

Problems

- Instability in causal model estimation.
- An immense number of possible causal model, i.e., $3^{n(n-1)/2}$, where n is the number of variables.

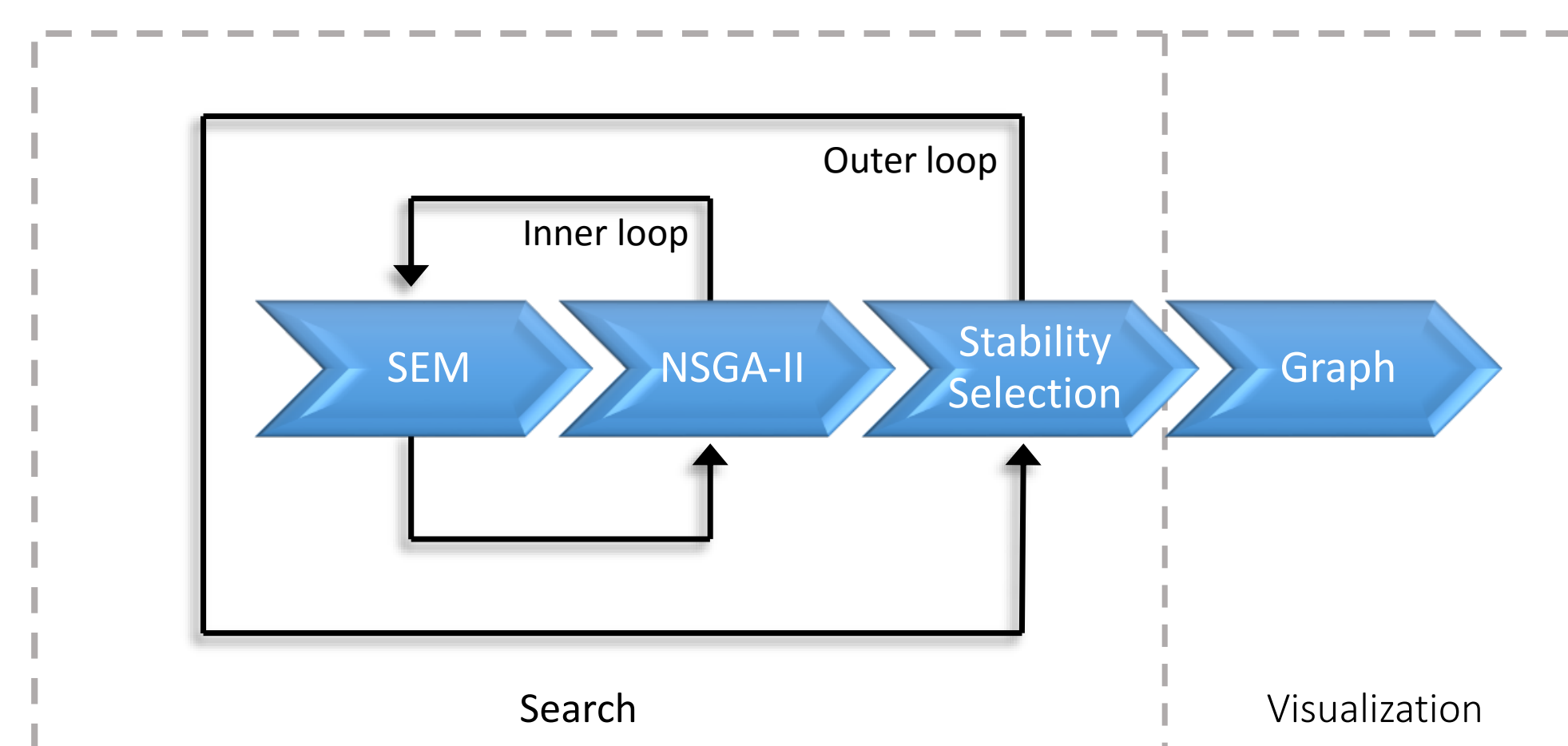
Objectives

To develop a causal method that aims to resolve the problems of an immense number of causal models and instability in model estimation.

