Simultaneous Causality: Part IV on Causality

James J. Heckman

Econ 312, Spring 2019

Econometric Causality Entertains the Possibility of Simultaneous Causality

Nonrecursive (Simultaneous) Models of Causality: Developed in Economics (Haavelmo, 1944)

- A system of linear simultaneous equations captures interdependence among outcomes Y.

- Linear model in terms of parameters (Γ, B), observables (Y, X) and unobservables U:

$$\Gamma Y + BX = U, \qquad E(U) = 0, \tag{1}$$

- Y is now a vector of internal and interdependent variables
- X is external and exogenous $(E(U \mid X) = 0)$
- Γ is a full rank matrix.

- This is a linear-in-the-parameters "all causes" model for vector Y, where the causes are X and $\mathcal{E}$.
- The "structure" is (Γ, B), Σ_U, where Σ_U is the variance-covariance matrix of U.
- In the Cowles Commission analysis it is assumed that Γ, B, Σ_U are invariant to general changes in X and translations of U.
- Autonomy (Frisch, 1938) also "SUTUA" in Holland, 1986

Nonlinear Systems Possible

- Thus we can postulate a system of equations $G(Y, X, U) = 0$ and develop conditions for unique solution of reduced forms $Y = K(X, U)$ requiring that certain Jacobian terms be nonvanishing.
- See Heckman et al. (2010).
- The structural form (1) is an all causes model that relates in a deterministic way outcomes (internal variables) to other outcomes (internal variables) and external variables (the X and U).
- **Are ceteris paribus manipulations associated with the effect of some components of Y on other components of Y possible within the model?**

- Consider a two-agent model of social interactions.
- Y_1 is the outcome for agent 1; Y_2 is the outcome for agent 2.

- $$\begin{aligned} Y_1 &= \alpha_1 + \gamma_{12} Y_2 + \beta_{11} X_1 + \beta_{12} X_2 + U_1, \qquad (2a) \\ Y_2 &= \alpha_2 + \gamma_{21} Y_1 + \beta_{21} X_1 + \beta_{22} X_2 + U_2. \qquad (2b) \end{aligned}$$

- Social interactions model is a standard version of the simultaneous equations problem.
- This model is sufficiently flexible to capture the notion that the consumption of 1 (Y_1) depends on the consumption of 2 if $\gamma_{12} \neq 0$, as well as 1's value of X if $\beta_{11} \neq 0$, X_1 (assumed to be observed), 2's value of X , X_2 if $\beta_{12} \neq 0$ and unobservable factors that affect 1 (U_1).
- The determinants of 2's consumption are defined symmetrically.
- Allow U_1 and U_2 to be freely correlated.
- Captures essence of "reflection problems."

- Assume

$$E\left(U_1 \mid X_1, X_2\right) = 0 \tag{3a}$$

and

$$E\left(U_2 \mid X_1, X_2\right) = 0. \tag{3b}$$

- Completeness guarantees that (2a) and (2b) have a determinate solution for (Y_1, Y_2).
- Applying Haavelmo's (1943) analysis to (2a) and (2b), the causal effect of Y_2 on Y_1 is γ_{12}.
- This is the effect on Y_1 of fixing Y_2 at different values, holding constant the other variables in the equation.

- Symmetrically, the causal effect of Y_1 on Y_2 is γ_{21}.
- Conditioning, i.e., using least squares, in general, fails to identify these causal effects because U_1 and U_2 are correlated with Y_1 and Y_2.
- This is a traditional argument.
- It is based on the correlation between Y_2 and U_1 (Haavelmo, 1943).
- But even if $U_1 = 0$ and $U_2 = 0$, so that there are no unobservables, least squares breaks down because Y_2 is perfectly predictable by X_1 and X_2.
- We cannot simultaneously vary Y_2, X_1 and X_2.
- Correlation with "*non-identifiability*;" the error is not the fundamental source of non-identifiability with the models.

