Causal Discovery of DSGE Models

Emmet Hall-Hoffarth

University of Oxford

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

April 29, 2021

Overview

- Causal discovery uses algorithms and data to identify causal models.
- Suppose that some observed data can be rationalised by some DSGE model. I use some concepts from this literature to provide a test and algorithm to identify the log-linear (state-space) approximation of that model.
- Test and algorithm are asymptotically consistent for a unique solution.
- Performs well in small sample simulations and real data.

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:

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

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

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

- 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?

- Reduction in the problem space
  - Given a set of observations over k variables there are 3^k possible state-space partitions. Reduce this to just one (or very small subset).
  - However, still doesn't reduce to a single structural model (microeconomic dissonance).
- Agnosticism
  - Assume nothing in particular about each variable — treat them all the same.
  - As a result the solution produced reflects the data to the greatest extent possible.
- Inference
  - 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.

(Massively Oversimplified) Methodology

- State-Space model is also a DAG, therefore, it implies a set of conditional independence relationships:

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

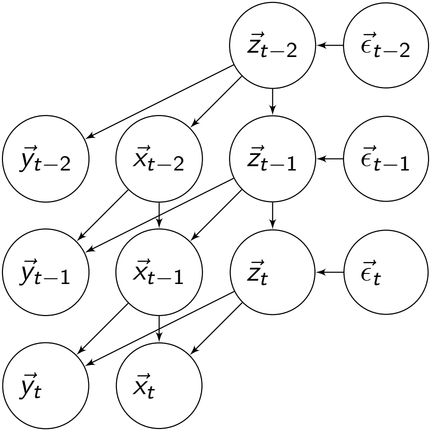
$$x_{t-1} \perp\!\!\!\perp z_t \,||\, \vec{z}_{t-1} \text{ for all } x_{t-1} \in \vec{x}_{t-1} \text{ and } z_t \in \vec{z}_t \tag{5}$$

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

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

- We test for these conditional independence relationships in the data.
- Iterate over all possible models until we find one that passes these tests.

DAGs


$\vec{z}_{t-2}$
$\vec{\epsilon}_{t-2}$
$\vec{y}_{t-2}$
$\vec{x}_{t-2}$
$\vec{z}_{t-1}$
$\vec{\epsilon}_{t-1}$
$\vec{y}_{t-1}$
$\vec{x}_{t-1}$
$\vec{z}_{t}$
$\vec{\epsilon}_{t}$
$\vec{y}_{t}$
$\vec{x}_{t}$

Github

https://github.com/e-hall-hoffarth/causal_dsge