Proposed method

We introduce stable specification search for cross-sectional data (S3C), which is a combination of:

- Non-dominated sorting genetic algorithm-II (NSGA-II)
- Stability selection
- Structural equation model (SEM)



Extensions:

- 1 To longitudinal data (S3L)
- 2 To latent model (S3C-Latent)
- 3 To survival causal model

Results

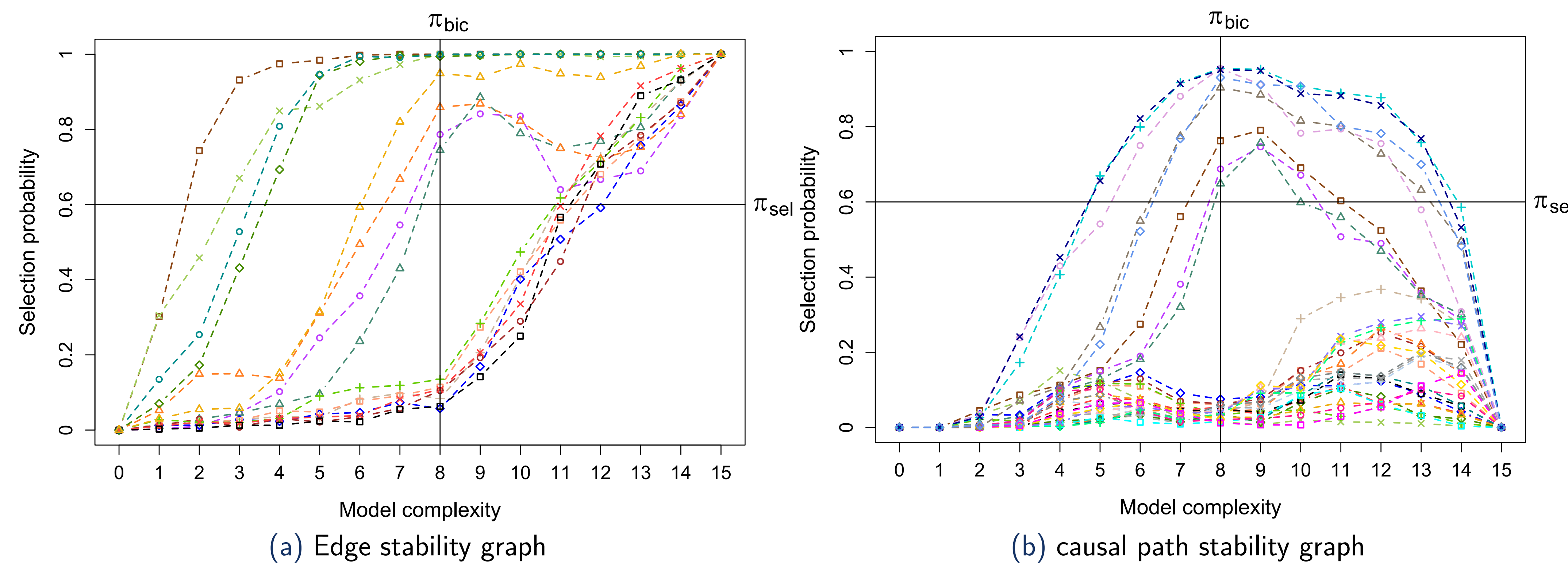


Figure: Causal structures with selection probability $\geq \pi_{sel}$ and model complexity $\leq \pi_{bic}$ are considered relevant (stable & parsimonious).