- Under completeness, the reduced form outcomes of the model after social interactions are solved out can be written as

$$
\begin{aligned}
Y_1 &= \pi_{10} + \pi_{11} X_1 + \pi_{12} X_2 + \mathcal{E}_1, &\text{(4a)}\\
Y_2 &= \pi_{20} + \pi_{21} X_1 + \pi_{22} X_2 + \mathcal{E}_2. &\text{(4b)}
\end{aligned}
$$

- Least squares can identify the *ceteris paribus* effects of X_1 and X_2 on Y_1 and Y_2 because $E(\mathcal{E}_1 \mid X_1, X_2) = 0$ and $E(\mathcal{E}_2 \mid X_1, X_2) = 0$.
- Simple algebra:

$$\pi_{11} = \frac{\beta_{11} + \gamma_{12}\beta_{21}}{1 - \gamma_{12}\gamma_{21}}, \qquad \pi_{12} = \frac{\beta_{12} + \gamma_{12}\beta_{22}}{1 - \gamma_{12}\gamma_{21}}, \pi_{21} = \frac{\gamma_{21}\beta_{11} + \beta_{21}}{1 - \gamma_{12}\gamma_{21}},$$

and

$$\begin{aligned} \mathcal{E}_1 &= \frac{U_1 + \gamma_{12} U_2}{1 - \gamma_{12}\gamma_{21}}, \\ \mathcal{E}_2 &= \frac{\gamma_{21} U_1 + U_2}{1 - \gamma_{12}\gamma_{21}}. \end{aligned}$$

- **Without any further information on the variances of (U_1, U_2) and their relationship to the causal parameters, we cannot identify the causal effects γ_{12} and γ_{21} from the reduced form regression coefficients.**
- This is so because holding X_1, X_2, U_1 and U_2 fixed in (2a) or (2b), it is not possible to vary Y_2 or Y_1, respectively, because they are exact functions of X_1, X_2, U_1 and U_2.
- This exact dependence holds true even if $U_1 = 0$ and $U_2 = 0$ so that there are no unobservables.

- There is no mechanism yet specified within the model to independently vary the right hand sides of Equations (2a) and (2b).
- The mere fact that we can write (2a) and (2b) means that we "can imagine" independent variation.
- Causality is in the mind.

- We "can imagine" a model

$$Y = \varphi_0 + \varphi_1 X_1 + \varphi_2 X_2,$$

  but if part of the model is $(*)\, X_1 = X_2$, no causal effect of X_1 holding X_2 constant is possible in principle within the rules of the model.
- If we break restriction $(*)$ and permit independent variation in X_1 and X_2, we can define the causal effect of X_1 holding X_2 constant.
- But we can imagine such variation.

- In some conceptualizations, no causality is possible; in others it is.
- Distinguish identification from causation.
- The X effects on Y_1 and Y_2, identified through the reduced forms, combine the direct effects (through β_{ij}) and the indirect effects (as they operate through Y_1 and Y_2, respectively).
- If we assume exclusions ($\beta_{12} = 0$) or ($\beta_{21} = 0$) or both, we can identify the *ceteris paribus* causal effects of Y_2 on Y_1 and of Y_1 on Y_2, respectively, if $\beta_{22} \neq 0$ or $\beta_{11} \neq 0$, respectively.

- Thus if $\beta_{12} = 0$, from the reduced form

$$\frac{\pi_{12}}{\pi_{22}} = \gamma_{12}.$$

If $\beta_{21} = 0$, we obtain

$$\frac{\pi_{21}}{\pi_{11}} = \gamma_{21}.$$

- In a general nonlinear model,

$$\begin{aligned} Y_1 &= g_1 (Y_2, X_1, X_2, U_1) \\ Y_2 &= g_2 (Y_1, X_1, X_2, U_2), \end{aligned}$$

exclusion is defined as $\frac{\partial g_1}{\partial X_1} = 0$ for all (Y_2, X_1, X_2, U_1) and $\frac{\partial g_2}{\partial X_2} = 0$ for all (Y_1, X_1, X_2, U_2).

