects or moderate ones given a large sample size when the number of variables is 10. Second, path strength was the deciding factor for the sensitivity performance in gVAR as well as in uSEM and huSEM models with MI estimation - these models showed a poor recovery for moderate cross-lagged effects even with a large sample size. Reg-uSEM and Reg-huSEM with either penalty method outperformed gVAR or MI-uSEM and MI-huSEM with respect to the recovery of moderate cross-lagged relations.

3.6.4 Specificity

Note that specificity in the current study refers to the ability to exclude false paths or edges that are *not* related with the true paths in the DGM. This means that the corresponding directed contemporaneous path recovered in uSEM as representative of a true undirected relation, or a PCC and a PDC edge in gVAR for a true directed relation will not be counted against specificity. Under this rule, all models except gVAR showed a specificity consistently higher than 95% (close to 100% when sample size got larger), indicating their exceptional ability to remove false relations. In comparison, specificity of gVAR model was overall lower with a larger variability, ranging from 81% to 98%. An interesting observation was that larger sample size was associated with poor specificity performance: all models with N = 1,000 had a specificity lower than 90% while those of models with N = 60 or 200 were mostly slightly above 90%. This finding that gVAR is more likely to introduce false positive edges as sample size increases may reflect that the model is attempting to accommodate the misspecification inherent in conducting a VAR on data generated with huSEMs. Recall that gVAR recovers a strong directed contemporaneous relation as both a directed lag-1 PDC edge (for direction) and an undirected lag-0 PCC edge (for presence) at the same time. The lost in the degree of freedom as a consequence potentially brings some challenge to the estimation of the entire model, an issue that is enlarged with an increased sample size. Therefore, although we found that the sensitivity performance of gVAR improved substantially as sample size gets larger, it is at the cost of decreasing specificity.

3.6.5 Supplemental Simulation

As a supplemental analysis to complete the story, we generated another two sets of data using models that represented a uSEM (with only directed contemporaneous relations) and a restricted VAR (with only undirected contemporaneous relations), respectively. The goal is to investigate how the hybrid uSEM (as a more general model) works when the true model is actually a more restricted one. The first data generating model (DGM(1)) is a five-variable uSEM model that contains the two directed contemporaneous relations in Figure 3.1, i.e., $V1 \rightarrow V2$ and $V1 \rightarrow V4$ and the lagged effect between $V3$ and $V4$. DGM(2) is a five-variable VAR that consists of the undirected relation $V3 \leftrightarrow V5$ and the lagged effect between $V3$ and $V4$. We used the norm sample size condition (i.e., $T = 200$), and the same setting for path weight factor, i.e., at each time one type of the relations is strong while the others moderate. The sensitivity result was similar to those of the main simulation (see Table **??** in Appendix C).

Briefly, we found that the presence recovery for the directed paths in DGM(1) was overall good when the relation was strong. However, performance dropped substantially when the relation was weak, gVAR performed relatively the best in this regard. Reg-huSEM with LASSO, MI-uSEM, gVAR, and MI-huSEM were most reliable in recovering the direction of a directed contemporaneous relation, though the direction recovery of a moderate relation was challenging for all models. For the recovery of undirected relations in DGM(2), it was apparent that gVAR and Reg-huSEM with LASSO can consistently recover the presence and direction of the relations, although the chance of missing a moderate one ï¼or recovering it as a directed path in Reg-huSEMï¼was still nontrivial. While the other models can recover the presence well, the uSEM models cannot recover the direction (because they are modeled as directed only) yet Reg-huSEM with aLASSO or MI-huSEM have half a chance to recover the direction of a strong relation and an even lower likelihood for that of a moderate one. These two sets of results pointed to Reg-huSEM with LASSO and gVAR as the best option when models are restricted.

3.7 Empirical Example

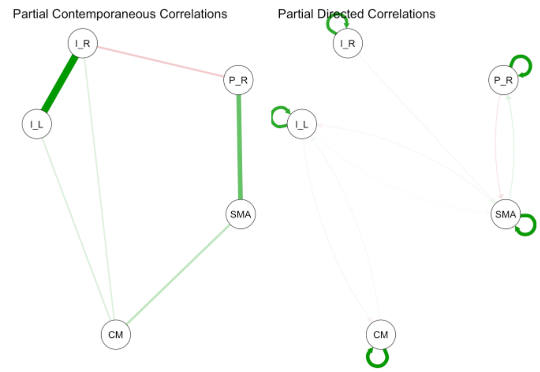
In this section, we illustrate the models using an empirical example. We fit the candidate models to a publicly available data set obtained from the Autism Brain Image Exchange (ABIDE; Craddock et al., n.d.; Di Martino et al., 2014) repository of fMRI data. Given that we do not know the ground truth or data generating models, we compare the models recovered by these approaches and discuss the different interpretations that result from each model choice. We also offer several precautions regarding potential issues arising from these model assumptions that researchers should be aware of when adopting one over the other.

For the processed brain data, we randomly picked two individuals from one of the sites that provide data to the ABIDE project. The number of observations for each individual was $T = 246$. The data used here were collected while the participants were in a resting state, meaning that they were not involved in a task and were instructed to stare at a cross-hair. We utilized data that underwent a standard preprocessing pipeline called the Configurable Pipeline for the Analysis of Connectomes (C-PAC, http://fcp-indi.github.com) which was made publicly available by Craddock et al. (n.d.). As a result of the preprocessing steps, the data are normally distributed and weakly stationary (i.e., the mean and variance are constant) across time with cyclical trends removed.

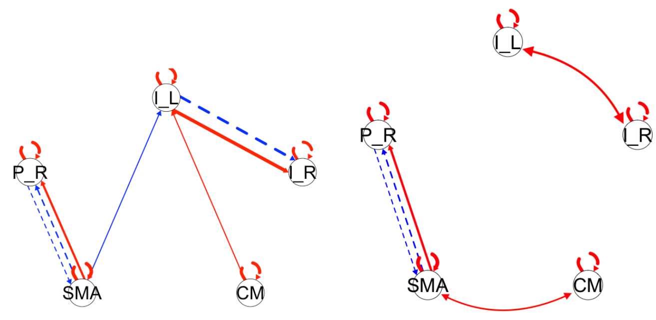
We used time series data of five brain regions from the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002): the left and right Insula (I_L and I_R), right median cingulate (CM), right precentral gyrus (P_R), and the left supplementary motor area (SMA). We selected the left and right Insula because bilateral brain regions typically evidence contemporaneous relations across time (e.g., Beltz & Molenaar, 2016; Gates et al., 2017) and we wanted to see how these methods may differ in the recovery and interpretation of paths that likely will exist. The CM is an integral part of the limbic system, and thus is likely to relate to at least one of the two Insula regions as they both have been shown to be related to aspects of consciousness and higher level brain processing such as emotion processing. We would also expect that the two motor regions, the P_R and SMA, would relate to each other. We standardized the variables to account for unequal variances that are common in fMRI data. We fit five models to the 246 x 5 standardized time series data of the two individuals respectively: gVAR, MI-uSEM, MI-huSEM, Reg-uSEM, and Reg-huSEM. For the Reg-uSEM and Reg-huSEM, we used LASSO penalty. The aLASSO penalty was abandoned for the issue we observed from the simulation study (R codes are available as supplemental materials). Figure 3.10 shows the result of the five models for person one, and Figure 3.11 for person two.

For person one, expected associations were found within the two bilateral brain regions (i.e., the left and right Insula) as well as within the motor regions (i.e., P-R and SMA) by all methods, but results differ with respect to the direction of the relation at the contemporaneous (i.e., lag-0) level or at the lag-1 level. The relations across bilateral, limbic, and motor regions also differed. For example, gVAR recovered a sparsest lag-1 network, all the five regions were contemporaneously connected at least by a second-order (i.e., through a third variable). This suggests an absence of Granger Causality from the network disclosed by gVAR. This was compared to a two-way lagged

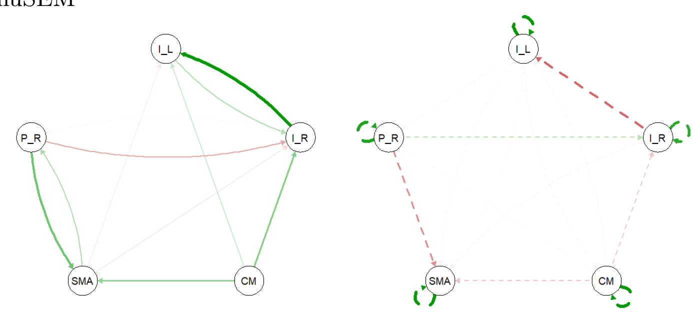
Figure 3.10: Result of fMRI Study Person One



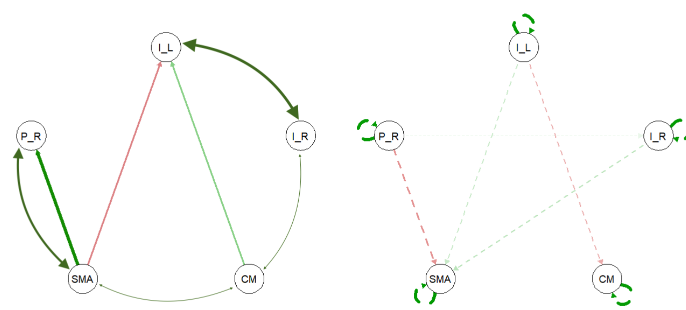
(a) gVAR



(b) MI-uSEM vs. MI-huSEM



(c) Reg-uSEM



(d) Reg-huSEM

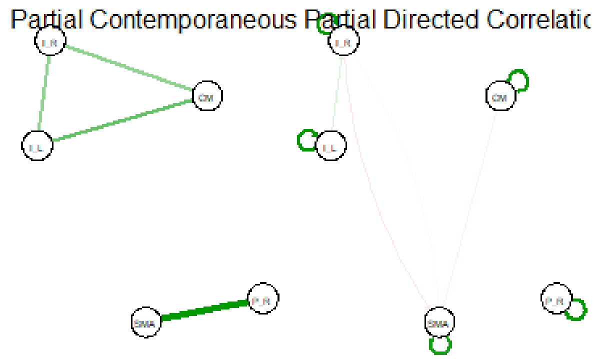
Note: 1) IL = left Insula, IR = right Insula, CM = right median cingulate, PR = right precentral gyrus, SMA = left supplementary motor area; 2) **regsem** does not provide a plotting function, we used **qgraph** to plot the result. Plot consists of solid lines on the left represents contemporaneous relations with two-headed curves represent covariance relations, plot with dotted lines on the right represents lagged effects. Green lines represent positive values, and red represent negative. Line width corresponds with coefficient estimate size. Plots only present > .01 relations, similar to the convention in programs like *graphicalVAR*.

relations (indicating a mutual Granger Causality) between the two motor regions (P-R and SMA) revealed in MI-uSEM and MI-huSEM, or a one-way relation (P-R Granger caused SM) in Reg-uSEM and Reg-huSEM, plus an additional instantaneous reciprocal effect in Reg-uSEM or a undirected association in Reg-huSEM.

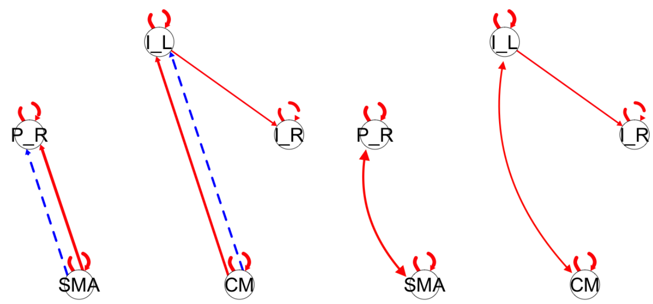
An interesting pattern showed up for the relation between the two bilateral regions across these models: while an unidirectional lagged effect was found in MI-uSEM and Reg-uSEM with an additional association on the contemporaneous level, this relation was estimated as undirected contamporaneous covariance in gVAR, MI-huSEM, and Reg-huSEM results. These together show that only the huSEM models have the capacity to recover both the Granger causality and common-cause relations that might exist simultaneously in neuroimaging data. Within the two huSEM models, MI-huSEM was sparser than Reg-huSEM, a pattern consistent to the simulation results in which we found models with MI searching procedure tend to be conservative in recovering moderate relations. Differences were also observed on the variable level in terms of the centrality or the importance of each variable in the network. For instance, the MI-uSEM and Reg-huSEM models pointed to Insula-L and SMA as central regions that showed more connections and also played a more critical role in linking (or Granger causing) the two brain clusters. From the gVAR results, Insula-L and Insula-R are equally central to the system in terms of processing, while in MI-huSEM the two clusters were by and large isolated.

Similar to person one, the overall result of person two revealed two “clusters" of these brain regions, i.e., the bilateral regions and the limbic system (CM), plus the motor regions. For person two, the two clusters were more isolated by most models. Once again, gVAR recovered a network consisted of contemporaneous connections rather than lagged relations. The MI-uSEM and Reg-uSEM further recovered the direction of these relations on both the lag-0 and lag-1 levels. This provided evidence for a Granger causality between SMA and P_R or CM and I_L. Additionally, a contemporaneous mediation path from CM to I_R via I_L was recovered in MI-uSEM, while a lagged one from CM to I_L via I_R was seen in Reg-uSEM. In contrast, the MI-huSEM and Reg-huSEM suggested a network with hybrid relations: while the contemporaneous relations between the pair SMA and P-R or that of CM and I_L were both undirected covariance, a directed, Granger causal relation was recovered from I_L to I_R. Once again, the hybrid models represent a network of relations that is more conservative than that of the fully directed uSEM models yet more

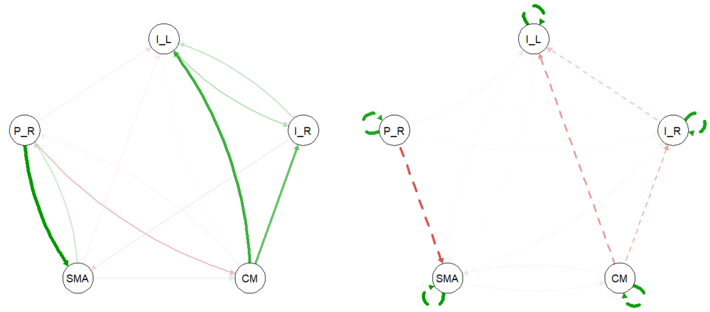
Figure 3.11: Result of fMRI Study Person One



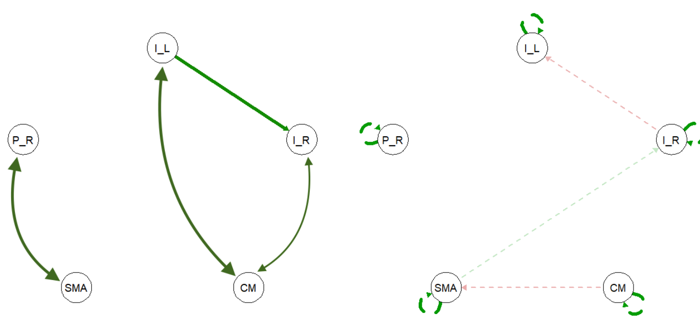
(a) gVAR



(b) MI-uSEM vs. MI-huSEM



(c) Reg-uSEM



(d) Reg-huSEM

Note: 1) IL = left Insula, IR = right Insula, CM = right median cingulate, PR = right precentral gyrus, SMA = left supplementary motor area; 2) **regsem** does not provide a plotting function, we used **qgraph** to plot the result. Plot consists of solid lines on the left represents contemporaneous relations with two-headed curves represent covariance relations, plot with dotted lines on the right represents lagged effects. Green lines represent positive values, and red represent negative. Line width corresponds with coefficient estimate size. Plots only present > .01 relations, similar to the convention in programs like *graphicalVAR*.

deterministic and informative than that of a fully undirected (on the lag-0 level) gVAR model. The difference in the same pair of variables across these two participants suggests interindividual differences in their intraindividual brain networks, which further points to the need of choosing a flexible modeling framework when the ground truth is unknown in empirical settings.

We can extend this to larger network or experiment with more individual data, but these two examples demonstrate some interesting patterns already. From these results, we found that the uSEM models recovered a brain map with more frequent and directed connections, while the network recovered by gVAR is sparse in the lagged network and less informative at the level of undirected relations. The fact that gVAR model returned a sparse temporal network may largely be due to the fast-speed nature of brain activities and limitation from neuroimaging data, which further points out the importance of including directed relations on the lag-0 level to account for structural functional connectivity that are commonly seen (Friston et al., 2013; Smith, 2012). The two hybrid uSEM models represent a picture that is somewhere in between, i.e., some relations might be association by nature, while others might be causal or even mutually causing each other. Sometimes such mutual or reciprocal effects even happen at a different rate with the possibility of having a combination of signs at each level, e.g., one may have an instant positive effect on the other while the effect the other way around is negative and takes place at a slightly slower pace as a lagged effect.

The empirical example further illustrates and complements to our findings under application settings. From the neuroimaging studies, we observed similarities and differences in the networks and inferences disclosed by the these approaches. Specifically, differences from the gVAR, uSEM, and huSEM models were seen at the variable level regarding the relative importance of each variable, or at the network level with respect to the sparsity, the presence and the direction of relations and hence the interpretation in the perspective of Granger causality. Particularly, the dynamic networks revealed by uSEM or huSEM trace back to the sources of variation (e.g., SMA or CM), and point to the end (e.g., P-R or I-R, respectively). These provide us with the direction of information flows as indicative of underlying causal pathways. The importance of each construct involved is described by where it stands on the pathway. The gVAR dynamic networks, by contrast, are characterized by mutual interactions among nodes rather than the information flow in a direction. The centrality of nodes is revealed by the quantity of related existing links (often measured by some defined statistics) instead of the relative position along a flow. In fact, pathways do not exist in undirected networks, and it is meaningless to call a node of an edge the “source" or the “outcome" because they are not distinguishable. Relatedly, constructs represented by nodes relate to each other to varying degrees but are not caused by any source information from within the system. For example, P-R and SMA, are roughly the same central to the system. They are both mutually related to each other and to other nodes, but we cannot tell if they are the causes or the outcomes.

