Causal Discovery of DSGE Models

Emmet Hall-Hoffarth

University of Oxford

emmet.hall-hoffarth@economics.ox.ac.uk

November 2, 2020

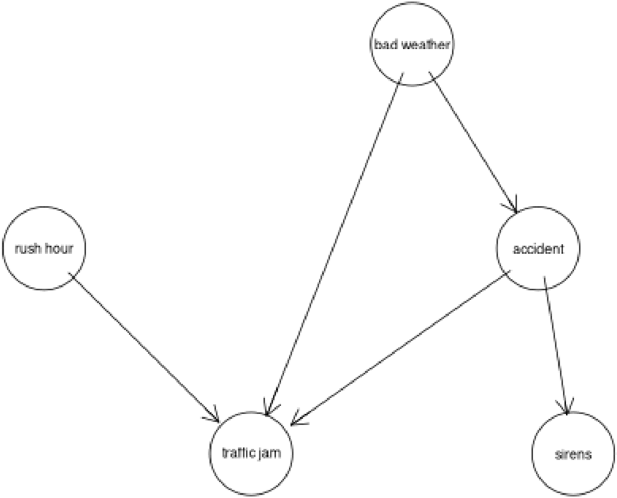
Overview

Introduction

- Causal discovery uses algorithms and data to identify causal models.
- I use this idea to construct a test and an algorithm that identifies log-linear solution to DSGE model from observational data.
- Test and algorithm are asymptotically unique and consistent.
- Performs well in small sample simulations.

What does this mean?

Causal Discovery



DSGE Models

Consumer problem

$$\max \sum_{t=0}^{\infty} \beta^t U(C_t, L_t, ...) \; s.t.$$

$$P_t C_t + ... \leq W_t L_t + ... \; \forall t$$

$$\lim_{T \to \infty} \beta^T \lambda_{x,T} x_T \to 0 \; \forall x \in \{K, B, ...\}$$

Firm problem

$$\max \; P_t f(L_t, ...) - W_t L_t \; \forall t$$

Competitive equilibrium, market clearing conditions ...

State Space Representation

- Log-linear approximation yields the following general solution:

$$\vec{y}_t = \vec{A}\vec{x}_{t-1} + \vec{B}\vec{z}_t \tag{1}$$

$$\vec{x}_t = \vec{C}\vec{x}_{t-1} + \vec{D}\vec{z}_t \tag{2}$$

$$\vec{z}_t = \vec{E}\vec{z}_{t-1} + \vec{\epsilon}_t \tag{3}$$

- Partition variables $\vec{w}_t$ into three categories:
  - $\vec{x}_t$: predetermined or endogenous state variables
  - $\vec{y}_t$: control variables or "jumpers"
  - $\vec{z}_t$: exogenous state variables
- Algorithm identifies partition (and $\vec{A}$ - $\vec{E}$).

Results (I)

- Baseline RBC model.
- 9 Observables, true partition is $\vec{z} = [g \;\; z]$, $\vec{x} = [k]$, $\vec{y} = [y \;\; c \;\; i \;\; r \;\; w \;\; l]$.
- With large sample of 10^6 observations algorithm uniquely identifies true partition.
- 834 models considered (models with $1 \leq \#$ state variables ≤ 3).
- 19683 possible models reduced to just 1.

Results (II)

- 1000 iterations, 100 sample size
- Wins = number of times model selected, Valid = number of times model is valid relative to CI test

Index	Exogenous States	Endogenous States	Wins	Valid
1	g z	k	944	944
2	g w	k	27	729
3	g y	k	27	571
4	c g	k	2	8
5	g l y		0	340
6	g r y		0	421
7	g r	k	0	576
8	g l z		0	716
9	g i r		0	781
10	g i l		0	629
11	g i	k	0	867
12	g r w		0	609
13	g r z		0	858
14	g k l		0	625
15	g l w		0	603
16	g k r		0	779
17	c g w		0	1

Why is this useful?

Context

"The most important issue holding back DAGs is the lack of convincing empirical applications. History suggests that those are what is driving the adoption of new methodologies in economics and other social sciences, not the mathematical elegance or rhetoric."

"The advantages of the formal methods for deriving identification results with DAGs are most pronounced in complex models with many variables that are not particularly popular in empirical economics."

— Imbens (2019)

Reduction of problem space

- Given a set of observations over k variables there are 3^k possible state-space partitions.
  - Choice of model is *trinomial*
- Reduce this to just one (or very small subset).
- Rules out many *invalid* models that the applied macroeconomist would have otherwise considered.
- However, still doesn't reduce to a single DSGE model.

Agnosticism

- Assumptions we use are very minimal:
  - True DGP is a log-linear DSGE model.
  - *Casual Markov Assumption* — All variables are independent of their non-descendants given their parents (in a DAG) (Spirtes & Zhang, 2016).
  - *Faithfulness* — observed conditional independence are real features of the DGP.
  - Linearity, Gaussian shocks — these can be relaxed.
- As a result the solution produced reflects the data to the greatest extent possible.