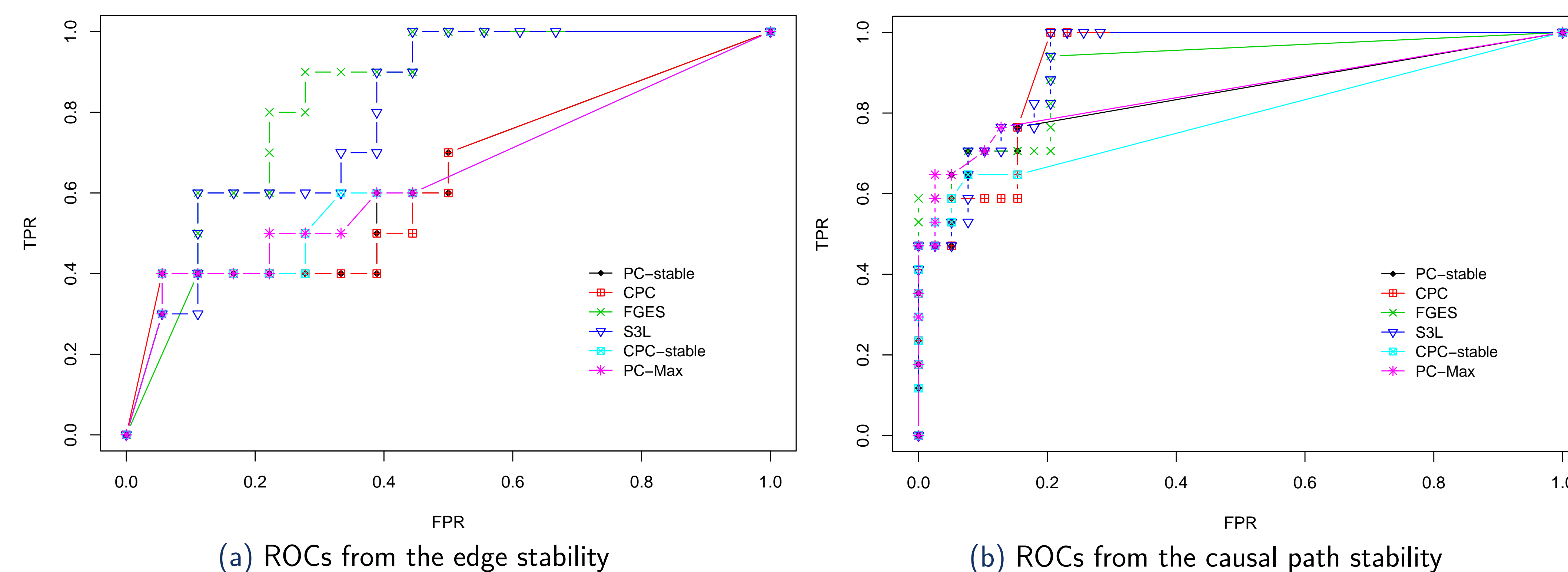


Figure: Comparison of ROCs of S3L (an extension of S3C to longitudinal data) to other methods