3.8 Discussion

In studying person-specific dynamic models using time series data, it is possible to observe the coexistence of instantaneous causal pathways and simultaneous connections beyond lagged effect. Researchers interested in intraindividual variability should consider modeling the intraindividual variability using a general VAR representation where hybrid forms of contemporaneous relations can be estimated (e.g., Molenaar & Lo, 2016). To overcome the limitation in the current practice for the identification and estimation of this hybrid VAR approach, we introduced a novel modeling framework, *the regularized hybrid unified SEM (Reg-huSEM)*. The hybrid uSEM, as an extension of uSEM that had been widely used to estimate SVAR, is flexible in specifying a hybrid VAR with contemporaneous relations entering the model either as directed regression weight or as undirected covariance. We utilized regularization under the hybrid uSEM framework in place of MI step-wise model search algorithm previously used by uSEM (Gates et al., 2010) and huSEM (Molenaar, 2019). Regularization provides a global, continuous model search and parameter estimation for the optimal sparse model. Our simulation validated that the sparse hybrid uSEM model obtained by regularization with LASSO penalty was the closest to the true model among all model specification and estimation methods. That is, the presence and direction of true paths of hybrid forms were recovered with the highest accuracy.This study of a person-specific model with both directed and bidirectional contemporaneous relations estimated under regularized SEM is, as far as we are aware, the first of its kind.

Two issues were addressed by our simulation study. First, our results confirmed that the hybrid uSEM is indeed superior to uSEM with respect to avoiding false directed contemporaneous relations or graphical VAR in terms of accurately specifying the direction of structural contemporaneous relations. This conveys the message that a hybrid uSEM, rather than any restricted form of VAR, should be the starting model to fit time series data when the data generating model could potentially contain both types of relations. Specifically, graphical VAR does not consider the direction of a given contemporaneous structural relation and uSEM tends to impose a random (half chance to be false) direction to a strong covariance relation regardless of the modeling searching algorithm. It becomes even more challenging for gVAR and uSEM models in recovering and specifying the correct direction of a moderate relation. Second, we found that regularization using LASSO outperformed the stepwise approach.

One may wonder why the choice of different representations of VAR models is important, given that they can be analytically transformed. The specification of directed contemporaneous relation is indeed important because while some of the VAR representations are likelihood equivalent (see Appendix A for our illustration on the model equivalency, see also Molenaar & Lo, 2016), they differ in their interpretations. Recall that a statistical correlation or covariance represents an existence of connection whose source is exogenous to the model, and barely informs the direction of information flow. By comparison, the property of Granger causality that offers the predictability of ongoing activity in one part from that in another, is often more attractive to researchers. The concept of Granger causality is very useful to interpret directed functional connectivity when applied to neuroimaging and neurophysiological time series data such as EEG or MEG signals (Friston et al., 2013). For example, Barrett et al. (2012) used EEG to show that anaesthetic loss of consciousness Granger caused a reliable increase in cingulate cortices, extending previous “phase synchrony" obtained using undirected models (Murphy et al., 2011).

We stress that a contemporaneous directed relation is indicative of a Granger causal effect that is failed to be captured by temporal dependency due to inappropriate time scaling, which often occurs in studies like fMRI where the underlying neuronal brain activity happens much faster (on the scale of milliseconds) than what is captured via the MRI (typically on the order of seconds) and also in ecological momentary assessment studies. Recall the observation of a very sparse PCC recovered by gVAR for both individuals in our empirical fMRI example. When this is the case, an instantaneous Granger causality, if allowed, can occur when the time lag of the effect of interest is so short that it approaches infinitely to zero. Indeed, mutual causation in a dynamic system can be interpreted as a limit form of the lagged effect schemes where the time lag is reduced towards zero (Strotz & Wold, 1960). Taking these into consideration, the inclusion of contemporaneous directed relations is essential to the completeness of potential causality.

We emphasize some major conceptual points in interpreting the directed and undirected contemporaneous networks. First, valid causal inferences can be made only from structural relations that carry casual hypotheses, such as lagged paths; whereas partial correlations or marginal correlations without causal assumptions do not correspond to a unique causal interpretation. This means that the undirected edges recovered from the lag-0 PCC network in gVAR lack the ability to accurately distinguish different causal relations or to reliably represent marginal independence. Second, not all Granger causalities can be depicted by the temporal dependencies between repeated observations and estimated as lagged relations. Particularly, the lag-1 PDC is insufficient to represent those that happen instantaneously (shorter than the measurement window), it is likely that short-term effects are missing from the temporal ordering information in the data. While the lag-0 PCC network can recover some remaining information as undirected partial correlations, the information about the direction (and hence the causal implication) is lost. This is when directed relations are necessary at the lag-0 level to capture the potential Granger causality relations.

Another essential contribution of our study is the development of a model search regime effective for the hybrid uSEM. Our results showed clearly that for the model search approach, regularization consistently outperformed forward-selection step-wise search with respect to reliably and accurately identified moderate relations in both the uSEM and the hybrid uSEM models. MI seems to only indicate the presence and direction of strong relations; however, the inclusion or deletion of moderate connections are unlikely to cause a significant change in MI. In terms of the type of penalty method, LASSO outperformed adaptive LASSO especially for the direction recovery of structural relations. It is worth noting that although gVAR also adopted regularization to search for a sparse model, it suffers from low sensitivity performance when number of observation is small yet decreasing specificity performance as number of observation increases. This comes down to the point that a good estimation method cannot make up for unnecessary constraints in the model specification. Altogether, the LASSO-regularized hybrid uSEM model is the best combination of modeling and estimation framework for intraindividual variability analysis on multivariate time series data.

Indeed, the adoption of regularization filled a crucial methodological gap in the recovery of the uSEM family of models (Lane, 2017). As the true sparse model is almost always unknown, an exploratory search for the optimal hybrid VAR model without a *priori* theory is unavoidable in empirical research. The model selection process for hybrid uSEM involves not only a search for the (global) best fit in the likelihood, but perhaps also one with the Granger causality structure closest to the truth. An inappropriate model choice or a non-global solution leads to misleading causal inference. This poses a particular challenge as the search might involve statistically isomorphic models over the parameter space, i.e., likelihood-equivalent yet Granger causality-distinguishable models. The choice of recovering a contemporaneous relation between a directed and a undirected (or bidirectional) brings additional challenge to the model selection. Although the forward-selection method for uSEM model has been shown effective when applied to fMRI studies (e.g., Gates et al., 2016), it is not without issues as are all step-wise approaches. The lack of sensitivity to moderate relations found in our simulation study is not negligible. The adoption of the regularized SEM method offers hope to overcome the identification and estimation challenge in hybrid uSEM. Our simulation results validated that LASSO regularization is more effective to recover both the presence and the direction of small to moderate dynamic relations than is hybrid uSEM with MI search.

3.8.1 Limitation and Future Direction

There were several limitations in the present study. First, our simulation design is modest in the random simulation factors to keep a manageable scope and provide proof of concept. For example, we are fully aware that the size of network chosen is likely to be smaller than what one might encounter in time series studies. For instance, brain network studies using fMRI data is often high dimensional and might involve a network as large as 300-400 nodes (Smith et al., 2011). The choice of a small amount of variables is by and large limited by the computational speed and feasibility. The current optimization algorithm is limited both in the computation speed and the number of regularized variables in the model.

Second, a few assumptions are made to the data generating model as a preliminary investigation. For one, data was generated using a weakly stationary multi-normal time series, suggesting constant estimates for relations with a fixed structure. In reality, relationships might undergo some change through the passage of time. The findings should be generalized with caution, given a lack of randomness and the adoption of a simple model structure (e.g., between every pair of variables can only be an contemporaneous or a temporal connection). For two, consistent with the candidate models, a lag-1 model is assumed, higher-order lagged effect was not accounted for. Additionally, missing data was not investigated. Although missing data is usually handled by full information maximum likelihood under the assumption of missingness at random in SEM literature

(Enders & Bandalos, 2001), neither regsem or `graphicalVAR` programs at the time of the study has the mechanism for handling missing data.

As it does in practice, nonconvergence became another issue to consider in our study. The fact that not all the models with simulated data converged raises issues about how to best report the results. Following a common practice, we eliminated all replications that did not converge and our success rate was based on using only replications that converged. If we counted nonconverged replications as failures, then the success rates would be lower, particularly for Reg-uSEM models under those conditions when nonconvergence was relatively frequent.

There are two methodological extensions for future directions. One immediate next step is to account for between-person similarities in the model search process. The present study focused on individual-level model (one individual per model). We did not consider group-level information. Prior work has shown improvements in model search when using shared information as starting model (e.g.,Gates & Molenaar, 2012; Varoquaux & Craddock, 2013; Luo et al., Under Review). Another advantage of this extension is that it reduces noise due to non-meaningfull intra-individual variability, which in turn would reduce the estimation of spurious relations between variables.

The other major methodological improvement is related to the package development, such as the regularization method, optimization algorithm, and etc. We are aware that it is very likely that there has already been some improvement in the program or some other newer methods that could be implemented after our study was completed. In fact, just upon completion of the current paper, we learnt that a new package called `psychonetrics` (Epskamp, 2020b) was released, which includes a function to run gVAR with potentially better model searching algorithms and an improved performance. However, it is out of time and space for the current paper to add more comparisons. Additionally, the current study did not focus on expanding the search for other algorithms or penalty methods for Reg-uSEM and Reg-huSEM beyond the built-in options in the `regsem` package. Within the package itself, there is room for improvement. For example, we believe it would be beneficial to have dual or hybrid penalization in the function for different types of parameters or those of different scales. In our case, it would be to penalize the regression type of parameters in a separate term from that of the covariance type of parameters, each with a corresponding tuning parameter. This dual penalty feature is now available for the *regsem* function by a collaborative work of the authors and the package team. But it is still under developed for an automatic grid search over a two-dimension tuning parameter space for an optimal combination such as the current fashion in function *cv_regsem* for a uniform penalty, which was used in our simulation. Once this feature is updated, the current study can be extended as well. The other issue to be solved for `regsem` is the adoption of MLE weights for adaptive LASSO when the model from `lavvan` is under-identified. For one thing, the low direction sensitivity for structural contemporaneous relations by aLASSO penalty is somewhat surprising. We suspect that this is related with the scaling of each regularized parameter using weights from corresponding MLE estimates. The problem is that these MLE estimates were obtained from an initial unpenalized, under-identified uSEM or huSEM model that are not reliable estimates of the generative process. Outside of psychometric, research has examined solutions such as using marginal regression estimators as the initial estimators for the alasso weights (Huang et al., 2008). But such advancement has not been available in the current package. We do intend to extend our work and generalize it to more realistic conditions in our future research with the development of more efficient and reliable regularization algorithm.

The present study, to our knowledge, is the first extension of regularization to the hybrid uSEM. The success of the application points to a promising future in exploratory SEM in this context. Just as the general trend of statistics outside the field of psychometric modeling where statistical learning methods have becoming dominate variable selection or model selection method, they might as well replace step-wise methods to be the primary model search methods for SEM in general. The major challenge, however, remains to be computational. As pointed in the recent work by Pruttiakaravanich and Songsiri (2018), the question lies in finding a powerful optimization algorithm for a model as complex as SEM. Solving the fitting function in SEM often involves nonconvex and nonlinear optimization, meaning that unique and estimable solution can only be reached under very constrained scenarios (Pruttiakaravanich & Songsiri, 2018). But with the rapid ongoing development in contemporary computational machine learning, we are confident for the future advancement in this area.

CHAPTER 4

STUDY III: Path and Directionality Discovery in Individual Dynamic Factor Models: A Regularized Hybrid Unified Structural Equation Modeling Approach with Latent Variables

4.1 Introduction

The previous chapters discussed psychometric modeling approaches for fitting individual dynamic models to time series data (TSD). Particularly, two limitations were discussed in the contemporary approaches: one is a restriction in the vector autoreggressive (VAR) models in the way that contemporaneous relations are either represented only as directed among observed variables or only as undirected relations among errors, the other is the step-wise forward model building method that is subject to local maximum solutions (Ye et al., 2021). To overcome these limitations, the author and colleagues (2021) proposed a novel modeling approach called the regularized hybrid unified SEM (or huSEM), where LASSO regularization is used to perform model selection and estimation for a more flexible VAR representation that allows both forms of contemporaneous relations. The simulation study in Ye et al. (2021) showed that regularized huSEM performed uniformly the best over alternative VAR representations and/or modeling approaches, with respect to accurately recovering the presence and directionality of hybrid relations and reliably removing false relations when the data are generated to have two types of contemporaneous relations.

However, so far the examinations have focused on VAR models with only a small number of manifest variables without accounting for measurement error. In practice, it is common that more than one, sometimes many, indicators are used to measure the same underlying dynamic constructure. When multiple indicators are available, latent time series variables could be formed to adjust for measurement error and to reduce the dimensions of the observed variables. The combination of the factor model and the time series model results in what we call the dynamic factor model (DFM; Browne & Nesselroade, 2005; Molenaar, 1985). In DFM, dynamic relations (e.g., lagged and contemporaneous relations) are allowed either within the factor series or amongst the factor and the observed time series.

In fact, current dynamic modeling approaches include a factor model within their current VAR version to estimate a DFM. For example, the unified SEM (uSEM) model in the GIMME framework (Gates & Molenaar, 2012) has been extended to the uSEM with latent variables (i.e., LV-uSEM), for which they call the Latent Variable GIMME (or LV-GIMME; Gates et al., 2019). LV-GIMME estimates a DFM as a LV-uSEM using the stepwise model building algorithm, and an option to estimate parameters by use of either the pseudo-maximum likelihood (i.e., pseudo-ML[1]; Molenaar & Nesselroade, 1998) or the model-implied instrumental variables with two-stage least squares (MIIV-2SLS; Fisher et al., 2019). In addition, the graphical VAR (gVAR) model has been combined with a factor model to form what is called the latent variable gVAR (or LV-gVAR; Epskamp, 2020a) model under the general Gaussian Graphical Model framework. The primary purpose in the last chapter of this dissertation is to extend the regularized huSEM in Ye et al. (2021) to the regularized huSEM with latent variables to estimate a sparse DFM that allows for hybrid contemporaneous dynamic relations between the latent factors themselves.

Three steps address this overarching goal: the first is to reform the structural model of the latent variable uSEM (LV-uSEM) to its hybrid uSEM version, which I will refer to as the latent variable hybrid uSEM (or LV-huSEM); the second is to perform model selection using the LASSO regularization in the search for the optimal sparse LV-huSEM; lastly, the final model, i.e., a sparse LV-huSEM, will be estimated via MIIV-2SLS to obtain final parameter estimates. To evaluate the proposed method with existing ones, a simulation study will be conducted to compare 1) model recovery performance of the LASSO regularization versus the pseudo-ML based stepwise model build when they are applied under the LV-huSEM context; and 2) biasedness and robustness of parameter estimates obtained by LASSO regularization, pseudo-ML, as well as the MIIV-2SLS estimation.

Below, I first introduce the basic concepts and specification of DFM, followed by a review of the LV-GIMME framework to estimate an individual DFM as the LV-uSEM. Next, I point out the issues and limitations of the current approach, and how the proposed method addresses these issues.

[1]The author notes that the term “pseudo-ML" in the current work specifically refers to the application of using maximum likelihood estimation for time series data where the independence assumption is violated due to temporal dependence between repeated observations, as used in time series literature such as Molenaar & Nesselroade (1998).

This leads to the design and conduct of a simulation study to evaluate and compare the proposed method with the existing ones. Finally, conclusions and discussion are drawn from the results of the simulation study.

4.1.1 Dynamic Factor Models

Dynamic Factor Models (DFMs) represent a class of models by including lagged relations within the latent variable approach (Browne & Nesselroade, 2005; Molenaar, 1985). It can also be seen as a factor analysis extension to the family of VAR models in the sense that latent variables (or, factor series) are defined in a measurement model and that lagged relations can be incorporated either in the measurement model, in the structural model, or both (Molenaar, 1985). Indeed, DFM is a synthesis of factor analysis and VAR. Such a synthesis is ideal for many psychological studies that aim to unravel hybrid relations in unobserved dynamic processes. A substantive question that can be investigated by a DFM approach is to what extent an increase in a latent dynamic construct (e.g., anxiety) predicts changes in another latent dynamic construct (e.g., depression) as well as changes in the indicators of the other latent construct. As another example, neuroscientists often aim to study functional connectivity in human brains. The latent constructs in this context could be some unknown "brain networks" formed by a cluster of disparate brain regions that tend to interact across time when performing a task (Gates et al., 2019). Integrating a factor model component in the dynamic model opens up the possibility to explore dynamics among latent constructs underlying what we could observe.

A general DFM for a single multivariate TSD is defined by two components, the measurement model and the structural or latent variable model (Molenaar, 1985; Zhang et al., 2008). Let $Y_t = [y_{1t}, y_{2t}, ..., y_{pt}]^T$ denote a vector of a p-variate time series at a given time point t, with $t = 1, ..., T$. Assuming Y_t represents a weakly stationary linear time series (i.e., with a constant mean, variance and covariance function). To ease the presentation, it is assumed that all the time series have zero mean function (i.e., no intercept term):

$$Y_t = \sum_{u=0} \Lambda(u)\eta(t-u) + \epsilon_t, \epsilon_t \sim N(0, \Theta). \tag{4.1}$$

$$\eta_t = \sum_{u=1} \Phi(u)\eta(t-u) + \zeta_t, \zeta_t \sim N(0, \Psi). \tag{4.2}$$

In the measurement model, the $\eta(t-u)$ is a q-variant set of latent factor series, with the (p, q)-dimensional time-invariant factor loading matrix $\Lambda(u), u = 0, 1, ..., l$ that denotes the linear relations between the original p-variant time series Y_t and the q-variant factor series η_t at the lag order of u. The ϵ_t is a p-variate measurement error process for the p-variate observed variables in Y_t. Like in cross-sectional factor analysis, usually local independency among the indicators (conditional on the corresponding latent factor) and no cross-loading is assumed. The structural equation in the DFM is a dynamic process of VARMA(m, n), i.e., a VAR of order n with a MA of order m. The $\Phi(u), u = 1, 2, ..., n$ is a sequence of (q, q)-dimensional matrices of AR and cross-lagged effect of the latent factors at the lag order of u. Assumptions in Model (3.2) applies to Model (4.2), and that the parameters are time invariant. In addition, we assume that the errors ϵ_t and ζ_t are uncorrelated with their respective latent variables $\eta(t-u)$, the errors are uncorrelated over time, i.e., $cov(\zeta_t, \zeta_{t-1}) = 0$. Here, moving average (MA) coefficient matrices (i.e., current latent variable values predicted by errors from the prediction of previous latent variable values) are not not considered $(m = 0)$. In this way, the structural equation in the DFM is a dynamic process of VAR(n), i.e., a VAR of order n. Taken together with the measurement model, this returns a DFM (p, q, m, n, l). The general DFM implies that lagged values of the latent variable can have loadings on future values of the indicators beyond the indirect effect via the factor at that concurrent time point.

