The final publication is available at Springer via http://dx.doi.org/10.1007/s10614-015-9521-8

Is it possible to visualise any stock flow consistent model as a directed acyclic graph?

March 19, 2015

Abstract

Yes it is. We rigorously demonstrate the equivalence of any stock flow consistent model to a directed acyclic graph using condensation graphs. The equivalence between stock flow models and directed acyclic graphs is useful both for visualising large-scale macroeconomic models of this type and for inferring causality within these models. We developed a new package to build and simulate any stock flow consistent model and generate the corresponding directed acyclic graphs, and we provide an example of this package using a well known model from the literature.

Keywords: Stock flow consistent models, directed graphs, macroeconomic modeling.
JEL Codes: E01, E17, E12, E17.

1 Introduction

An open problem in the emerging stock flow consistent macroeconomic literature concerns how best to represent these models graphically and causally. The goal of this paper is to solve both the representational problem and the specification of causal structure problem using directed acyclic graphs. We prove that for any stock flow consistent models there is a corresponding directed acyclic graph. Using a newly developed software package, we show this correspondence in action. Our approach simplifies the process of making any stock flow consistent model, as well as visualising and inferring causality once the model has been built.

Every stock flow consistent (SFC) macroeconomic model is built to mimic the flow of funds data for an individual economy (Godley and Lavoie, 2007). Sectoral interlinkages are explicitly modelled to ensure consistency of stock and flows, so that every quantity comes from somewhere, and goes somewhere (Caverzasi and Godin, 2015). These models often feature precision regarding time, several financial assets and rates of return, budget constraints and adding up constraints (Tobin, 1982). At the representational level, stocks and flows are tracked using balance sheet and transaction matrices. Several hundred behavioural and identity equations build and balance the model. These models are unwieldy as a result. Once the model is defined, it can be calibrated or estimated[1] in order to determine the parameter values. The model is then numerically simulated and eventually shocked in order to compute out-of sample values for the endogenous variables (Godley, 1999).[2] For models estimated from real-world data, it is generally possible to compute confidence intervals around the predicted values of the model using exogenous innovation terms Dos Santos and Zezza (2008).[3] However, even for models developed purely for the purpose of simulation, causal identification can be difficult.

An example may aid intuition. Typically SFC models use linear consumption functions to understand the household sectors consumption decisions from current income and past wealth. In this linear consumption function, causality runs from right to left, so, when disposable income increases, consumption increases. But why should this be? At times it must be the case that increases in consumption cause increases in disposable income as people work more to afford to consume. There is no justification within the model for such a causal choice, other than an appeal to convention and the literature. Using directed acyclic graphs, it is possible to infer the most likely causal structure, especially for behavioural equations which are not as straightforward as consumption functions.

Directed acyclic graphs (DAGs) were developed by Pearl (2000), Morgan and Winship (2007), Lauritzen (2001) and others to capture the insight that in an interacting system, the expression of one variable can cause an effect in other variables. This effect is generally unknown when the behaviour of the system is observed as a whole. In controlled experimental environments variables can of course be isolated and their causality inferred, but

[1] Estimation is the process of discovering constant parameters using econometric methods such as Ordinary Least Squares (OLS), maximum likelihood, bayesian techniques etc. Calibration, on the other hand, consists in finding a value for each parameter, in each period, such that the model replicates the data set.

[2] This process is similar to the more well-known Dynamic Stochastic General Equilibrium class of models, see (Heer and Maussner, 2009, chapter 1) for an introduction.

[3] These exogenous innovation terms will not be included in our formal description of the models

in macroeconomic models such isolation is rarely possible. The aim of most directed acyclic graph modelling is to reverse engineer causality by inferring variable interactions from observations of the entire system. In economics Bessler et al. (2003) and Hoover (2001) have papers applying the directed graph concept to economic issues.