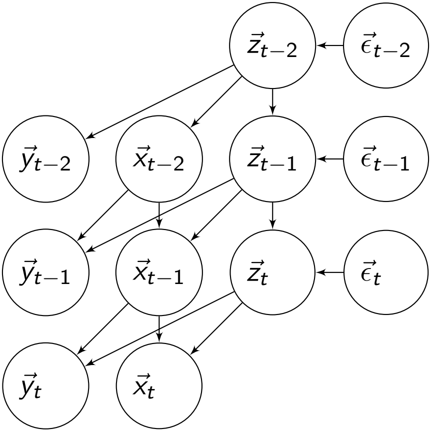
Inference

- Cannot pin down unique DSGE, but can still make some inferences about possible models.
- Is consumption predetermined? If so then habits important to explain behaviour.
- Is inflation predetermined? If so then it is probably not fully rational / forward looking.
- These types of inferences are particularly meaningful because of agnosticism.

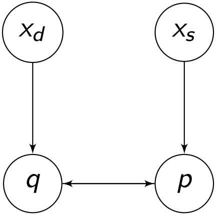
How is this even possible?

DAGs

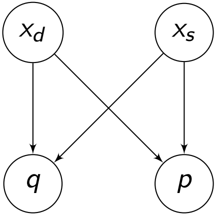
- We can represent the state-space solution as a DAG.



(Aside) Is this sensible?



$\iff$



$$p = \alpha_p + \beta_{ps} x(s) + \beta_{pq} q + \epsilon_p$$
$$q = \alpha_q + \beta_{qd} x(d) + \beta_{qp} p + \epsilon_q$$

$$p = \frac{1}{1 - \beta_{pq}\beta_{qp}}[(\alpha_p + \beta_{pq}\alpha_q) + \beta_{ps}x(s) + \beta_{pq}\beta_{qd}x(d) + (\epsilon_p + \beta_{pq}\epsilon_q)]$$
$$q = \frac{1}{1 - \beta_{qp}\beta_{pq}}[(\alpha_q + \beta_{qp}\alpha_p) + \beta_{qd}x(d) + \beta_{qp}\beta_{ps}x(s) + (\epsilon_q + \beta_{qp}\epsilon_p)]$$

- Understanding response to shocks does not require inference of "deep" parameters (β_{xy}).

Conditional Independence (I)

- Given the Causal Markov assumption we can infer the following relationships from this DAG:

$$x_t \perp\!\!\!\perp x_t' \,||\, [\vec{x}_{t-1}, \vec{z}_t] \text{ for all } (x_t, x_t') \in [\vec{x}_t, \vec{y}_t] \tag{4}$$

$$x_{t-1} \perp\!\!\!\perp z_t \,||\, \vec{z}_{t-1} \text{ for all } x_{t-1} \in \vec{x}_{t-1} \text{ and } z_t \in \vec{z}_t \tag{5}$$

$$x_t \perp\!\!\!\perp z_{t-1} \,||\, [\vec{x}_{t-1}, \vec{z}_t] \text{ for all } x_t \in [\vec{x}_t, \vec{y}_t] \text{ and } z_{t-1} \in \vec{z}_{t-1} \tag{6}$$

$$z_t \perp\!\!\!\perp z_t' || \vec{z}_{t-1} \text{ for all } z_t \neq z_t' \in \vec{z}_t \tag{7}$$

- (4), (5) and the minimum state variable criterion (MSV) (McCallum, 1999) uniquely identify state-space model.
- I refer to this as *validity*.

Conditional Independence (II)

- In reality conditional independence is not known, so we implement a feasible test.
- Gaussian variables so conditional independence $\iff$ 0 correlation (of partial residuals).
- Possible to combine the 4 conditions into a single test.
- In particular, test whether $\vec{x}_{t-1}$, $\vec{y}_{t-1}$, $\vec{z}_t$, and (optionally) $\vec{z}_{t-2}$ are *completely independent* (covariance matrix is diagonal), conditional on $\vec{x}_{t-2}$ and $\vec{z}_{t-1}$.
- Use test for this from Srivastava (2005).

Scoring

- In small samples power is limited so more than one model can be *valid*.
- Use score function (Likelihood/AIC/BIC) rank models and pick one.
- Could just maximise score function over all models, but performs worse than combined approach.

Algorithm

- To find the valid model(s), iterate over all possible models and apply the test (brute force).
- If we find more than one valid model choose a winner using the score function.
- This algorithm will consistently identify the unique valid model as $n \to \infty$ because the power of the CI test to reject incorrect models goes to 1.
  - We will still have a significance level α probability of rejecting the correct model.
  - α is the only tuning parameter of this algorithm.

Limitations

- Does not solve microeconomic dissonance (Levin et al., 2008).
- Type II error can be a problem on realistic data sizes.
- Computational complexity

Conclusion

- I introduce a test and algorithm to identify a unique state-space model (and associated family of DSGE models) which is *valid* relative to some observed data.
- This procedure is asymptotically consistent and empirical results show it performs well on realistic sample sizes.
- One approach to applying causal discovery tools in the context of economics.

Citations

Imbens, G. (2019). *Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics* (tech. rep.). National Bureau of Economic Research.

Levin, A. T., López-Salido, J. D., Nelson, E., & Yun, T. (2008). Macroeconometric equivalence, microeconomic dissonance, and the design of monetary policy. *Journal of Monetary Economics*, *55*, S48–S62.

McCallum, B. T. (1999). Role of the minimal state variable criterion in rational expectations models. *International finance and financial crises* (pp. 151–176). Springer.

Spirtes, P., & Zhang, K. (2016). Causal discovery and inference: Concepts and recent methodological advances. *Applied informatics*, *3*(1), 3.

Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of the Japan Statistical Society*, *35*(2), 251–272.