For model identification and substantive purposes, however, restriction is often imposed to allow for only one type of the lagged relations. For example, a general DFM is reduced to a simpler, more restrictive version containing only lag relations among the factors, i.e., DFM(p,q,0,1,0), called the first-order lag-1 process factor analysis or direct autoregressive factor models (DAFM; Browne & Nesselroade, 2005); alternatively, it reduces to what is called the shock factor analysis or white noise factor model when lagged relations are only in the measurement model, i.e., DFM(p,q,0,0,1); McArdle, 1982; Zhang & Browne, 2007).

4.1.2 Latent Variable Group Iterative Multiple Model Estimation

Latent Variable Group Iterative Multiple Model Estimation (LV-GIMME; Gates et al., 2019) adopts the DAFM model or the DFM(p,q,0,1,0), i.e., with lagged relations only between latent factors in the structural or latent variable model. The DAFM makes the interpretation and the implementation of LV-uSEM simpler, because the dynamic relations are between factors, which can be separate from the measurement model where factors are extracted from indicators at the same time. The dynamic relations between the latent variables in the LV-uSEM are comparable to those among the observed variables in the uSEM model, with the exception that these dynamic relations have accounted for measurement errors in the observed variables. Because the aim of this dissertation is to compare approaches of individual dynamic models with hybrid relations, I focus on individual case time series, that is, no group-level modeling or aggregation of individual models will be performed. In the remaining part of this chapter, I refer to this model as LV-uSEM instead of the LV-GIMME from the paper that applies to multiple subjects TSD.

In the measurement model, LV-uSEM adopts the general form of DAFM, i.e., a first-order processing factor series, or DFM(p,q,0,1,0). This part adopts a confirmatory factor approach in order to obtain latent construct with the same qualitative meaning. For each individual:

$$Y_t = \Lambda\eta_t + \epsilon_t, \epsilon_t \sim N(0, \Theta). \tag{4.3}$$

In the structural model, LV-uSEM inherits a uSEM structure (Gates et al., 2011; Kim et al., 2007) that “unifies" temporal ordering dependency and contemporaneous associations among the latent factors. The following model specification and matrix notations follow Gates et al. (2017) and is also consistent with the previous paper Ye et al. (2021).

$$\eta = B\eta + \zeta, \zeta \sim N(0, \Psi). \tag{4.4}$$

where $\eta = [\eta_{t-1}, \eta_t]$ is a $2qxT$ matrix. That is, the factor variables are time-embedded by appending the data at $t-1$ to the data at t. The data are thus expanded to two consecutive time points $t-1$ and t, so time series vector of lagged (exogenous) factor variables and those of contemporaneous (endogenous) factor variables are appended horizontally. This requires that the error vector ζ also be extended, as well as each of the corresponding matrices.

The contemporaneous and lagged regression coefficients collapse into a single $2p \times 2p$ B regression coefficient matrix:

$$
B = \begin{pmatrix}
0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \ddots & & & \vdots & \vdots & \ddots & & & \vdots \\
\vdots & & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\
\vdots & & & \ddots & \vdots & \vdots & & & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
\phi_{11} & \cdots & \cdots & \cdots & \phi_{1p} & 0 & a_{12} & \cdots & \cdots & a_{1p} \\
\vdots & \ddots & & & \vdots & a_{21} & \ddots & \ddots & & \vdots \\
\vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & a_{(p-1)p} \\
\phi_{p1} & \cdots & \cdots & \cdots & \phi_{pp} & a_{p1} & \cdots & \cdots & a_{p(p-1)} & 0
\end{pmatrix}_{2p \times 2p}
$$

in which the upper left $p \times p$ and the upper right $p \times p$ matrix block coefficients are set to zero for lagged (exogenous) factor variables. That is, no factor variable at time t directly predicts a factor variable at time $t-1$, and factor variables at time $t-1$ cannot predict each other. The lower left and right blocks are Φ parameters for lag-1 relations and A for contemporaneous factor variables at time, respectively.

The error covariance Ψ is a diagonal matrix representing an independent white noises:

$$
\Psi = \begin{pmatrix}
\psi_{11} & & & & & & & & & \\
\psi_{21} & \ddots & & & & & & & & \\
\vdots & \ddots & \ddots & & & & & & & \\
\vdots & & \ddots & \ddots & & & & & & \\
\psi_{p1} & \cdots & \cdots & \psi_{p(p-1)} & \ddots & & & & & \\
0 & \cdots & \cdots & \cdots & 0 & \ddots & & & & \\
\vdots & \ddots & & & \vdots & 0 & \ddots & & & \\
\vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & & \\
\vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & \\
0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \psi_{(2p)(2p)}
\end{pmatrix}_{2p \times 2p}
$$

where the upper left triangle contains freely estimated variance and covariances among the lagged (exogeneous) factor variables, and the lower right triangle is the variance and covariance matrix of factor residuals. Note that the covariances among factor residuals are set to zero as in the traditional uSEM framework. These matrices provide the foundation from which the expansion to latent variable hybrid uSEM is developed.

A note on the level of heterogeneity across individuals. It is possible that individual vary both in the measurement of constructs as well as the relations among latent factors. But the interpretation of the dynamic relations is hard to generalize across people if the latent factors stand for completely different constructs. For this reason, a typical practice in GIMME framework is to have a qualitatively homogeneous and quantitatively heterogeneous measurement models (i.e., same nonzero pattern yet individual-specific estimates in Λ matrix) in order to ensure the same substantive meaning of the latent factor with some levels of individual variability (Gates et al., 2019). To maximize heterogeneity in the dynamic patterns, the structural models represent entirely idiographic dynamic models (i.e., person-specific patterns and estimates). Note that this setting is opposite to the practice in idiographic filter DFM (Molenaar, 1985; Nesselroade et al., 2007), where the measurement model structure and factor loading are allowed to vary across individuals on the condition that the structural relations among the latent variables remain invariant across people. This is critical since previous literature has pointed to great variability across individuals both in how constructs are measured and how they relate to each other (Epskamp, Waldorp, et al., 2018; Gates et al., 2019; Hamaker et al., 2005). GIMME represents a very flexible approach that maximizes heterogeneity in dynamic patterns across individual DFMs.

The pseudo-ML model selection and the MIIV-2SLS estimation. The LV-uSEM operates under the SEM framework for DFM that uses a pseudo-ML based model build algorithm and the model-implied instrumental variables with two-stage least squares (MIIV-2SLS; Bollen, 1996) for parameter estimation. Ideally, the specification for the structural model 4.3 should be guided by a priori theory. Unfortunately, very little is known about the individual dynamic pattern, besides that research has shown that the patterns vary largely across people (Nichols et al., 2014; Wright et al., 2015, e.g.,). GIMME uses a data-driven forward selection algorithm where for every individual, it starts with a null model, and one path with the highest and significant modification indices is added iteratively until the model arrives at an acceptable fit (Gates et al., 2010). This model building procedure is automatic in the free open-source R package `gimme` (Lane, Gates, Fisher, et al., 2019).

The recommended estimation approach is a sequential, three-step procedure: in the first stage, measurement model is estimated using MIIV-2SLS (default), or pseudo-ML, or PCA, and latent scores of factor series is derived through the regression method; the second stage is consistent to the general GIMME approach which involves a pseudo-ML based stepwise search to identify the sparse individual structural models (i.e., uSEMs) with individualized patterns of dynamic relations amongst the latent factor series obtained from the previous step; finally, parameter estimates of the measurement and the sparse structural model are obtained via MIIV-2SLS. MIIV-2SLS has been shown to be more robust to model misspecification than the system wise ML estimations for traditional SEM (Bollen et al., 2007) and particularly for SEM under DFM (Fisher et al., 2019).

The use of equation-by-equation estimation in MIIV-2SLS is particularly advantageous because it separates the estimation of the measurement model from the structural model so that the measurement model parameter estimates remain unaffected by heterogeneous structures in the structural model (Gates et al., 2019). Enabling individual variability in the pattern of contemporaneous or dynamic relations among latent constructs allows for a better understanding of individual-centered processes as they unfold over time. However, MIIV-2SLS is not a model selection tool, hence we apply the MIIV-2SLS estimation of the final sparse LV-uSEM after the stepwise model selection procedure.

Technical details of the MIIV-2SLS are provided in a few articles (e.g., Bollen, 1996; Bollen et al., 2021; Bollen et al., 2007). Here I give a brief idea of the MIIV-2SLS. Unlike system-wise full-information estimators such as maximum likelihood that estimates all parameters simultaneously, the MIIV-2SLS estimates one equation at a time. To start, the MIIV-2SLS transforms the structural model into one that contains only the observed variables (i.e., L2O) by substituting each latent variable with its scaling indicator minus its error as specified in the measurement model (Bollen et al., 2021). One assumption for any OLS estimator is that the composite error cannot correlates with one or more of the covariates of that equation, which is typically violated after the L2O transformation. A common way to solve the correlated error issue is to use instrumental variables, i.e., variables that are uncorrelated with the composite error but correlate with the covariates that are associated with these errors. The special advantage of the MIIV-2SLS is that qualified instruments were drawn from other equation(s) within the system itself based on the model structure, hence it is called model-implied instrumental variables (or MIIVs; Bollen et al., 2021). Finding qualified MIIVs is done automatically using the algorithm from Bollen

(1996) and is implemented in the R package *MIIVsem* (Fisher et al., 2020). Another advantage of the equation-by-equation MIIV-2SLS estimation is that, besides the χ^2 tests of goodness-of-fit for the whole model, equation-wise overidentification test (e.g., Sargan's χ^2 test; Sargan, 1958) is available when there exists more than the minimum number of MIIVs for that equation. A rejection to the null hypothesis of the overidentification test suggests that at least one of the MIIVs fail to meet the conditions based on the current model specification, which is an evidence that some specifications of the model is in error (Bollen et al., 2021).

To estimate a LV-uSEM using MIIV-2SLS estimator under the GIMME approach, one convenience is that the latent factor variables are obtained prior to the estimation of the structural model and are treated as observed variables in the model selection and estimation for structural relations. In addition, the lagged latent factor variables are predetermined (exogenous) variables because no backward predictions from factor scores at time T to those at the previous time are allowed (recall the zero elements in the B matrix in Equation 4.4). This setting of the lagged factor variables guaranteed a minimum set of qualified MIIVs for the parameter estimates in the structural model.

4.2 The Current Study

The uSEM framework is used to estimate individual dynamic models with latent factors such as the DAFM. However, there are a couple of areas that I propose to extend the individual modeling under the GIMME framework. First, it is imperative to move from the restrictive VAR representation in the DAFM from a uSEM to the more flexible hybrid uSEM. That is, directed regressions and undirected error covariances among contemporaneous latent factor variables should be incorporated simultaneously. Because not only can they co-exist, they can carry different causal interpretations as well as practical implications. Therefore, the first goal is to extend the structural model in LV-uSEM to the hybrid representation, i.e., LV-huSEM, by altering the residual covariance matrix Ψ in Equation 4.4 to Ψ^* following Ye et al. (2021):

$$
\Psi^* = \begin{pmatrix}
\psi_{11} & & & & & & & & & \\
\psi_{21} & \ddots & & & & & & & & \\
\vdots & \ddots & \ddots & & & & & & & \\
\vdots & & \ddots & \ddots & & & & & & \\
\psi_{p1} & \cdots & \cdots & \psi_{p(p-1)} & \ddots & & & & & \\
0 & \cdots & \cdots & \cdots & 0 & \ddots & & & & \\
\vdots & \ddots & & & \vdots & \psi_{(p+1)(p)} & & & & \\
\vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & & \\
\vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & \\
0 & \cdots & \cdots & \cdots & 0 & \psi_{(2p)(p)} & \cdots & \cdots & \psi_{(2p)(2p-1)} & \psi_{(2p)(2p)}
\end{pmatrix}_{2p \times 2p}
$$

Note that the lower-right matrix block of Ψ^* is now a symmetric matrix with contemporaneous variances $\psi_i^*, i = p + 12p$ of the factor residuals on the diagonal as seen in Ψ (cited from Ye et al. 2021). By substituting the off-diagonal 0's in Ψ with parameters of contemporaneous residual covariance $\psi_{ij}^*, i, j = p + 12p, i \neq j$, Ψ is turned to Ψ^*. Each element in the lower right hand corner are now candidates in a model search procedure. By doing so, we relax the conditional independence assumption on the contemporaneous errors of uSEM, and allow the errors to be correlated. Additionally, the regression coefficient matrix B in Equation 4.4 is turned to B^* as the relations and estimates may change from what they were previously, and the model-implied covariance matrix is now Σ^* derived by:

$$
\Sigma^* = (I - B^*)^{-1} \Psi^* (I - B^*)^{-1\prime} \tag{4.5}
$$

Second, as discussed in Ye et al. (2021), the forward selection method of model building is highly dependent on the starting model and the intermediate steps, and can arrive at an arbitrary final model. Results from the simulation study in Ye et al. (2021) also showed that this approach tends to miss relations with moderate to medium strengths even with the correct starting model and a large sample size. Another critical downside unique to the current LV-GIMME is the sequential analysis. That is, factor scores of latent variable series are obtained through the traditional Bartlett method (Bartlett, 2011) or regression methods (Thurstone, 1935) in a separate step prior to the model building and are treated as observed variables in the estimation of the structural model. However, it has been shown analytically and numerically that a naïve use of factor score as observed variables without further correction leads to inconsistent and biased parameter estimates in the context of linear regressions (Skrondal & Laake, 2001) or simultaneous equations (Croon, 2002). In addition, there is a lack of knowledge about the impact of the measurement errors and random errors from the factor scores, regardless of the method of calculation, on model selection and estimation in simultaneous equations.

Regularization, in contrast, is a global, continuous model selection and a simultaneous estimation method. Regularization introduces sparsity by imposing a penalty term, the level of which is gauged by searching across a pre-specified range of λ values until the optimal λ (hence the sparsity level) is reached such that the model has the least mean square error or the lowest BIC (Jacobucci, 2017; Ye et al., 2021). When using the least absolute shrinkage and selection operator (LASSO, aka the $L1$-norm penalty; Tibshirani, 1996), the sum of the absolute values of the parameters are shrunken towards zero as λ increases, and they can eventually reach exactly zero. Hence, LASSO penalty is often used in favor of a sparse model or to perform model and variable selection. Ye et al., (2021) simulation study demonstrated success in adopting the LASSO regularization to identify a sparse huSEM with a high sensitivity (identify true paths) regardless of the magnitude as well as a high specificity (eliminating spurious relations). However, to the author's knowledge, LASSO regularization has not been implemented under the LV-uSEM context. Therefore, the current method seeks to replace the pseudo-ML stepwise searching and sequential estimation with the LASSO regularization for a simultaneous identification and estimation of the extended LV-huSEM.

To obtain the optimal solution, a ML regularized cost function is derived by adding the user-defined penalty function to the unregularized ML cost function:

$$F_{Reg}(\theta) = F_{ML}(\theta) + \lambda P(\theta^*) \tag{4.6}$$

in which $F_{ML}(\theta)$ is the unregularized cost function computed from the model implied covariance Σ^* given in Equation 4.5. The set θ includes all the parameters estimated in the model, while θ^* is the subset containing user-specified parameters under penalization (Notes cited from Ye et al., 2021).

Lastly, there is a lack of evaluation and comparison of these methods for the parameter estimation under data-driven model building procedures. Previous researchers have found that pseudo-ML estimates of individual DFMs obtained by the SEM approach showed higher biases and a tendency for inaccurate statistical conclusions compared with true ML estimates obtained from methods such as the state-space modeling approach with Kalman Filter estimator (Chow et al., 2010). But such evaluation was done on the correctly specified model, without potential biases associated with the model selection procedure. In addition, as an alternative model selection and estimation method, the property of LASSO regularized estimates in the LV-uSEM context has not been studied. In theory, regularization methods have sacrificed some level of unbiasedness for efficiency, because all the parameters under penalty (including the unknown true ones) are shrunk at the same time. But this does not mean that the LASSO estimates are always more biased than unregularized pseudo-ML estimates, because penalization eliminates unnecessary variables and false relations that can also bias the estimates of the correct parameters.

Importantly, both the pseudo-ML and the LASSO regularization are system-wide estimator that is not robust to any structural misspecification in the model because errors in one place can spread out to other parts of the model including those that are correctly specified (Bollen et al., 2021). Using a non-robust estimator to select the structural model would potentially impact final estimates in both the measurement model and the structural model. In contrast, the MIIV-2SLS is a limited-information equation-by-equation estimator that is more robust to structural misspecifications in the model (Bollen et al., 2007). Indeed, the MIIV-2SLS has been shown to be more robust than the pseudo-ML for the estimation of a DFM when estimated under the LV-uSEM framework (Fisher et al., 2019; Gates et al., 2019). Further, Bollen et al. (2018) has illustrated the analytic robustness conditions of the MIIV-2SLS estimator in SEM. Specifically, they found that misspecification errors from the structural model should not contaminate MIIV-2SLS estimates in the measurement model, whereas the impact of misspecifications in the measurement model on the structural model depends (see Table 7, page 858 in Bollen et al., 2018). This robustness property makes MIIV-2SLS an excellent choice for the estimation of sparse LV-huSEM when the heterogeneous, exploratory individual structural models are conditioned on a common, confirmatory measurement model. Because under the assumption of the correctly specified measurement model, the only source of misspecification comes from the structural model (as a result of model selection), which according to the robustness condition, will not bias the MIIV-2SLS estimates of the measurement model. This property, however, will not be guaranteed when system-wide estimators were used. That is, even though the measurement model is correctly specified, the errors from the misspecified structural model will likely bias pseudo-ML and LASSO regularization estimates of both the structural as well as the measurement model. Hence, the LV-GIMME research group adopted the MIIV-2SLS approach for the final parameter estimation post to the pseudo-ML based stepwise model search.