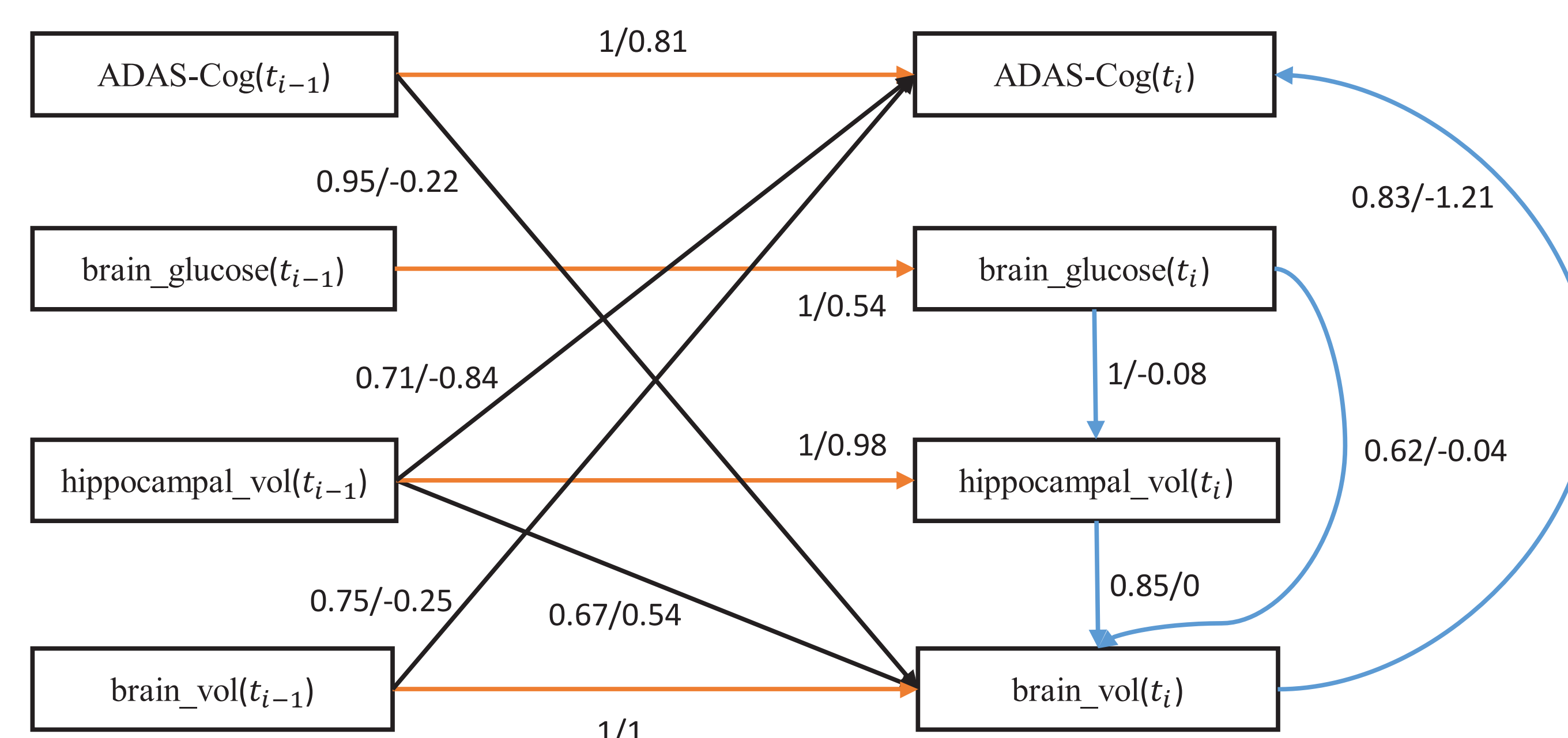


Figure: The longitudinal causal model of patients with the Alzheimer's disease.

Applications to real-world data

Chronic fatigue syndrome, chronic kidney disease, Alzheimer's disease, Mental abilities, Tuberculous meningitis, ADHD.

R Package

We implemented S3C as an R package
<https://cran.r-project.org/web/packages/stablespec/index.html>

Publications

R. Rahmadi, P. Groot, M. Heins, H. Knoop, T. Heskes and the OPTIMISTIC consortium (2017). Causality on cross-sectional data: Stable specification search in constrained structural equation modeling. *Applied Soft Computing*, 52:687–698.

R. Rahmadi, P. Groot, M. HC van Rijn, J. AJG van den Brand, M. Heins, H. Knoop, T. Heskes, the Alzheimer's Disease Neuroimaging Initiative, the MASTERPLAN Study Group and the OPTIMISTIC consortium (2018). Causality on longitudinal data: Stable specification search in constrained structural equation modeling. *Statistical Methods in Medical Research*, 27(12):3814–3834.

R. Rahmadi, P. Groot, T. Heskes (2018). The stablespec package for causal discovery on cross-sectional and longitudinal data in R. *Neurocomputing*, 275:2440–2443.

R. Rahmadi, P. Groot, T. Heskes (2019). Stable specification search in structural equation models with latent variables. *Intelligent Systems and Technology; special issue on causal discovery* (accepted).

Contact Information

Center of Data Science, Informatics, UII
Email: ridho.rahmadi@uii.ac.id