In this paper we make three contributions to the literature. First, we show that every stock flow consistent model has a corresponding directed acyclic graph. Second, we demonstrate how to infer causality directly from the graphical representation of the model. Third, we have developed software to simulate any stock flow consistent model and show its directed acyclic graph.

The rest of the paper is organized as follows. The equivalence between any stock flow consistent model and its DAG is derived in Section 2, while an example is given in Section 3 of the *BMW* model, developed by Godley and Lavoie (2007). Finally, we conclude and give directions for future research.

2 Marrying DAGs to SFC Models

Definition A *stock flow consistent model* is a macroeconomic model based on both the balance sheet of an economy and a transaction flow matrix T which describes the monetary transactions (current account or capital account) between all sectors of that economy. The model variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ (ex. $(C, G, I, \dots)$) are the non-zero entries of T and the relationships between the variables are described by the system of equations $\mathbf{X} = \mathbf{f}(\mathbf{X})$ i.e.

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{X}) \\ f_2(\mathbf{X}) \\ \vdots \\ f_3(\mathbf{X}) \end{pmatrix} = \mathbf{f}(\mathbf{X}). \tag{1}$$

These equations ensure consistency of stocks and flows over time.

A SFC model can be converted to a directed graph in the following manner. From the set of equations of the SFC we define the Jacobian as

$$\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial f_1}{\partial X_1} & \frac{\partial f_1}{\partial X_2} & \cdots & \frac{\partial f_1}{\partial X_n} \\ \frac{\partial f_2}{\partial X_1} & \frac{\partial f_2}{\partial X_2} & \cdots & \frac{\partial f_2}{\partial X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial X_1} & \frac{\partial f_n}{\partial X_2} & \cdots & \frac{\partial f_n}{\partial X_n} \end{pmatrix}. \tag{2}$$

This Jacobian gives us a medium to infer causality in the model through the mathematical operation of differentiation. A variable X_i is directly dependent – or caused by – another variable X_j if and only if

$$\frac{\partial X_i}{\partial X_j} \neq 0. \tag{3}$$

Since X_i is fully determined by the equation $X_i = f_i(\mathbf{X})$, then Eq. (3) is equivalent to the condition that $\partial f_i / \partial X_j \neq 0$. Thus X_j causes an effect in X_i if and only if J_{ij}, the $(i, j)^{\text{th}}$ entry of $\mathbf{J}$, is non-zero. These causal dependancies are encoded as a binary matrix $\mathbf{A}$ that depends on the SFC model through $\mathbf{J}$ by

$$A_{ij} = \begin{cases} 1 & \text{if } J_{ji} \neq 0, \\ 0 & \text{if } J_{ji} = 0. \end{cases} \tag{4}$$

This causality will be the foundation of the DAG, but before continuing with the derivation we introduce and formally define the necessary graph theoretical concepts while directing the reader to Newman (2010) for an accessible overview of basic graph theory.

Definition A *directed graph* $G = (V, E)$ is a set of nodes $V = \{v_1, v_2, ..., v_n\}$ along with a set of directed edges $E = \{e_1, e_2, ..., e_n\}$ that link the nodes. Each edge e_i is of the form $e_i = (v_{i_1}, v_{i_2})$ indicating that there is a link from node v_{i_1} to node v_{i_2}.

A useful representation of a directed network is the *adjacency matrix*. This is a binary matrix $\mathbf{A}$ whose elements A_{ij} satisfy

$$A_{ij} = \begin{cases} 1 & \text{if the is a link between } v_i \text{ and } v_j, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The adjacency matrix is a great mathematical tool used to calculate a range of properties about the network. One such property, which is important in our setting, is the existence of cycles.

Definition In a directed network, a *cycle* is a closed loop of edges where the direction of each edge points the same way around the loop. A directed network that has no cycles is called a *directed acyclic graph*, or *DAG*.