Therefore, to investigate unbiased and robust estimation under possible heterogeneous misspecified structural models, I included the MIIV-2SLS as a post model selection estimation approach as well (i.e., parameter estimation after the optimal sparse LV-huSEM model is selected). Besides model recovery property, the current study also seeks to compare unbiasedness and robustness behaviors of the pseudo-ML, the LASSO regularized, and the MIIV-2SLS estimators under the context of data-driven dynamic modeling. The current investigation will include the evaluations of these properties under the LV-huSEM with entirely idiographic structural models conditioning on a unified, confirmatory measurement model. Having an identical factor structure in the final LV-huSEM models across these methods ensures that the comparison of model recovery and parameter estimation for the dynamic relations are not contaminated by differences in the structure of latent factors.

In sum, the primary goal of the current chapter is to evaluate the different model building methods (pseudo-ML vs. regularization) with respect to model recovery as well as estimation approaches (e.g., pseudo-ML, regularization, and MIIV-2SLS) for unbiased and robust parameter estimations for a DAFM with a hybrid VAR representations (i.e., LV-huSEM). First, I extend the model specification for DAFM implemented in methods such as GIMME. That is, the LV-uSEM is extended to the LV-huSEM in the same way that uSEM is extended to huSEM in Ye et al. (2021), with the distinction that the huSEM is now constructed on the factor series instead of the observed time series variables. Second, the LV-GIMME approach used a pseudo-ML based stepwise model search on the factor scores. Instead, the proposed method adopts the LASSO regularization for model search and estimation in the extended LV-huSEM. One advantage of using LASSO regularization in this LV-huSEM context is that measurement model and structural model can be estimated simultaneously rather than separately. The simultaneous estimation avoids potential biases introduced by the errors in the factor scores. Finally, neither the unbiasedness or robustness of parameter estimates obtained by regularization has yet been evaluated under the current context. Indeed, the focus of existing literature is primarily on the recovery of true path presence and path direction (e.g., Epskamp, Waldorp, et al., 2018; Gates et al., 2019; Ye et al., 2021). However, it is vital to compare regularization with pseudo-ML and MIIV-2SLS for parameter estimates in the LV-huSEM. Therefore, in addition to model recovery, the current study also seeks to evaluate properties of parameter estimates using the current three estimation approaches (i.e., pseudo-ML, regularization, and MIIV-2SLS), especially in the presence of various misspecifications in the structural model recovered through model build procedures. Particularly, the author is interested in comparing the model search and estimation method using the proposed LASSO regularized LV-huSEM approach as well as a combination of LASSO regularization and MIIV-2SLS for LV-huSEM, versus the existing LV-GIMME approach, with the pseudo-ML and MIIV-2SLS estimation. The author hypothesize that LASSO regularization will be a more efficient and reliable method than the pseudo-ML based forward selection in the search for the optimal sparse model; and for parameter estimates, MIIV-2SLS is the most unbiased and robust estimator (under model structural misspecifications).

4.2.1 The Simulation Design

I designed a Monte Carlo simulation study to evaluate LASSO regularization and pseudo-ML approach with respect to model recovery as well as their properties for parameter estimation compared to those of MIIV-2SLS under the LV-huSEM context. The goal is to investigate to what extent the novel estimation method for the LV-huSEM models, i.e., with a LASSO regularization model build and post model selection MIIV-2SLS estimation, is superior to the existing approach similar to the individual model in the *LV-GIMME* framework, with respect to their 1) sensitivity of finding the true dynamic relations in the structure model; 2) the specificity of excluding the false dynamic relations; and 3) the unbiasedness and robustness of parameter estimates in response to misspecifications in the structural model.

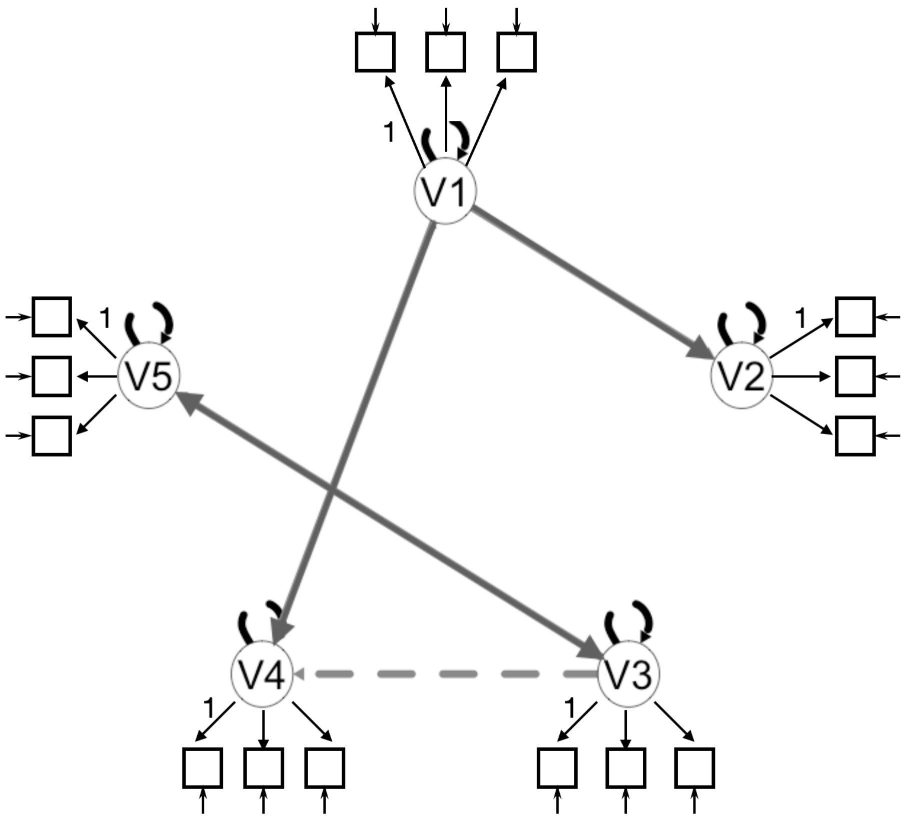
The Data Generating Model (DGM). The DGM is a five-factor DFAS with lag-1 effects as well as hybrid types of contemporaneous relations among the latent factors. In the measurement model, each factor is composed with three exclusive contemporaneous indicators (i.e., no cross loading or lagged relations in the measurement model), with one scalar indicator of factor loading equals 1, one with medium factor loading (.7), and one large (.9). To investigate the recovery for hybrid dynamic relations, I include paths of a mixed type and strength as I had between observed time series variables. This includes a lag-1 autoreggressive process within each factor (.5), a cross-lagged effect from lagged factor three to contemporary factor four (.7). For the contemporaneous relations amongst the latent factors, I included a moderate directed path from factor one to two (-.25), a strong directed path from factor one to four (-.7), as well as a moderate covariance between factors one and three (.25) and a strong covariance between factors three and five (.7). The varying magnitudes of coefficients within one DGM take into consideration that path strength is a potential factor that matters to the recovery of the true relations. However, the author does not claim that the combination of these parameter values returns a common LV-huSEM model in practice. In fact, this setting is chosen for an illustration purpose, which represents an overly simplified and sparse relations than what one would observe in reality. I will return to this point in the limitation section.

To investigate the influence of sample size on the performance: data is generated from the same DGM using time lengths varying from 60, 200, to 1,000, representing respectively a typical range from small to large in practice. This is to be consistent with the simulation design in previous evaluation on regularized huSEM (Ye et al., 2021). That is, the choice of these design factors are decided such that they represent data structure and characteristics of time series data in psychological and psychophysiological research. For example, although 60 might appear small in panel or cross-sectional data, it would be moderately large in time series such as daily dairy. Note that only the number of timepoints is crossed design, the other factors are investigated within one model. All the DGMs will be replicated 1,000 times, resulting in 3,000 datasets. The weakly stationary test will be performed in the data generating process, i.e., I will test that all eigenvalues of Φ have modulus less than one (Lütkepohl, 2005). All analyses will be performed in R, codes will be released and made publicly available on Open Science Framework (OSF).

4.2.2 Analytic Procedure

For the pseudo-ML approach, confirmatory five-factor measurement models are estimated by pseudo-ML in *lavaan* or by MIIV-2SLS in *MIIVsem*, and factor scores are obtained by the default regression method of the 'lavPredict' function in *lavaan*. These factor score series will enter the subsequent structural model for model selection using pseudo-ML forward search in the GIMME

Figure 4.1: DFM: A Time-invariant Five-Factor DFAM With A Hybrid uSEM Structure



package, function *indSEM*. The difference from the original setting in LV-GIMME is that here the starting structural model is a huSEM (with the covariance matrix ψ^*) instead of the more restricted uSEM (with ψ). Additionally, I focus on individual models only, i.e., no group level model is involved. For this reason, I refer to this method "pseudoML-FS-huSEM" to indicate it uses modification indexes for the search of sparse huSEM model using the factor scores.

For the proposed method, the LV-huSEM under LASSO regularization (i.e., LASSO-LV-huSEM) will be implemented under the regularized SEM framework. After the LV-huSEM model structure is specified in *lavvan*, *regsem* can import the *lavvan* output, i.e., the unregularized ML cost function $F_{ML}(\theta)$ derived from 4.6, and perform LASSO regularization with the user-defined list of parameters in the penalty function $\lambda P(\theta^*)$. Note that in LASSO-LV-huSEM the model selection and estimation are performed simultaneously on both the measurement and the structural model. To ensure that factor series represent latent constructs that are consistent with those of "pseudoML-FS-huSEM", the same confirmatory factor structure is estimated without penalty. Parameters in the measurement model (e.g., factor loading) belong to the freely estimated set in θ but not in set θ^*. Parameters in the set θ^* are regression coefficients for cross-lagged effects and contemporaneous effects (coefficients in the B matrix except the diagonal elements of the lower left block matrix Φ to free up AR coefficients) as well as the error covariance among contemporaneous latent factors (i.e., off-diagonal elements in the lower right block matrix Ψ^*. Ideally, the optimal of λ (with the lowest BIC) penalizes all unnecessary parameter(s) to zero and estimate the remaining parameter(s), unraveling the true type of relation between any two latent factors from five possibilities: two cross-lagged effects, two directed contemporaneous regression coefficients, and one undirected contemporaneous error covariance. For more details on the technical and steps in *regsem*, see Ye et al. (2021).

Two additional methods were included in the analysis to account for the confounding factor from the use of factor scores in the "pseudoML-FS-huSEM" method. The first one is to repeat the huSEM model search and estimation using factor scores obtained from the DGM (i.e., LV-huSEM) which I call "pseudoML-DGM-FS". By using the population parameters from the DGM, it should reduce the biases because some measurement errors in relation to the difference between the factor model and the DGM are accounted for. But note that factor score estimates would still differ from the latent variables - the problem is the measurement errors that are part of the indicators that form the factor scores. In reality, however, the true model is unknown and so it is impossible to implement this method in practice. It is included in the simulation just for the purpose of evaluation. Another confounding factor lies in the comparison of "pseudoML-FS-huSEM" (i.e., using psuedo-ML) and "LASSO-LV-huSEM" in their ability to select and estimate true relations and eliminate false ones in the structural model is the fact that the pseudo-ML is subject to measurement errors and random errors in the factor scores, while LASSO regularization simultaneously estimates the measurement and structural models without calculating factor scores. To account for that difference that confounds the comparison between pseudo-ML and LASSO regularization, I included an additional analysis to apply LASSO regularization on a huSEM using the same factor scores from the five-factor measurement model as did in the pseudo-ML approach, i.e., "LASSO-FS-huSEM". In that case, "LASSO-FS-huSEM" and "pseudoML-FS-huSEM" only differ in their model selection and estimation methods but not how the measurement model and starting structural model are constructed. While the difference between "LASSO-FS-huSEM" and

Table 4.1: Two framework for LV-huSEM: Model Build and Estimation Approaches

Modeling Framework	LASSO Regularization		pseudo-ML	
Method Name	LASSO-LV-huSEM	LASSO-FS-huSEM	pseudoML-FS-huSEM	pseudoML-DGM-FS
Analysis Procedure	Simultaneous	Sequential	Sequential	Sequential
Measurement Model	confirmatory	factor scores	factor scores	factor scores (DGM)
VAR Model Build	LASSO penalty	LASSO penalty	Forward stepwise	Forward stepwise
Parameter Estimate	LASSO/MIIV-2SLS	n/a	Pseudo-ML/MIIV-2SLS	n/a

the generic proposed method "LASSO-LV-huSEM" informs the impact of measurement errors from factor scores instead of latent variables.

Finally, recall that both pseudo-ML and LASSO regularization are for both model selection and estimation. But given the potential biases introduced by the model selection procedure and the robustness property of the MIIV-2SLS, the latter is also included as a post model selection estimation. That is, after the final sparse LV-huSEM is selected, it will be estimated again using the MIIV-2SLS to obtain the final parameter estimates of both the measurement and the structural models. In terms of the selection of MIIVs for each equation when the number of MIIVs exceed the minimum number required for model identification, previous simulations studies (e.g.,; Bollen et al., 2007) found that using one additional MIIV than the minimum number produces the least biased estimation at small sample size conditions, but matters less in large samples. The author chose to adopt this approach for the MIIV-2SLS estimation in the current simulation, given that the examination includes small to moderate sample sizes. We refer to this approach MIIV-2SLS-DF1 to indicate the one degree of freedom in the overidentification Sargen's test. With these investigations, we could examine which combination of model selection and estimation regime is the overall optimal practice, accounting for the treatment of the measurement model. The analytical steps and differences for the four methods are summarized in Table 4.1.

4.2.3 Evaluation Measures

For the reliability and accuracy in recovering relations with a correct direction, I used the same evaluation metrics: sensitivity and specificity. Sensitivity and specificity are the most common outcome measures for evaluating capability of recovery of connections in network research (e.g., Abegaz & Wit, 2013; Epskamp & Fried, 2016). Sensitivity is calculated by the ratio of the true positive count discovered in the search over the sum of all true relations in the DGM (i.e., true positives and false negatives). Sensitivity represents the power to detect true relationships. In this chapter, because the starting model is the more flexible LV-huSEM with all the free parameters in the extended Ψ^* and B^*, I do not distinguish path sensitivity from direction sensitivity as I did previously. That is, only the relations that are recovered with the correct direction are recorded. Essentially, the sensitivity concept here is equivalent to the direction sensitivity in Ye et al. (2021). Specificity, in comparison, is calculated by the ratio of true negative count over the sum of negatives in the DGM (i.e., the sum of true negative count and false positive count). This represents the percentage of non-existing paths in the DGM that the search procedure accurately omitted in the final model. These measures allow for a global evaluation of a model's ability to detect true recovery and to reject false ones. In both sensitivity and specificity measures, higher values indicate better performance in terms of the selection of true data-generating relations.

However, from previous observations, sometimes a relation between two variables will still be recovered with a misspecified direction. For example, at the presence of a directed relation between two contemporaneous factors (e.g., $Y_{t,1} \rightarrow Y_{t,2}$), an alternative relation such as a reversed sign ($Y_{t,1} \leftarrow Y_{t,2}$) or as lagged ($Y_{t-1,1} \rightarrow Y_{t,2}$ or $Y_{t-1,2} \rightarrow Y_{t,1}$) or as undirected covariance ($Y_{t,1} \leftrightarrow Y_{t,2}$) might be selected by the model. In some scenarios, it is better to have another form of relation from the true form than completely missing the relations, but not always. For example, if a direct path between two variables is missed, having a covariance or a lagged path with the correct direction in the selected model tells some information more than having no relation at all; however, if the directed path with the reversed sign is recovered, the information is misleading (i.e., a wrong causal implication). Sometimes, a relation with one or two wrong directions could be selected *instead* when the sample size is too small and the sampling error is large, but other times one or more could also be selected in addition to the true direction when the sample size is large. To investigate this behavior in relation to model selection methods and across sample sizes, I calculate relation-specific "direction false positive": the percentage of time where there are at least one (out of four possible) misspecified directions in relation to a given true relation being selected in the final model. Accordingly, I distinguish the overall model specificity with the one that eliminate the "direction false positive" related to true relations, I refer to them as direction specificity (more stringent) and path specificity (more lenient), respectively.

In terms of the assessment for unbiasedness and robustness in parameter estimations, I will exam the relative bias measure: mean relative bias (RB) for each parameter is calculated as the difference of the actual estimate and the true value divided by the true value, averaged across the cases where the path is recovered by the model (i.e., nonzero). Hence, this is a RB rate conditional on the path recovery. When MIIV-2SLS is used, the equation level over-identification test (i.e., Sargan's χ^2 test) is available. The Sargan's χ^2 test informs whether the MIIVs are indeed not uncorrelated with the error term of the corresponding equation. Rejection of the null hypothesis suggests that one or more of the MIIVs are inappropriate. This could occur if 1) only the equation is misspecified, 2) only the model is elsewhere misspecified, but the equation is correct, or 3) both the equation and other parts of the model are misspecified. A significant Sargan test cannot definitely tell which of these three possibilities is true, but it does alert the researcher to the potential of inconsistent coefficients estimators. Following previous literature (Fisher et al., 2019), we are interested in two properites: 1) what is the statistical power of the Sargan's χ^2 test when at least one of the MIIVs for an equation is wrong, and 2) Is the Type I error rate of the Sargan's χ^2 test accurate when all MIIVs are valid. Consistent with literature, I will use an α of .05 for both circumstances. The aim is to evaluate the RB and robustness behaviors under different misspecified structural models as recovered by pseudo-ML, regularization, versus MIIV-2SLS. From a practical point of view, the convergence behaviors for each method will be recorded and compared.