- Assuming the existence of local solutions, we can solve these equations to obtain

$$\begin{aligned} Y_1 &= \varphi_1(X_1, X_2, U_1, U_2) \\ Y_2 &= \varphi_2(X_1, X_2, U_1, U_2) \end{aligned}$$

- By the chain rule we can write

$$\frac{\partial g_1}{\partial Y_2} = \frac{\partial Y_1}{\partial X_1} \Big/ \frac{\partial Y_2}{\partial X_1} = \frac{\partial \varphi_1}{\partial X_1} \Big/ \frac{\partial \varphi_2}{\partial X_1}.$$

- We may define causal effects for Y_1 on Y_2 using partials with respect to X_2 in an analogous fashion.

- Alternatively, we could assume $\beta_{11} = \beta_{22} = 0$ and $\beta_{12} \neq 0$, $\beta_{21} \neq 0$ to identify γ_{12} and γ_{21}.
- These exclusions say that the social interactions only operate through the Y's.
- Agent 1's consumption depends only on agent 2's consumption and not on his value of X_2.
- Agent 2 is modeled symmetrically versus agent 1.
- Observe that we have *not* ruled out correlation between U_1 and U_2.

- When the procedure for identifying causal effects is applied to samples, it is called **indirect least squares** (Tinbergen, 1930).
- The analysis for social interactions in this section is of independent interest.
- It can be generalized to the analysis of N person interactions if the outcomes are continuous variables.

- The intuition for these results is that if $\beta_{12} = 0$, we can vary Y_2 in Equation (2a) by varying the X_2.
- Since X_2 does not appear in the equation, under exclusion, we can keep U_1, X_1 fixed and vary Y_2 using X_2 in (4b) if $\beta_{22} \neq 0$.
- Notice that we could also use U_2 as a source of variation in (4b) to shift Y_2.
- The roles of U_2 and X_2 are symmetric.
- However, if U_1 and U_2 are correlated, shifting U_2 shifts U_1 unless we control for it.
- The component of U_2 uncorrelated with U_1 plays the role of X_2.

- Symmetrically, by excluding X_1 from(2b), we can vary Y_1, holding X_2 and U_2 constant.
- These results are more clearly seen when $U_1 = 0$ and $U_2 = 0$.

- A hypothetical thought experiment justifies these exclusions.
- If agents do not know or act on the other agent's X, these exclusions are plausible.
- An implicit assumption in using (2a) and (2b) for causal analysis is invariance of the parameters $(\Gamma, \beta, \Sigma_U)$ to manipulations of the external variables.

- This definition of causal effects in an interdependent system generalizes the recursive definitions of causality featured in the statistical treatment effect literature (Holland, 1988, and Pearl, 2009.
- The key to this definition is manipulation of external inputs and exclusion, not randomization or matching.

Frisch, R. (1938). Autonomy of economic relations: Statistical versus theoretical relations in economic macrodynamics. Paper given at League of Nations. Reprinted in D.F. Hendry and M.S. Morgan (1995), *The Foundations of Econometric Analysis*, Cambridge University Press.

Haavelmo, T. (1943, January). The statistical implications of a system of simultaneous equations. *Econometrica 11*(1), 1–12.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica 12*(Supplement), iii–vi and 1–115.

Heckman, J. J., R. L. Matzkin, and L. Nesheim (2010, September). Nonparametric identification and estimation of nonadditive hedonic models. *Econometrica 78*(5), 1569–1591.

Holland, P. W. (1988). Causal inference, path analysis and recursive structural equation models. In C. Clogg and G. Arminger (Eds.), *Sociological Methodology*, pp. 449–484. Washington, DC: American Sociological Association.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys 3*, 96–146.

Tinbergen, J. (1930, October). Bestimmung und deutung von angebotskurven ein beispiel. *Zeitschrift für Nationalökonomie 1*(5), 669–679.