To find out if a network is acyclic it is sufficient to examine the eigenvalues of the adjacency matrix. If all of the eigenvalues of the adjacency matrix are equal to zero then the network is acyclic. Otherwise cycles exist, each of which can be mapped to a strongly connected component.

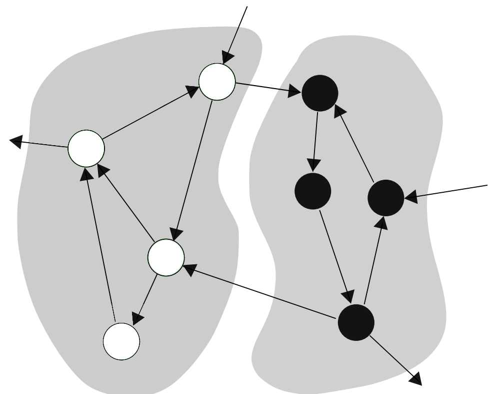


Figure 1: Schematic of the proof of Lemma 1. In the condensation graph, the strongly connected components S_1 (white) and S_2 (black) are replaced by metanodes. If these two metanodes form a cycle, then the nodes in S_1 and S_2 are linked be directed paths in both directions, which contradicts the maximal property of stronlgy connected components.

Definition A *strongly connected component (SCC)* in a directed network $G = (V, E)$ is a maximal subset $V_S \in V$ of nodes such that every pair of nodes in V_S are connected by directed paths in both directions. Maximal here means that no additional nodes in V can be included in V_S without breaking its property of being strongly connected. Every node in the network belongs to one and only one strongly connected component, and strongly connected components of only one node may exist. The set $S = \{S_1, S_2, \ldots, S_m\}$ of strongly connected components of G forms a partition of G and this partition is unique.

In a strongly connected component of more than one node, every node is part of a cycle. This is intuitive, as to be part of the SCC there must be a directed path between the node itself and every other node in the SCC in both directions. An implication of this is that the set of nodes forming any cycle is a subset of exactly one strongly connected component.

Lemma 1 *Let $G = (V, E)$ be a directed graph and $S = \{S_1, S_2, \ldots, S_m\}$ be the set of m strongly connected components of G. The* condensation graph *G_C of G is the graph whose strongly connected componenents are contracted into single nodes called* metanodes.[4] *Then the condensation graph is a DAG.*

[4]Efficient algorithms which identify the strongly connected components exist such as Trajan's algorithm Tarjan (1972) or Kosaraju's algorithm Hopcroft (1983).

Proof We aim to show that there are no cycles in G_C in which case it is a DAG. This will be proved by contradiction. A schematic of the proof is given in Fig. 1.

Assume that there exists a cycle in G_C. Then by definition, there is a directed path between at least two vertices in G_C in both directions. Because the nodes in G_C are the strongly connected components in G, this implies that there is directed path in both directions between at least two strongly connected components of G. However this cannot be the case. The strongly connected components of G are maximal implying that for every node outside the SCC there is not a directed path in both directions to the SCC. Thus two strongly connected components cannot be linked by paths in both directions, and so we arrive at a contradiction.

The construction of a DAG from the SFC model now follows from the graph theoretical concepts introduced here. Recall that the behaviour of the variables in the SFC is governed by the system of equations $\mathbf{X} = \mathbf{f}(\mathbf{X})$ with Jacobian $\mathbf{J} = \partial \mathbf{f}/\partial \mathbf{X}$. The dependencies between variables in the system was encoded in the matrix $\mathbf{A}$ which is related to $\mathbf{J}$ through Eq. (4). We now construct a directed graph G from the SFC as the graph whose adjacency matrix is $\mathbf{A}$. The condensation graph of this directed graph is taken and by Lemma 1 this is a DAG. Since there is only one unique partitioning of a graph into its strongly connected components then the DAG is unique.