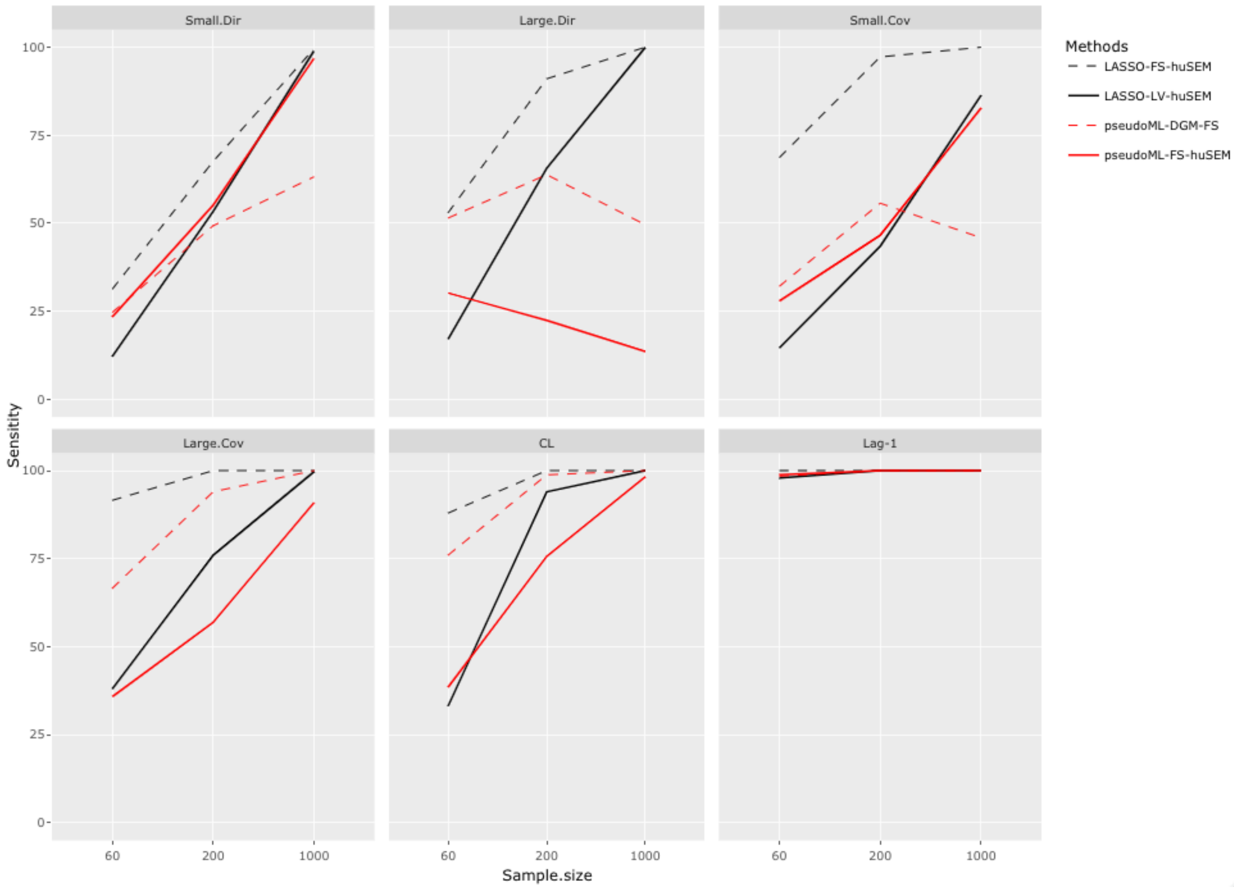
4.3 Result

4.3.1 Model Convergence

Some datasets caused nonconvergence when the LASSO regularization or the pseudo-ML approach estimates a LV-huSEM. These datasets were dropped from the analysis of the outcome measures below. It was observed that out of the 1000 datasets, there were 10.9 %, 7.3%, and 7.1% that did not converge for the one-step LASSO-LV-huSEM method at sample sizes N = 60, 200, 1000, respectively. These rates were increased to 21.9%, 14.5%, 8.3% (respectively) when the two-step LASSO-FS-huSEM method was used. The pseudo-ML method using factor scores from a five-factor measurement model did not converge for 2% datasets at N = 60. All the other conditions were converged. It is clear that the model using LASSO regularization using a starting model of a full LV-huSEM (with all the parameters included) has a higher chance of model nonconvergence compared to those using the pseudo-ML method that starts with a null model and a much small model specification (as the estimation of the measurement model is separate from that of the structural model).

4.3.2 Sensitivity and Specificity

Figure 4.2: Sensitivity of Path Recovery by Path Type and Strength across Sample Size



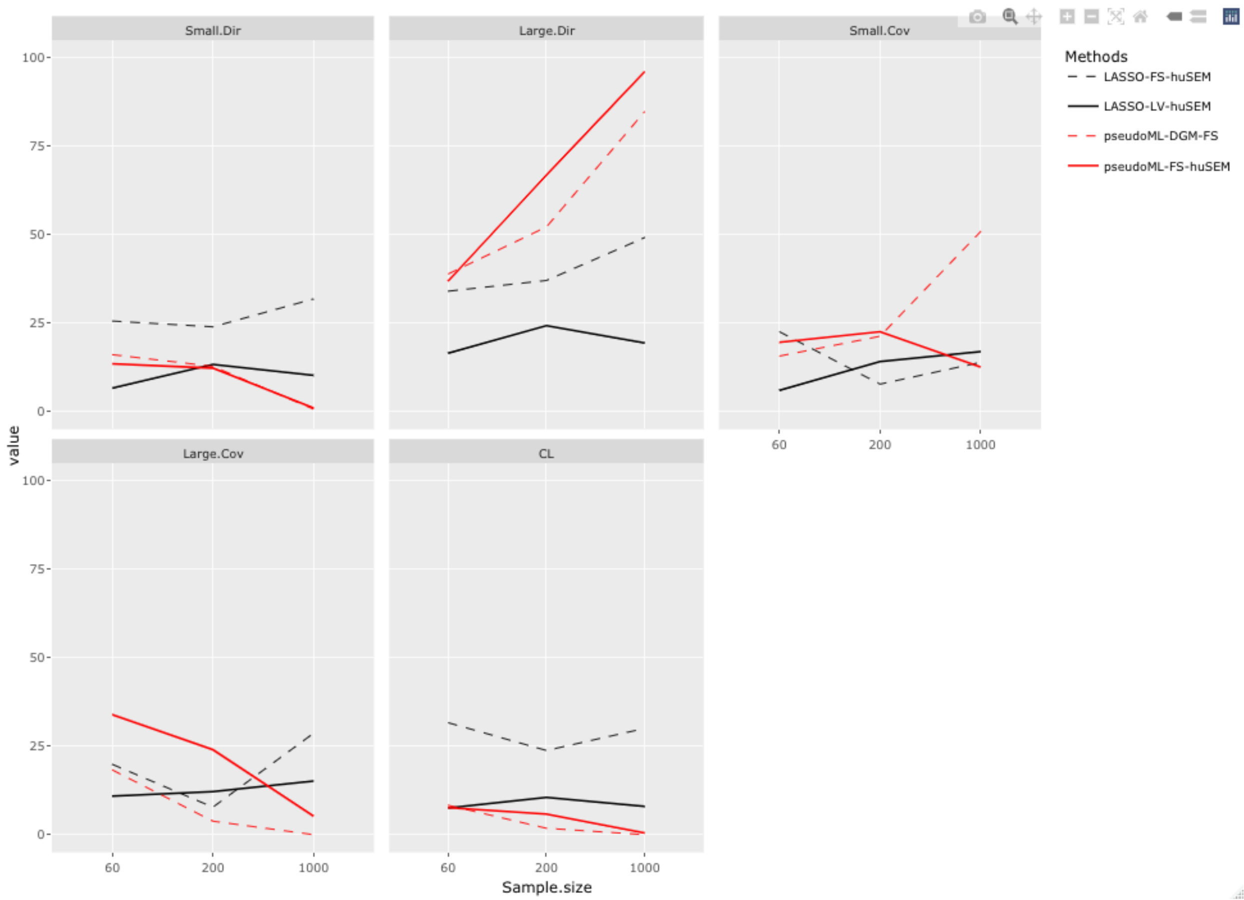
Note: Small.Dir = small directed path, Large.Dir = large directed path, Small.Cov = small covariance relation, Large.Cov = large covariance relation, CL = cross-lag effect, Lag-1 = lag-1 effect, Large.FL refers to factor loadings of .9, Med.FL refers to factor loadings of .7.

Let us first turn to the sensitivity for recovering true relations of the DGM from the starting LV-huSEM (i.e., confirmatory measurement model and an exploratory structural model with all the free parameters denoted in and B and ψ^*). All the methods showed an excellent sensitivity for lag-1 effects regardless of the sample size. Besides lag-1 relations, the probability to recover another true path by any method depends largely on the sample size: the recovery rates were low when the sample size was small (N = 60), overall acceptable at a medium sample size (N = 200) and satisfactory given a large sample size (N = 1000). Specifically, between the two generic methods of our interest, i.e., the pseudo-ML using factor scores (pseudoML-FS-huSEM) and the proposed LV-huSEM under LASSO regularization (LASSO-LV-huSEM), the performance of recovering a moderate directed path or a covariance relation were similar; however, LASSO-LV-huSEM showed an overall higher sensitivity to strong relations (i.e., directed, covariance, and cross-lagged relations) when given a medium or large sample size. Surprisingly, pseudoML-FS-huSEM performed poorly in recovering the strong directed path even with a large sample size. A closer examination revealed that the majority of time the model tended to recover a true strong directed path as a covariance relation and sometimes as a reversed sign directed path (hence a high rate of direction false positive, see Figure fig:III.false.positive). This is a scenario of a recovery that counted as a "path presence recovery" but not as a "direction recovery" in the simulation of Ye et al. (2021). Note that the distinction was emphasized there because the investigation involved more restricted starting models such that some types of relation were misspecified one way or another, their presence can only be recovered by an alternative form between the two variables. It is not the case here where both methods used the true starting structural model (i.e., huSEM) for the search.

For the other two methods for the investigation of the impact from the use of factor scores, it seemed that when a true DGM model (i.e., LV-huSEM) was used to obtain the factor scores for the subsequent pseudo-ML analysis (i.e., pseudoML-DGM-FS), the overall model recovery performance of the strong relations was much better than those from using the factor scores of a confirmatory measurement model alone, which suggests that measurement errors in the factor scores affect the recovery for the structural relations amongst the factors (when treated as observed in step two). Interestingly, in LASSO-FS-huSEM when LASSO regularization was used to perform the same sequential analysis as did in pseudoML-FS-huSEM, i.e., LASSO regularization on a hybrid uSEM of the factor score time series, the sensitivity outperformed all the methods across all conditions. This suggests that separating the measurement model from the structural model using factor scores as observed actually increased the chance of recovering the structural relations amongst the latent factors.

In order to investigate the misspecification in the direction of a recovered path, the author examined the path-specific "direction false positive rate", defined as the chance of recovering a true relation yet with a wrong direction (refer to the Evaluation Measures section for the more detailed description). Not surprisingly, it was observed that some relations were recovered with a wrong direction when the sample size was small. However, even when the sample size was sufficient and in many cases the true path was recovered, sometimes additional paths might still be selected when

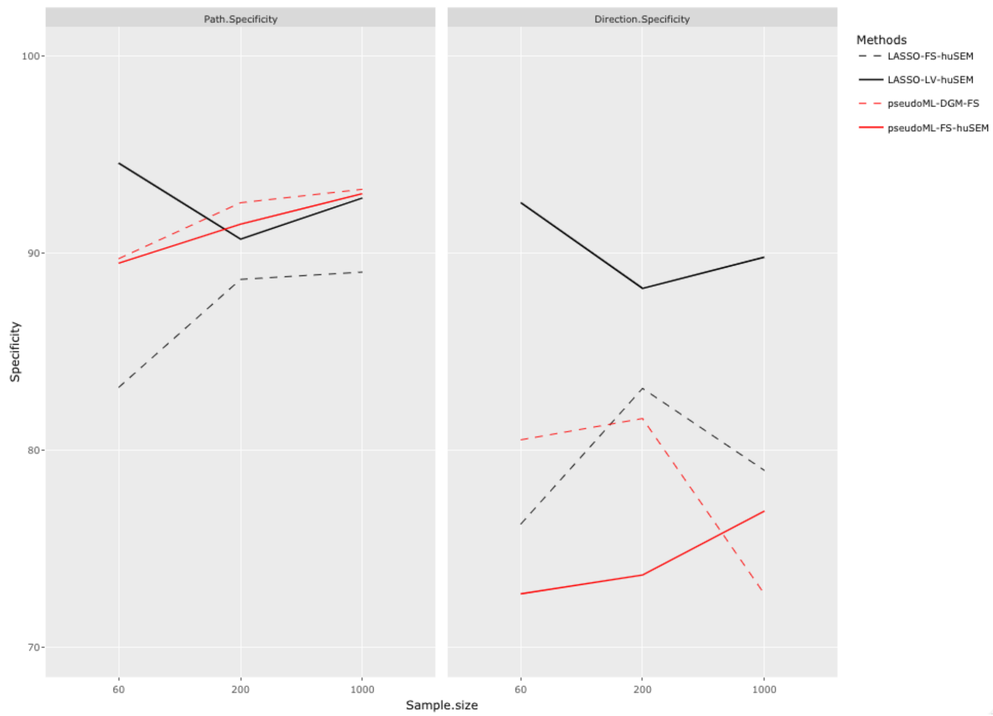
Figure 4.3: Direction False Positive by Path Type and Strength across Sample Size



Note: Small.Dir = small directed path, Large.Dir = large directed path, Small.Cov = small covariance relation, Large.Cov = large covariance relation, CL = cross-lag effect, Lag-1 = lag-1 effect, Large.FL refers to factor loadings of .9, Med.FL refers to factor loadings of .7.

there existed a strong correlation between the two variables. Hence, direction false positive rates did not necessary go down with the increase of sample size. Overall, except for cross-lagged relations, pseudo-ML methods had higher direction false positive rates in relation to the true paths in the DGM than did LASSO methods. This is partly the reason that pseudoML-FS-huSEM had a very poor sensitivity under some conditions. That is, some relations were recovered only with a wrong direction or type of relation. For instance, at sample size of 200 and 1000, both the LASSO-LV-huSEM (around 12-17%) and pseudoML-FS-huSEM (around 20-34%) methods had some tendency to recover a directed path at the presence of a true covariance relations between two contemporaneous factors. More problematically, pseudo-ML showed a high chance (67% at N = 200 or 96% at N = 1000) of recovering a reversed sign directed path or a covariance when there existing

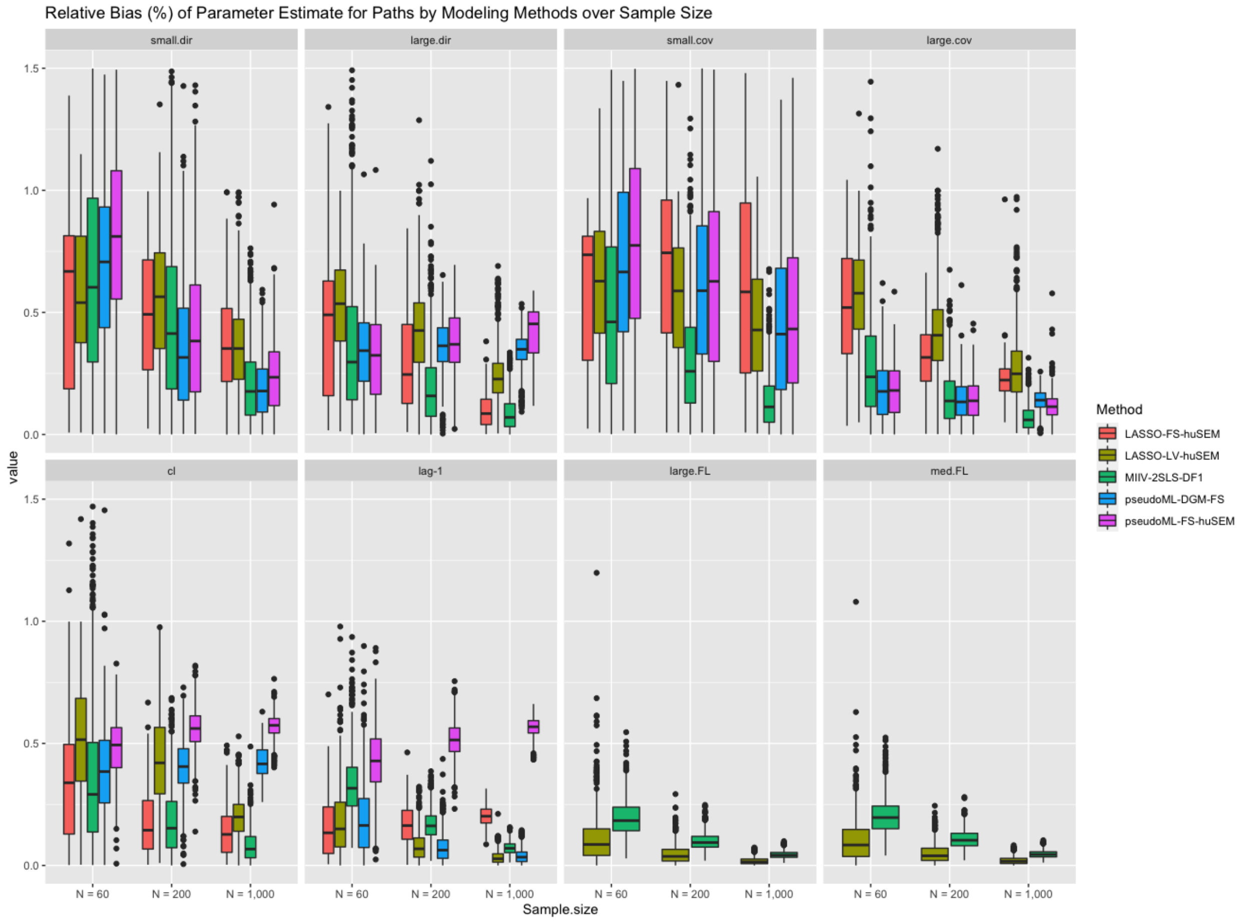
Figure 4.4: Path and Direction Specificity in the Model by Sample Size



a strong directed path. Using the factor scores from the DGM (i.e., pseudoML-DGM-FS) did not seem to decrease the chance of false positive directions, if not worse. In fact, using LASSO on the factors scores to select the structural model seemed to also introduce more direction false positive than the one-step LASSO-LV-huSEM model. These consistent observations that all methods using the factor scores showed a higher rate of direction false positive than their counterparts suggested that the issue of a wrong direction recovery of true relations is very likely resulted from the errors in factor scores of the latent variables.

Both generic methods reached a path specificity above 90%, suggesting they are reliable in rejecting false paths that were unrelated with those pairs of variables that have a true relation of another form or direction. However, the direction specificity (i.e., the odds of ruling out any path when it is truly false) dropped quite a bit for pseudoML-FS-huSEM (to around 72-77%) or any method that used factor scores. This is again because there was quite an amount of direction false positive paths in relation to true paths.

Figure 4.5: Conditional Relative Bias (%) by Path Type and Strength across Sample Size



Note: Small.Dir = small directed path, Large.Dir = large directed path, Small.Cov = small covariance relation, Large.Cov = large covariance relation, CL = cross-lag effect, Lag-1 = lag-1 effect, Large.FL refers to factor loadings of .9, Med.FL refers to factor loadings of .7.

4.3.3 Conditional Relative Bias

Overall, the post model selection MIIV-2SLS estimation was the least biased for parameter estimates of contemporaneous relations in the structural model, even though the estimates of moderate relations from any method were highly biased at a sample size of 60. In comparison, the LASSO penalized LV-huSEM methods produced the most unbiased estimates for freely estimated parameters, i.e., lag-1 autoregressive effects within each latent factor as well as factor loading in the measurement model, even at a small sample size as low as 60. Two methods produced similar results for cross-lagged effect among latent factors. Particularly, at a sample size of 200 and 1000, the conditional mean relative biases of lag-1 from LASSO-LV-huSEM were only 4.6% and 1.6%, respectively. The mean RBs of factor loading estimates were as low as under 1.8% at N = 200 or under 0.4% at N = 1000. When sample size was large, MIIV-2SLS-DF1 also produced estimates with small biases on average for factor loading estimates (under 5%), but the mean RBs were higher than those from LASSO-LV-huSEM at a small to moderate sample size (e.g., around 20% for MIIV-2SLS-DF1 compared to around 10% for LASSO-LV-huSEM at N = 60, or 12% versus 2% at N = 200, respectively). However, for the parameters under penalty in LASSO-LV-huSEM or LASSO-FS-huSEM (i.e., contemporaneous effects in the structural model), estimates were on average more biased (e.g., ranged from 20% to 40% at N = 1000) than those from MIIV-2SLS-DF1 (e.g., ranged from 5% to 18% at N = 1000), which was as expected given they were under penalty in LASSO methods while being freely estimated by the MIIV-2SLS after the model selection.

The comparisons between LASSO methods and pseudo-ML methods were mixed across different types and strengths of relations. For example, for lag-1 and cross-lagged estimates, pseudoML-FS-huSEM estimates produced much more biases (e.g., mean RBs was between 50 to 60% even at N = 200 or 1000) than did LASSO estimates (e.g., mean RBs ranged from 2% to 13% at N = 200 or 1000). One evidence that some of the biases come from errors in factor scores is that pseudoML-DGM-FS using factors scores from the DGM produced much less biased estimates than did pseudoML-FS-huSEM (using factor scores from a five-factor measurement model). Although in reality, the true DGM is unknown so the method pseudoML-DGM-FS is never available, here we include that only for the purpose of investigation on the impact of errors in the factor scores. However, this pattern did not apply to all types of parameters. For instance, the two pseudoML methods produced similar and slightly less biased estimates than did the two LASSO for moderate directed and large covariance relations at each sample size level, although all of them were still more biased than those from MIIV-2SLS. This suggests that errors in factor scores could introduce additional biases in some (e.g., lagged relations) but not all types of relations. However, all the pseudoML and LASSO estimates were largely biased for moderate covariance relations even with a large sample size (mean RBs ranged from 40% to 70%). The RBs of these estimates were also associated with a large variability across replications, indicating a substantial amount of influence from the sampling error. Therefore, MIIV-2SLS-DF1 is particularly useful to obtain less biased estimates for moderate or small structural relations.

One interesting observation is that mean RBs of pseudo-ML estimates did not show an asymptotic trend as did both LASSO estimates or the MIIV-2SLS estimates. That is, an increased sample size was associated with a decreasing trend of mean RBs in LASSO estimates and the MIIV-2SLS estimates across all parameters, but such trend was not consistently observed for mean RBs of pseudo-ML estimates. For example, the mean RBs of lag-1 and cross-lagged even went up as sample size increased. This suggests that the source of biases in pseudo-ML estimators do not just come from sampling error. This is more evidence of impact of systematic errors in the factor scores. As another way to investigate the shortcoming of using the factor scores instead of latent variables, we need to examine results between the two LASSO methods in which the only difference is that one uses factor scores (LASSO-FS-huSEM) while the other estimates a latent variable model simultaneously with a measurement model (LASSO-LV-huSEM). It seemed that the use of factor scores did not necessarily affect biases in the parameter estimates, as the overall mean RBs were similar between the two LASSO methods; however, in some cases the variability of RBs was larger in LASSO-FS-huSEM than that from LASSO-LV-huSEM, particularly when sample size was small. This seemed to suggest that sampling errors and measurement errors in factor scores might not affect the average accuracy of the parameter estimates of the structural model in a systematic way, but it affects the consistency of the estimates such that they are even less consistent at an increasing sampling fluctuation than methods using latent variables.

4.3.4 The Overidentification Test

The author evaluated the finite sample properties of Sargan's χ^2 overidentification test of MIIVs. The current examination involves the specificity (i.e., true negative) as well as sensitivity to wrong MIIVs (i.e., the statistical power of the Sargan's χ^2 test when at least one of the MIIVs for an equation is wrong). I investigated the case where a wrong MIIV was included as a result from an omitted true relation in the model. For instance, the omission of the directed contemporaneous relation from factor one to factor four will render a wrong inclusion of the scale indicator of factor one in the structural equation of factor four. Even though a significant Sargan's Test does not suggest the equation per se is incorrect, it offers evidence that one or more MIIVs of that equation are incorrect. This in turn indicates errors in the model specification that led to these MIIVs. In contrast, passing the Sargan test is consistent with a correctly specified equation and valid MIIVs for that equation. The author found that the specificity was around 95-97%. In other words, as expected there was less than 5% of rejection rate in which Sargan's χ^2 test incorrectly identified a wrong selection of MIIVs when in reality the model was correctly specified and all MIIVs were correct. This suggests accurate Type I error across the sample sizes considered here. For the test sensitivity, i.e., when a true relation was omitted from the model and hence a wrong MIIV set would be included in the corresponding equation, the Sargan's test rejected the problematic equation at the rates of 59% at N = 60, 66% at N = 200, and 92% at N = 1000, respectively. This suggests that the test has a moderate power to detect a wrong MIIV at a small to medium sample size, but can do so at a very high rate when the sample size is large.

4.4 DISCUSSION

In this chapter, the author extended the regularized hybrid uSEM framework in Ye et al. (2021) to the regularized latent variable hybrid uSEM (i.e., regularized LV-huSEM). In this way, the extended regularized LV-huSEM can be used to estimate a dynamic factor model with a hybrid VAR representation in the structural model. Three goals are accomplished here. First, restrictions on the contemporaneous relations between the latent factors are relaxed such that structural and covariance relations can be simultaneously estimated. This is similar to the extension from the uSEM to the hybrid uSEM using observed variables in Ye et al. (2021). Second, LASSO regularization is used to perform both model selection for the optimal sparse latent variable hybrid uSEM, and a simultaneous estimation for a freely estimated confirmatory measurement model and an exploratory structural model (with LASSO penalty on the structural paths and covariances between the contemporaneous latent factors). Compared to the current framework in GIMME, where measurement model and structural model are estimated sequentially with a stepwise model search procedure using factor scores obtained prior to the model selection, the current method provides a model search on a continuum and a simultaneous estimation without calculating factor scores. Finally, in order to reduce biases in the parameter estimates introduced by 1) model structural misspecifications, or 2) the LASSO penalty in the regularized LV-huSEM, or 3) errors of factor scores in GIMME, post model selection estimation is performed using the limited-information estimator MIIV-2SLS. The goal is to obtain less biased final parameter estimates of the selected sparse latent variable hybrid uSEM. The simulation study compared the two approaches and supplemental methods with respect to sensitivity to identify true paths, specificity to eliminate false ones, mean relative bias of each parameter estimate, as well as unbiasedness in the post model selection MIIV-2SLS estimators.

For model recovery, the simulation results revealed that the two approaches have comparable recovery rates for some relations such as lagged effects and moderate contemporaneous effects among factors, and they both are reliable in recovering a close-to-true structural model when the sample size is medium to large. The impact of factor scores on model selection under the SEM context was largely unknown. Even though that the LV-GIMME study (Gates et al., 2019) found that the path recovery performance does not seem to be related to what approach was used to derive the factor scores, there is no comparison with a simultaneous estimation method without the use of factor scores. In addition, the evaluation was on a restricted DFM, i.e., LV-uSEM with only directed contemporaneous relations among factors, which might not apply to the recovery of the more complicated model, LV-huSEM in which the recovery of a true relation between two variables involves a selection among five possible parameters. Indeed, the author found that the pseudo-ML methods using factor scores have a higher chance to commit a direction false positive on strong directed relations, that is, a tendency to recover a strong directed as one with a reversed direction or as an undirected covariance relation. This low direction specificity is likely results from errors in the factor scores. And it is shown that the performance is improved when factor scores from the DGM are used instead, although the use of such factor scores is still different from the latent variables (due to the measurement errors from the part of the indicators that form the factor scores). Further, in the additional analysis of applying LASSO regularized hybrid uSEM on factor scores, the likelihood to commit a direction false positive is also higher than that of the simultaneous LASSO regularization hybrid uSEM with latent variables. The result suggests that the use of factor scores instead of latent variable approach is more likely to select a model with a higher false positive rate.

This tendency of recovering a relation with a wrong direction undermines the purpose of adopting the more flexible VAR representation, i.e., to accurately represent the hybrid forms of contemporaneous relations that might coexist in practice. If we recall the concepts reviewed in chapter one, causal implications represented by a model with only directed paths or only undirected covariance could be very different. However, if we choose to use a more flexible model representation, but the chance of selecting false positive relations with a wrong direction is high by the model selection method, we still end up with misleading causal interpretations. In this sense, when there might exist some strong contemporaneous relations, LASSO regularization seems to have a higher tendency to eliminate false positive relations and avoid misleading causal interpretations than does the pseudo-ML method using factor scores. To the author's knowledge, this is the first evaluation on the impact of using factor scores for model selection under the uSEM context.

In terms of parameter estimation, as expected, the post model selection MIIV-2SLS estimator is the least biased for parameters in the exploratory structural model. But slightly surprisingly, the proposed LASSO regularized hybrid uSEM with latent variable approach produced the least biased estimates for free parameters including factor loading coefficients in the confirmatory measurement model and the lag-1 autoregressive effect between factors. This suggests that the biases for other regularized parameters are mainly from the LASSO penalty. In practice, if we are using regularized SEM method, we should only penalize the uncertain (i.e., exploratory) part of the model. However, both the pseudo-ML estimator and the LASSO regularized estimator introduce quite some biases in the parameter estimates of the structural model, especially for moderate relations. Hence, post model selection estimation might indeed be a more practical choice to obtain final estimates.

We also observe at least two major downsides of using factor scores in parameter estimation, besides its impact on model selection (i.e., the tendency to select a directed path with a wrong direction). First, the estimation of the lagged effects seem to be particularly largely biased, regardless of sample size. This is not surprising because no matter what estimation method is used to obtain the factor scores, the sample variance matrix of estimated factor scores is an inconsistent and biased estimate of the true variance matrix of factors (Croon, 2002; Skrondal & Laake, 2001). In addition, the time embedding process of the data to obtain lagged factor variables introduces more random error on top of the measurement error within the contemporaneous factors variables themselves, thus causing additional biases in the estimates of relations between the lagged factor variables and the contemporaneous factors variables. Second, besides a high average level of biases in estimates of moderate contemporaneous relations regardless of which factor score methods are used, methods using factor scores also tend to produce less consistent estimates that is more subject to sampling fluctuations. The substantial amount of variability in the biases across samples especially with a small to medium sample size is very likely a consequence of the sampling error in the sample variance matrix of estimated factor scores.

4.4.1 Limitations and Future Directions

Limitation in the two approaches under investigation. The simultaneous analysis using LASSO regularization under the LV-huSEM is easy to implement and can avoid biases from the use of factor scores, but it might be more limited in the size and complexity of the model (e.g., number of variables, factors, density of the structural paths, etc.) than is sequential analysis like the GIMME approach. This is because use of factor scores reduces the dimension of the parameter space - the number of parameters is higher in the simultaneous model as it includes estimates for the measurement model besides the structural model. Optimizing the covariance matrix of observed variables with a higher dimension is more difficult than that of the latent factors. Another advantage of a sequential analysis is that it is at a better chance to avoid improper solutions or nonconvergence issues that is not uncommon in simultaneous estimation methods. Nevertheless, the issue with a high direction of false positives using factor scores makes it a less appealing choice for model selection.

Although the use of MIIV-2SLS estimation reduces biases from model misspecification and model selection, the naïve post model selection statistical inference is subject to another source of bias due to the duplicate use of the data, that is, the same dataset is used for model selection and model estimation. Since statistical inference is established under the assumption that the fitted model is known in advance, which is clearly violated here, the naïve inference of a regularized LV-huSEM after the data driven selection process for variables and relations is no longer valid. Because both the randomness introduced by the selection process and the sample space restriction implied by the chosen model influence the sampling distribution of the estimator and need to be accounted for (Huang, 2020). Indeed, it has been found that the naïve method after model selection using regularized SEM tends to obtain significant results for selected zero parameters, resulting in numerous false positive findings in psychology (Huang, 2020). However, post model selection inference methods are extremely difficult to perform, and few has been developed for regularized SEM or been implemented in the regsem package upon the time of this work. This is probably because the primary goal of statistical learning method such as LASSO regularization is to achieve the least biased prediction, while the aim of traditional psychology and of the current study is to identify the optimal model from which to obtain statistical inference for individual parameters. This is one of the biggest gap yet to be filled in the future development of statistical learning under SEM framework. It would be useful to make these methods available in the software and packages for regularized SEM, so that future studies could be conducted to evaluate and validate their properties under the regularized hybrid uSEM with latent variables.

Another aspect that is out of the scope of the current study is DFMs for multiple subject time series. The focus of my dissertation is on individual dynamic models, for which there is no consideration of between-person effects nor attempt to aggregate individual models. However, the use of group-level or between-person information (i.e., similarities and variances across individuals) has been shown as an effective way to extract true effects from noise information so that it avoids the risk of over-fitting individual dynamic models (Asparouhov et al., 2018; Gates et al., 2019). In fact, one of the strengths in the GIMME algorithm is the construct of a group-level model with the most shared information is found and be used as the starting model for each individual model. This would likely to address to some extent the high direction false positive rate. Besides, one recent method called “multi-VAR" is proposed (Fisher et al., 2022) that used LASSO penalization on multiple subjects multivariate TSD for the forecast of dynamic processes at the individual level. The goal of multi-VAR is to identify an optimal sparse VAR that achieves the best prediction, rather than recovering the true model or statistical inferences of the selected model. Another framework is to aggregate individual dynamic results using a multilevel structure, such as the Dynamic Structural Equation Modeling (DSEM; Asparouhov et al., 2018), where multiple subjects time series are modeled simultaneously to estimate population mean and individual differences (deviations from the mean) in the parameters governing a dynamic process. In future studies, the current LASSO regularized hybrid uSEM with latent variable can be extended to multiple subject cases either in the form of the group-level model or the multilevel structure. When that is done, we can compare these methods under LASSO regularization, GIMME, as well as DSEM in modeling multiple subject DFM with hybrid VAR representations.

Limitation in the simulation design Like all simulations, the simulation factors do not represent a comprehensive list of empirical situations. To keep the scope manageable, the author did not adopt a cross design of factors such as path strength, model size (i.e., number of variables), level of measurement error in the latent factors, etc. The DGM might also represent an over-simplified, over-sparse DFM, with a very standard measurement structure without cross-loadings or local dependency structures. In practice, the structural relations in a DFM could be much denser with many weak to moderate relations. In addition, the measurement model could also be extended both in the sense of containing some error correlations as well as to include lagged effects between the observed indicators and latent factors, i.e., a white noise factor model or a hybrid of WNFM and

DAFM. These added complexities pose a bigger challenge for the model selection and parameter estimation, because a denser model means that the model search and optimization for solution are performed at a larger parameter space with higher dimensions. This might require a harsher penalty term, which in turn might lead to larger biases in the estimates of correct parameters, although the current proposal to use the post model selection MIIV-2SLS estimation can compensate the biases. Future development in the optimization algorithm for regularized SEM in general is the key to estimate a LV-huSEM with more complex structures and higher dimensions.

Our results of the LASSO regularized hybrid uSEM with latent variables highlights the flexibility of the LASSO regularized SEM in estimating individual DFMs. The data-driven LASSO penalty opens up a variety of possibilities in the development and appraisal of individual dynamic theories. The penalization structure relies on which part of the model is more supported by theory, and which part is more uncertain and needs to be explored by the data. For instance, when the latent factor structure among the candidate indicators is not fully determined by the theory, partially exploratory model selection could be implemented on the measurement model. Specifically, if some indicators of a factor are confirmed by the theory, but the rest of the indicators are not, we can adopt the “semi-confirmatory" factor model from Huang (2020) in the way such that the factor loadings from the uncertain indicators are included as parameters under penalization while the factor loadings of the confirmed indicators are free parameters. This way, only the uncertain part of the measurement model is under the data-driven LASSO selection. This idea could also be used to explore the hybrid form of WNFM and DAFM when the theory is not enough to determine which DFM is more appropriate. That is, in the case where the measurement model may contain either the lagged factor loading, or the contemporaneous factor loading, or some combination thereof. One way is to penalize the lagged factor loading and the contemporaneous factor loading simultaneously, as an automatic search between a WNFM and a DAFM with respect to the factor loading structure across time. Alternatively, we can retain the confirmed structure of the contemporaneous factor loading and only penalize the lagged factor loading if the latter is optional. Depending on the knowledge of the theory available to us, we can choose which relations we feel confident enough to be included as free estimates, and which we are less so and thus allow the data to decide by imposing a LASSO penalty on the corresponding parameters. Therefore, the flexibility of regularization with user-defined estimation and penalization structure lifts the dichotomous boundary between the exploratory approach and the confirmatory one and allows for an expansion and refining of theory on a continuum.

CHAPTER 5

CONCLUSION

5.1 Summary of the Unifying Theme

My dissertation aims to evaluate, advance, and extend current psychometric modeling approaches for person-specific individual dynamic assessments based on Vector Autoregressive (VAR) models. Studies that focus on person-specific individual dynamic assessment emphasize individual differences in the dynamics and characteristics unpacked by multivariate time series data. I reviewed and compared two representative approaches commonly used in current psychological studies: the unified Structural Equation Model (uSEM) framework, as a time series extension of SEM, and the graphical Vector Autoregression (gVAR) model, as a time series extension of the network psychometric model. I found these approaches differ in the variant of VAR representation used in the model (i.e., often restricted in one way or another), as well as the estimation framework to select the optimal model. My dissertation started by discussing the differences in the interpretations and causal implications carried from the adoption of different VAR variants in these psychometric approaches (Chapter II). Given the limitation in the current model representation and model building routine, I then proposed a novel modeling framework using machine learning method for the selection and estimation of the optimal dynamic model (Chapter III). Finally, I extended the proposed framework under the latent variable contexts and evaluated not only model recovery performance, but also (asymptotic) statistical properties of the proposed method versus those of the current approaches (Chapter IV). The conclusion, limitations, and future directions are discussed for each study. Below, I summarize the purpose, design, and major findings of each study, followed by a general conclusion of the dissertation.

5.2 Summary of Chapter II Study

Despite the vast amount of methodological development for individual dynamic model, few studies discussed the differences in the interpretations and causal implications carried from the adoption of different VAR variants in these psychometric approaches. The first study of my dissertation aims to explicate the interpretative caveats of choosing the options for modeling contemporaneous relations among variables. To do so, I provided a series of small simulation examples and two empirical examples to demonstrate the differences from DAG, Granger causality, and exogeneity perspectives. For each example, I fitted time series models with either directed and undirected contemporaneous VAR, and discussed how causal implications are altered as a result of the contrasting way of representing contemporaneous relations. This focused discussion will fill a timely and vital gap given the increasing application of these dynamic models in psychological and neuroscience research.

I emphasized major conceptual points in interpreting the directed and undirected contemporaneous VAR models. First, causal inferences can be made only from structural relations that carry casual assumptions, whereas partial correlations or marginal correlations without causal hypothesis do not correspond to a unique causal interpretation. Second, not all Granger causality relations can be accurately depicted by the temporal dependencies between repeated observations and estimated as lagged relations. For instance, the lag-1effect might not be sufficient to represent instantaneous relations that occur when short-term effects are missing from the temporal ordering information in the data. While the contemporaneous network can recover some remaining information as undirected partial correlations, the information of the direction is lost, in which case causal implication is intractable. Further, at the presence of simultaneous causality or two-way instantaneous Granger causality relations, nonrecursive structures should be included, which can also be understood as an approximation to bidirectional cross-lagged effect unfolding over time. Otherwise, if nonrecursive relations are modeled as bidirectional covariances or as undirected partial correlations, biased parameter estimates and misleading inferences would occur.

The two empirical examples further illustrate and complement to these points under psychopathological and the neuroimaging settings, respectively. By comparing the differences in the networks disclosed by contrasting models, it is seen that the model revealed by directed VAR traces back to the sources of variation (i.e., cause), and provides us the direction of information flow in the underlying causal pathways. On the variable level, the importance of each variable involved is described by where it stands on the pathway. The undirected VAR model, by contrast, is characterized by mutual interactions among variables rather than the information flow in one direction. The centrality of a variable is revealed by the quantity of existing links related to other variables in the model.

In sum, I reviewed traditional causality concepts in the interpretation of individual-level directed and undirected contemporaneous models, and illustrated how person-specific inferences are altered by the model choice that can inform both psychometric and neuroscience research. This pedagogical piece highlights the extent to which inferences of an idiographic dynamic network obtained by one approach differs from that of the other, without making claims as to which is absolutely superior to recover the data generating model. I remind the readers that this study does not target at a comprehensive methodological advancement of the modeling approaches. It is out of the scope to provide a new model choice procedure. However, the theoretical discussion implies that there is an increasing demand of an active methodological research area, e.g., the development of some types of hybrid approaches that encompass the current ones. In addition, due to limited space and time, I do not aim for providing an exhausting list of data examples in this paper, but future work can extend the evaluation to diverse contexts or perhaps with more complex data structures.

5.3 Summary of Chapter III Study

To the best of the author's knowledge, the majority of current psychometric dynamic approaches are limited in at least two ways: 1) VAR models are restricted in that contemporaneous relations are represented either as only undirected or as only directed, but not both, despite the fact that they carry different casual interpretations that could coexist in a dynamic system; 2) the current estimation frameworks are heavily reliant on stepwise model building procedures, and few research has compared the properties of these traditional methods of model selection with those recently adapted from statistical learning, i.e., regularized SEM (Jacobucci et al., 2016). To fill in these gaps, I proposed and evaluated a novel modeling approach in Chapter III of the dissertation, where LASSO regularization is used for the identification and estimation of a more flexible VAR representation. First, in studying person-specific dynamic models, it is possible to observe the coexistence of instantaneous causal pathways and simultaneous connections beyond lagged effect. Researchers should consider modeling the intraindividual variability where hybrid forms of contemporaneous relations can be estimated. I extended the current unified SEM to a hybrid representation (i.e., hybrid uSEM), which incorporates both the undirected and directed contemporaneous effects to be identified by a data-driven algorithm. Second, for a more efficient automatic model search tool, I replaced the stepwise procedures with a regularization method implemented under the unified SEM framework. This new framework allows for a global and continuous search for the optimal sparse model. A simulation study is conducted to investigate whether the proposed approach is indeed superior to alternatives in terms of recovering true relations and removing false ones.

The simulation study served two major goals: 1) to validate whether the hybrid uSEM can correctly recover the data generating relations when directed and undirected contemporaneous relations are both presented in addition to AR and cross-lagged effects among multivariate time series, regardless of the model search procedures. At the same time, I assessed to what extent the other VAR variations (i.e., uSEM and gVAR) can recover the presence of relations even though the data generation does not match the modeling approach; and 2) to evaluate whether regularization performs better than does forward selection for the uSEM and hybrid uSEM models. I tested if model recovery performance differs between the adaptive LASSO and the LASSO penalty when applied to uSEM or to hybrid uSEM, respectively. I evaluated these modeling framework with the consideration of the impact from conditions including the number of variables, strength of relations, as well as number of timepoints.

The simulation results showed that LASSO regularized hybrid uSEM performed uniformly the best over alternative VAR representations and/or modeling approaches, with respect to accurately recovering the presence and directionality of hybrid relations and reliably removing false relations when the data are generated to have two types of contemporaneous relations. This finding validated that the sparse hybrid uSEM model obtained by regularization with LASSO penalty was the closest to the true model among all model specification and estimation methods. That is, the presence and direction of true paths of hybrid forms were recovered with the highest accuracy and specificity. This conveys the message that the more flexible hybrid uSEM, rather than any restricted form of VAR, should be the starting model to fit time series data when the data generating model is unknown but could potentially contain both types of relations. Further, it is stressed that a statistical correlation or covariance represents an existence of connection whose source is exogenous to the model, and barely informs the direction of information flow as those in casual pathways. In contrast, a contemporaneous directed relation is indicative of a Granger causal effect that is failed to be captured by temporal dependency due to inappropriate time scaling. The empirical example offered more evidence, where the undirected VAR model revealed a very sparse contemporaneous model using the fMRI data. This is likely because that the underlying neuronal brain activity happens much faster than what is captured via the fMRI data.

The other important contribution of this study is the use of regularization under the extended SEM framework to accomplish a data-driven model selection for dynamic models. Specifically, the results showed clearly that the LASSO regularization consistently outperformed forward-selection step-wise search with respect to reliably and accurately identified moderate relations regardless of which starting model is used. The modification index method, in contrast, seems to be only sensitive to the presence and direction of strong relations. This suggests that the inclusion or deletion of moderate connections are unlikely to cause a significant change in the modification index. Therefore, the LASSO-regularized hybrid uSEM model is the best combination of modeling and estimation framework for intraindividual variability analysis using multivariate time series data.

However, like many other simulation studies, there were several limitations in the design and scope of the current examination. First, as a preliminary investigation, the simulation design is modest in the random simulation factors (e.g., sparsity of the data generating model, number of variables, etc.), this is for the sake of keeping a manageable and (time) economical scope. The choice for many of these factors is, to a large extent, limited by the complexity in the optimization algorithm, computational speed and feasibility. Second, the generalization of the result is limited by the few assumptions made in the data generating model or the estimation model (e.g., weakly stationary multivariate time series, time-invariant dynamic relations, first-order lagged relation, etc.). In practice, it is likely that there exists nonnegligible nonstationary trends in the time series, higher-order dynamic relations, or time-varying relations through the passage of time. The findings should be generalized to these situations with caution, given a lack of consideration for more complex dynamics and the adoption of a simple model structure.

The author notes that the future direction of this methodological study is mainly on the advancement of the optimization algorithm and software development for the regularized SEM. The current examination does not represent an exhausting investigation of available penalty methods and optimization algorithms, but only as a first attempt to apply the recent regularized SEM under a time series context. For instance, in future development of the regularized SEM, more reliable adaptive versions of LASSO penalty should be examined under scenarios when the unregularized

SEM is unidentifiable (i.e., the unregularized ML estimates are intractable). Indeed, solving the fitting function in SEM often involves nonconvex and nonlinear optimization, meaning that unique and estimable solution can only be reached under very constrained scenarios (Pruttiakaravanich & Songsiri, 2018). But with the rapid ongoing development in contemporary computational machine learning, we should remain confident for the future advancement in this area. Nevertheless, the present study, to the author's knowledge, is the first application of the recently developed regularized SEM technique to the estimation of a hybrid version of time series SEM. The success of the application points to a promising future in exploratory SEM using statistical learning techniques.

5.4 Summary of Chapter IV Study

The previous examinations have focused on VAR-based psychometric models with dynamic relations among manifest variables without accounting for measurement error. In practice, however, it is common that the dynamic processes are at the underlying latent construct level. The combination of a latent variable model and a VAR model results in what we know as the dynamic factor model (DFM; Browne & Nesselroade, 2005). In the last study (Chapter IV), I extended the proposed hybrid uSEM to the latent variable hybrid uSEM, and examined model recovery and estimation properties under a process factor model (Browne & Nesselroade, 2005) where hybrid dynamic relations are only among latent factor time series variables.

Therefore, Chapter IV of my dissertation serves to advance and investigate the model search and estimation for a DFM with a hybrid VAR representations. Three steps were taken in the proposed method: first, transform the structural model of the latent variable uSEM (LV-uSEM) to its hybrid uSEM version, i.e., the latent variable hybrid uSEM (or LV-huSEM); second, perform model selection using the LASSO regularization in the search for the optimal sparse LV-huSEM; lastly, the final selected sparse LV-huSEM is estimated via the model-implied instrumental variables with two-stage least squares (MIIV-2SLS; Bollen, 1996; Fisher et al., 2019) to obtain final parameter estimates. The MIIV-2SLS estimation is used for its robustness property under model structural misspecification (Bollen et al., 2007), which is particularly advantageous for the estimation of sparse LV-huSEM selected in a data-driven manner. This is compared to the latent variable GIMME framework (Gates et al., 2019), where a LV-uSEM is estimated in a sequential fashion with the measurement model estimated first, followed by a structural model of the computed factor scores using pseudo-ML based forward stepwise model search and a pseudo-ML or MIIV-2SLS estimation.

However, the current focus is on individual models without using the group level modeling in GIMME, hence I referred to the compared method as pseudo-ML method. A simulation study was conducted to investigate to what extent the novel estimation method for the LV-huSEM models, i.e., with a LASSO regularization model build and post model selection MIIV-2SLS estimation, is superior to the pseudo-ML approach similar to the individual model in the LV-GIMME framework with respect to 1) model recovery performance (sensitivity and specificity) of the LASSO regularization versus the pseudo-ML based stepwise model build when they are applied under the LV-huSEM context; and 2) biasedness of parameter estimates in sparse LV-huSEM obtained by LASSO regularization, pseudo-ML, versus the post model selection MIIV-2SLS estimation.

The simulation results revealed that for model recovery, the two approaches have comparable sensitivity to lagged effects and moderate contemporaneous effects among factors. However, the pseudo-ML methods using factor scores is more likely to commit a direction false positive on strong directed relations (i.e., a tendency to recover a strong directed as one with a reversed direction or as an undirected covariance relation). The high false positive rate is likely a result of using factor scores instead of latent variables approach for two reasons: 1) it is shown to be improved if factor scores were drawn from the true data generating LV-huSEM (rather than a factor model), although such factor scores would still differ from the latent variables due to the measurement errors from the indicators that form the factor scores; and 2) in the additional investigation, I found that the likelihood to commit a direction false positive is higher when the LASSO regularized hybrid uSEM is estimated using factor scores than that using the latent variable approach. In terms of parameter estimation, the post model selection MIIV-2SLS estimator is the least biased for parameters in the exploratory structural model, while the proposed LASSO regularized hybrid uSEM with latent variable approach produced the least biased estimates for free parameters (i.e., factor loading coefficients in the confirmatory measurement model and the lag-1 autoregressive effect between factors). This suggests that the biases for other regularized parameters are mainly from the LASSO penalty. In practice, if we are using regularized SEM method, we should only penalize the uncertain (i.e., exploratory) part of the model. In addition, both the pseudo-ML estimator and the LASSO regularized estimator introduce substantial biases in the structural parameters, especially for moderate relations. Hence, post model selection estimation is particularly useful for weak to moderate parameters.

The findings on the LASSO regularized hybrid uSEM with latent variables highlight the flexibility of the LASSO regularized SEM in estimating individual DFMs. The data-driven LASSO penalty opens up a variety of possibilities in the development and appraisal of intraindividual dynamic theories. Because the user-defined penalization extends the dichotomous choice between the exploratory approach and the confirmatory approach to one on a continuum. It is found that the simultaneous analysis using LASSO regularization is easy to implement and can avoid biases from the use of factor scores in the current pseudo-ML approach; however, it might be more limited in the size and complexity of the model (e.g., number of variables, factors, density of the structural paths, etc.) than is sequential analysis like the pseudo-ML approach. Another advantage of a sequential analysis is that it better avoids improper solutions or nonconvergence issues that are not uncommon in simultaneous estimation methods. Nevertheless, the issue with a high direction false positive using factor scores makes it a less appealing choice for model selection. Another limitation in the current work is the use of naïve post model selection MIIV-2SLS estimation, which is subject to another source of bias due to the duplicate use of the data for both model selection and estimation. Future studies need to develop reliable and easy to implement methods of post model selection statistical inference. Another future direction is to develop methods of LV-huSEM for multiple subjects time series. The current dissertation on individual dynamic models has not considered between-person effects nor attempted to aggregate individual models. However, accounting for similarities and variances across individuals has been shown to be informative to extract true effects from noise and to avoid the risk of over-fitting individual dynamic models (Asparouhov et al., 2018; Gates et al., 2019). In future studies, the current LASSO regularized hyrbid uSEM with latent variable can be extended to multiple subject cases either in the form of the group-level model similar to that of GIMME or as multilevel structure such as in the Dynamic Structural Equation Modeling (Asparouhov et al., 2018).

5.5 General Conclusion

In conclusion, the major contribution of the current dissertation research to the methodology literature is that the regularized hybrid unified SEM offers a flexible model search and estimation framework for individual dynamic models. In terms of the application to substantive research in psychology, the author reminds researchers to keep in mind the strengths and shortcomings of the proposed approach. First and foremost, the adoption of the hybrid uSEM representation is particularly useful for cases where there are a handful of variables or latent constructs and causal implications is of interest, especially only some causal assumptions can be made from the available literature. In other words, (with some uncertainty) one would expect some relations are causal by nature (e.g., mediation, common cause, etc.), some are pure association, and the overall structure is subject to some levels of exploration. Second, this type of model requires a certain amount of sample size. It might be challenging for some studies such as daily diary to have a few hundreds to thousands of time points, while in neuroimaging or biometric studies this might be less an issue. Third, because the property we found is conditional on the weakly stationary assumption in the time series data and time-invariant parameters are estimated, it might not be very useful for studies such as developmental psychology with meaningful nonstationary trend or time-varying relations. For instance, if the trend of growth (or decrease) carries crucial developmental meanings, the removing of the trend takes away the focus of interest; in addition, the stage-specific time-varying dynamic relations might be more informative for developmental studies. In these cases, DFMs that can handle nonstationary and time-varying parameters (e.g., Chow et al., 2011; Molenaar et al., 1992) should be used. In contrast, this limit has the minimal impact on functional connectivity studies where both the time intervals and the entire duration of the brain scan are very short and they do not carry developmental meanings. In sum, dynamic researchers need to be aware of their research purposes and the characteristics of the time series data they have at hand when making the modeling choice.

BIBLIOGRAPHY

Abegaz, F., & Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics*, *14*(3), 586–599.

Asparouhov, T., Hamaker, E. L., & MuthÃÂ©n, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Barrett, A. B., Murphy, M., Bruno, M.-A., Noirhomme, Q., Boly, M., Laureys, S., & Seth, A. K. (2012). Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. *PloS one*, *7*(1), e29072.

Bartlett, M. (2011). The statistical concept of mental factors. *British Journal of Psychology. General Section*, *28*, 97–104. https://doi.org/10.1111/j.2044-8295.1937.tb00863.x

Beltz, A. M., & Molenaar, P. C. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate behavioral research*, *51*(2-3), 357–373.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, *88*(3), 588.

Bollen, K. (1996). An alternative two stage least squares (2sls) estimator for latent variable equations. *Psychometrika*, *61*(1), 109–121. https://EconPapers.repec.org/RePEc:spr:psycho:v:61:y:1996:i:1:p:109-121

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, *17*(3), 303–316.

Bollen, K. A., Fisher, Z. F., Giordano, M. L., Lilly, A. G., Luo, L., & Ye, A. (2021). An introduction to model implied instrumental variables using two stage least squares (miiv-2sls) in structural equation models (sems). *Psychological methods*. https://doi.org/10.1037/met0000297

Bollen, K. A., Gates, K. M., & Fisher, Z. (2018). Robustness conditions for miiv-2sls when the latent variable or measurement model is structurally misspecified [PMID: 30573943]. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(6), 848–859. https://doi.org/10.1080/10705511.2018.1456341

Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., & Chen, F. (2007). Latent variable models under misspecification: Two-stage least squares (2sls) and maximum likelihood (ml) estimators. *Sociological Methods & Research*, *36*(1), 48–86. https://doi.org/10.1177/0049124107301947

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology [PMID: 23537483]. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Boschloo, L., van Borkulo, C. D., Rhemtulla, M., Keyes, K. M., Borsboom, D., & Schoevers, R. A. (2015). The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PLOS ONE*, *10*(9), 1–12. https://doi.org/10.1371/journal.pone.0137621

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological methods*, *22*(3), 409.

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.

Bringmann, L., Lemmens, L., Huibers, M., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the beck depression inventory-ii. *Psychological medicine*, *45*(4), 747–757.

Browne, M., & Nesselroade, J. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of arma time series models.

Chen, G., Glen, D. R., Saad, Z. S., Hamilton, J. P., Thomason, M. E., Gotlib, I. H., & Cox, R. W. (2011). Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis. *Computers in biology and medicine*, *41*(12), 1142–1155.

Chou, C.-P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, lagrange multiplier, and wald tests. *Multivariate Behavioral Research*, *25*(1), 115–136.

Chou, C.-P., & Huh, J. (2012). Model modification in structural equation modeling.

Chow, S.-M., Ho, M.-h. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(2), 303–332. https://doi.org/10.1080/10705511003661553

Chow, S.-M., Zu, J., Shifren, K., & Zhang, G. (2011). Dynamic factor analysis models with time-varying parameters [PMID: 26741330]. *Multivariate Behavioral Research*, *46*(2), 303–339. https://doi.org/10.1080/00273171.2011.563697

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks. a tutorial on recent developments in network psychometrics [Dynamic Personality Psychology]. *Personality and Individual Differences*, *136*, 68–78. https://doi.org/https://doi.org/10.1016/j.paid.2017.06.011

Craddock, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B. S., Lewis, J. D., Li, Q., Milham, M., Yan, C., & Bellec, P. (n.d.). The neuro bureau preprocessing initiative: Open sharing of preprocessed neuroimaging data and derivatives

. *Frontiers in Neuroinformatics*, (41). https://doi.org/10.3389/conf.fninf.2013.09.00041

Cramer, A. O., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and brain sciences*, *33*(2-3), 137–150.

Croon, M. (2002). Using predicted latent scores in general latent structure models. *Top.*

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., et al. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*(6), 659–667.

Eichler, M. (2005). A graphical approach for evaluating effective connectivity in neural systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1457), 953–967.

Elwert, F. (2013). Graphical causal models. In S. L. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 245–273). Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable [PMID: 30111904]. *Annual Review of Sociology*, *40*(1), 31–53. https://doi.org/10.1146/annurev-soc-071913-043455

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural equation modeling*, *8*(3), 430–457.

Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica*, *51*(2), 277–304. http://www.jstor.org/stable/1911990

Epskamp, S. (2018). *Graphicalvar: Graphical var for experience sampling data* [R package version 0.2.2]. https://CRAN.R-project.org/package=graphicalVAR

Epskamp, S. (2020a). Psychometric network models from time-series and panel data. *Psychometrika*, *85*. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S. (2020b). *Psychonetrics: Structural equation modeling and confirmatory network analysis*. http://psychonetrics.org/

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2016). A primer on estimating regularized psychological networks. *arXiv preprint arXiv:1607.01367*.

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.

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of consulting and clinical psychology*, *83*(4), 825.

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment*, *23*(4), 496–506.

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of abnormal psychology*, *126*(8), 1044.

Fisher, Z. F., Bollen, K. A., & Gates, K. M. (2019). A limited information estimator for dynamic factor models [PMID: 30829065]. *Multivariate Behavioral Research*, *54*(2), 246–263. https://doi.org/10.1080/00273171.2018.1519406

Fisher, Z. F., Bollen, K. A., Gates, K. M., & Rönkkö, M. (2020). Model implied instrumental variable (miiv) estimation of structural equation models [r package miivsem version 0.5.5].

Fisher, Z., Kim, Y., Fredrickson, B., & Pipiras, V. (2022). Penalized estimation and forecasting of multiple subject intensive longitudinal data [Funding Information: Vladas Pipiras was supported in part by the NSF grant DMS-1712966. Publisher Copyright: © 2021, The Author(s) under exclusive licence to The Psychometric Society.]. *Psychometrika*. https://doi.org/10.1007/s11336-021-09825-7

Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2017). Further evidence that psychopathology networks have limited replicability and utility: Response to borsboom et al. (2017) and steinley et al. (2017). *Journal of abnormal psychology*, *126*(7), 1011–1016. https://doi.org/10.1037/abn0000313

Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. https://doi.org/10.17605/OSF.IO/6CTS9

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Friedman, J., Hastie, T., & Tibshirani, R. (2019). *Glasso: Graphical lasso: Estimation of gaussian graphical models* [R package version 1.11]. https://CRAN.R-project.org/package=glasso

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441.

Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with granger causality and dynamic causal modelling. *Current opinion in neurobiology*, *23*(2), 172–178.

Gates, K., Fisher, Z., & Bollen, K. (2019). Latent variable gimme using model implied instrumental variables (miivs). *Psychological Methods*, *25*. https://doi.org/10.1037/met0000229

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A monte carlo evaluation of weighted community detection algorithms. *Frontiers in Neuroinformatics*, *10*, 45.

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guiskewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate behavioral research*, *52*(2), 129–148.

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319.

Gates, K. M., Molenaar, P. C., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fmri connectivity mapping: An alternative to granger causality testing using formal equivalences among sem path modeling, var, and unified sem. *NeuroImage*, *50*(3), 1118–1125.

Gates, K. M., Molenaar, P. C., Hillary, F. G., & Slobounov, S. (2011). Extended unified sem approach for modeling event-related fmri data. *NeuroImage*, *54*(2), 1151–1158.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American statistical association*, *77*(378), 304–313.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Greenland, S., Robins, J. M., & Pearl, J. (1999). Confounding and collapsibility in causal inference. *Statist. Sci.*, *14*(1), 29–46. https://doi.org/10.1214/ss/1009211805

Hallquist, M., Wright, A. G., & Molenaar, P. C. M. (2019). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. https://doi.org/10.31234/osf.io/pg4mf

Hamaker, E., Dolan, C., & Molenaar, P. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, *40*. https://doi.org/10.1207/s15327906mbr4002_3

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2002). On the nature of sem estimates of arma parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(3), 347–368. https://doi.org/10.1207/S15328007SEM0903{_}3
doi: 10.1207/S15328007SEM0903_3

Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). Princeton university press Princeton, NJ.

Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: The lasso and generalizations*. Chapman; Hall/CRC.

Henry, T. R., & Ye, A. (2019). Endogeneity in network psychometrics. osf.io/srd3a

Hillary, F. G., Roman, C. A., Venkatesan, U., Rajtmajer, S. M., Bajo, R., & Castellanos, N. D. (2015). Hyperconnectivity is a fundamental response to neurological disruption. *Neuropsychology*, *29*(1), 59.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Huang, J., Ma, S., & Zhang, C.-H. (2008). Adaptive lasso for sparse high-dimensional regression models. *Statistica Sinica*, *18*(4), 1603–1618. http://www.jstor.org/stable/24308572

Huang, P.-H. (n.d.). Lslx: Semi-confirmatory structural equation modeling via penalized likelihood. *Journal of Statistical Software*.

Huang, P.-H. (2020). Postselection inference in structural equation modeling [PMID: 31304796]. *Multivariate Behavioral Research*, *55*(3), 344–360. https://doi.org/10.1080/00273171.2019.1634996

Huang, P.-H., Chen, H., & Weng, L.-J. (2017). A penalized likelihood method for structural equation modeling. *psychometrika*, *82*(2), 329–354.

Jacobucci, R. (2017). Regsem: Regularized structural equation modeling.

Jacobucci, R., Grimm, K. J., Brandmaier, A. M., Serang, S., Kievit, R. A., & Scharf, F. (2019). *Regsem: Regularized structural equation modeling* [R package version 1.3.9]. https://CRAN.R-project.org/package=regsem

Jacobucci, R., Grimm, K. J., & McArdle, J. J. (2016). Regularized structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, *23*(4), 555–566.

Jöreskog, K. G., & Sörbom, D. (1981). *Lisrel 5: Analysis of linear structural relationships by maximum likelihood and least squares methods;[user's guide]*. University of Uppsala.

Jöreskog, K. G., & Sörbom, D. (1986). *Lisrel vi: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Scientific Software.

Kaplan, D. (1988). The impact of specification error on the estimation, testing, and improvement of structural equation models. *Multivariate Behavioral Research*, *23*(1), 69–86.

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional mri data. *Human Brain Mapping*, *28*(2), 85–93.

Koopmans, T. C., Hendry, D. F., & Morgan, M. S. (1995). When is an equation system complete for statistical purposes? (in statistical inference in dynamic economic models, t. c. koopmans (ed.), cowles commission monograph 10, wiley, new york, 1950, pp. 393â409). *The foundations of econometric analysis* (527â537). Cambridge University Press. https://doi.org/10.1017/CBO9781139170116.047

Lane, S. (2017). Regularized structural equation modeling for individual-level directed functional connectivity.

Lane, S., Gates, K., Fisher, Z., Arizmendi, C., Molenaar, P., Hallquist, M., Pike, H., Henry, T., Duffy, K., Luo, L., & Beltz, A. (2019). *Gimme: Group iterative multiple model estimation* [R package version 0.6-1]. https://github.com/GatesLab/gimme/

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*, *24*(1), 54.

Lauritzen, S. L. (1996). *Graphical models* (Vol. 17). Clarendon Press.

Logothetis, N. K. (2008). What we can do and what we cannot do with fmri. *Nature*, *453*(7197), 869–878.

Luo, L., Gates, Z. F., Fisher, Arizmendi, C., Molenaar, P. C. M., & Beltz, K. M., Adriene Gates. (Under Review). Estimating both directed and bidirectional contemporaneous relations in time series data using hybrid-gimme. *Psychological Methods*.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological bulletin*, *100*(1), 107.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological bulletin*, *111*(3), 490.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological reports*, *66*(1), 195–244.

Molenaar, P. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181–202. https://doi.org/10.1007/BF02294246

Molenaar, P., De Gooijer, J. G., & Schmitz, B. (1992). Dynamic factor model analysis of nonstationary multivariate time series. *Psychometrika*, *57*, 333–349. https://doi.org/10.1007/BF02295422

Molenaar, P. C. (2017). Equivalent dynamic models. *Multivariate behavioral research*, *52*(2), 242–258.

Molenaar, P. C. (2019). Granger causality testing with intensive longitudinal data. *Prevention Science*, *20*(3), 442–451.

Molenaar, P. C., & Lo, L. L. (2016). Alternative forms of granger causality, heterogeneity and non-stationarity. *W., Wiedermann, A. von Eye,(Eds.), Statistics and causality: Methods for applied empirical research*, 205–230.

Molenaar, P. C., & Nesselroade, J. R. (1998). A comparison of pseudo-maximum likelihood and asymptotically distribution-free dynamic factor analysis parameter estimation in fitting covariance-structure models to block-toeplitz matrices representing single-subject multivariate time-series [PMID: 26782717]. *Multivariate Behavioral Research*, *33*(3), 313–342. https://doi.org/10.1207/s15327906mbr3303_1

Murphy, M., Bruno, M.-A., Riedner, B. A., Boveroux, P., Noirhomme, Q., Landsness, E. C., Brichant, J.-F., Phillips, C., Massimini, M., Laureys, S., et al. (2011). Propofol anesthesia and sleep: A high-density eeg study. *Sleep*, *34*(3), 283–291.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & Van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological medicine*, *39*(9), 1533–1547.

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Focus article: Idiographic filters for psychological constructs. *Measurement: Interdisciplinary Research and Perspectives*, *5*(4), 217–235. https://doi.org/10.1080/15366360701741807

Nichols, T. T., Gates, K. M., Molenaar, P. C., & Wilson, S. J. (2014). Greater bold activity but more efficient connectivity is associated with better cognitive performance within a sample of nicotine-deprived smokers. *Addiction biology*, *19*(5), 931–940.

Pearl, J. (2009). *Causality*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Price, R. B., Lane, S., Gates, K., Kraynak, T. E., Horner, M. S., Thase, M. E., & Siegle, G. J. (2017). Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biological psychiatry*, *81*(4), 347–357.

Pruttiakaravanich, A., & Songsiri, J. (2018). Convex formulation for regularized estimation of structural equation models.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Ram, N., & Gerstorf, D. (2009). Time-structured and net intraindividual variability: Tools for examining the development of dynamic characteristics and processes. *Psychology and aging*, *24*(4), 778.

Ramsey, J. D., Hanson, S. J., & Glymour, C. (2011). Multi-subject search correctly identifies causal connections and most causal directions in the dcm models of the smith et al. simulation study. *NeuroImage*, *58*(3), 838–848.

Rhemtulla, M., Fried, E. I., Aggen, S. H., Tuerlinckx, F., Kendler, K. S., & Borsboom, D. (2016). Network analysis of substance abuse and dependence symptoms. *Drug and Alcohol Dependence*, *161*, 230–237. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2016.02.005

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. http://www.jstatsoft.org/v48/i02/

Rothman, A. J., Levina, E., & Zhu, J. (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics*, *19*(4), 947–962.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, *26*(3), 393–415. http://www.jstor.org/stable/1907619

Schmittmann, V. D., Cramer, A. O., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena [On defining and interpreting constructs: Ontological and epistemological constraints]. *New Ideas in Psychology*, *31*(1), 43–53. https://doi.org/https://doi.org/10.1016/j.newideapsych.2011.02.007

Shapiro, M. D., & Watson, M. W. (1988). *Sources of business cycle fluctuations* (Working Paper No. 2589). National Bureau of Economic Research. https://doi.org/10.3386/w2589

Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With r examples*. Springer.

Sims, C. A. (1981). An autoregressive index model for the u.s., 1948-1975.

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563–575. https://EconPapers.repec.org/RePEc:spr:psycho:v:66:y:2001:i:4:p:563-575

Smith, S. M. (2012). The future of fmri connectivity. *Neuroimage*, *62*(2), 1257–1266.

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., & Woolrich, M. W. (2011). Network modelling methods for fmri. *Neuroimage*, *54*(2), 875–891.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, *25*(2), 173–180.

Strotz, R. H., & Wold, H. O. A. (1960). Recursive vs. nonrecursive systems: An attempt at synthesis (part i of a triptych on causal chain systems). *Econometrica*, *28*(2), 417–427. http://www.jstor.org/stable/1907731

Thurstone, L. L. (1935). *The vectors of mind*. University of Chicago Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/https://doi.org/10.1006/nimg.2001.0978

Varoquaux, G., & Craddock, R. C. (2013). Learning and comparing functional connectomes across subjects. *NeuroImage*, *80*, 405–415.

Weigard, A., Lane, S., Gates, K., & Beltz, A. (Under Review). The influence of autoregressive relation strength and search strategy on directionality recovery in gimme.

Wigman, J., Van Os, J., Borsboom, D., Wardenaar, K., Epskamp, S., Klippel, A., Viechtbauer, W., Myin-Germeys, I., Wichers, M., et al. (2015). Exploring the underlying structure of mental disorders: Cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological medicine*, *45*(11), 2375–2387.

Wild, B., Eichler, M., Friederich, H.-C., Hartmann, M., Zipfel, S., & Herzog, W. (2010). A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC medical research methodology*, *10*(1), 28.

Wright, A. G., Beltz, A. M., Gates, K. M., Molenaar, P., & Simms, L. J. (2015). Examining the dynamic structure of daily internalizing and externalizing behavior at multiple levels of analysis. *Frontiers in psychology*, *6*, 1914.

Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: An fmri study. *Journal of Neurolinguistics*, *33*, 29–49.

Ye, A., Gates, K. M., Henry, T. R., & Luo, L. (2021). Path and directionality discovery in individual dynamic models: A regularized unified structural equation modeling approach for hybrid vector autoregression. *Psychometrika*, *86*(2), 404–441. https://doi.org/10.1007/s11336-021-09753-6

Zhang, Z., Hamaker, E., & Nesselroade, J. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling A Multidisciplinary Journal*, *15*, 377–402. https://doi.org/10.1080/10705510802154281

Zou, H., & Hastie, T. (n.d.). Tibshirani., r.,(2004). *Sparse Principal Component Analysis, Technical Report, Statistics Department, Stanford University.*

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.