The implementation of the DAG construction is performed in R and is available in a package from GitHub. This package takes the system of equations in the SFC model and returns various outputs. These include the initial directed graph, the strongly connected components which contain all of the cycles and the unique DAG where the cycles have been replaced by metanodes. In the next section we illustrate an example of the conversion of a SFC into its DAG. We explain why this is important and show the insights into the model that this can afford.

3 Applications

3.1 The BMW Model

In this section, we give an illustrative example of how to construct a graphical representation of a stock flow consistent model. As our example we use the *bank-money world* (*BMW*) model introduced in (Godley and Lavoie, 2007, chapter 7). This is a model of a closed economy where there are three sectors – households, production firms and banks. The sectors are linked through both the balance sheet and the transactions-flow matrix. The full details of the model can be found in of Godley and Lavoie (2007); meanwhile the variables and equations that govern their evolution are summarized in Table 1.

For each of the equations given in Table 1, a change in the value of a variable on the right hand side of the equation will cause a corresponding change in the value of the variable in the left hand side of the equation. The Jacobian matrix, which calculates the partial derivative of each variable with respect to each other variable, quantifies this. If we assign to each variable a row and column i then the Jacobian matrix satisfies

$$J_{ij} = \frac{\partial X_i}{\partial X_j}. \tag{6}$$

where X_i and X_j are the variables corresponding to row i and column j respectively. For example in the case of the supplied wage bill WB_s, which satisfies the equation $WB_s = W \times N_s$, we have

$$\frac{\partial WB_s}{\partial X_j} = \begin{cases} N_s & \text{if } X_j = W, \\ W & \text{if } X_j = N_s, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Thus the only two non-zero elements of rows corresponding to WB_s are in the rows corresponding to W and N_s. With this in mind, the full Jacobian matrix for the BMW model is given by

$$\mathbf{J} = \begin{array}{c} \\ AF \\ C_d \\ C_s \\ \vdots \\ YD \end{array} \begin{array}{c} \begin{array}{ccccc} AF & C_d & C_s & \dots & YD \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & \alpha_1 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \end{array}. \tag{8}$$

The Jacobian is a square, asymmetric matrix. It is sparse because the equation for each variable generally depend on a small number of the other variables. The Jacobian is used to obtain the adjacency matrix of the directed graph though Eq. (4), i.e., by taking the transpose of the Jacobian and replacing the non-zero terms with ones:

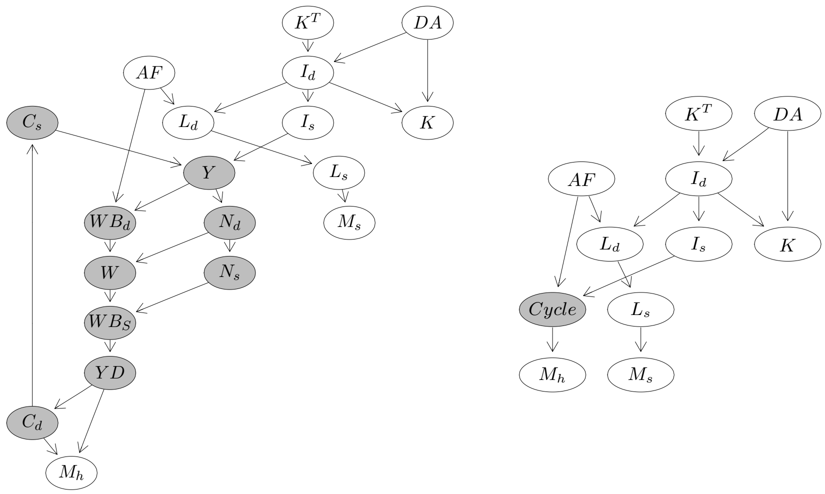
$$\mathbf{A} = \begin{array}{c} \\ AF \\ C_d \\ C_s \\ \vdots \\ YD \end{array} \begin{array}{c} \begin{array}{ccccc} AF & C_d & C_s & \dots & YD \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{pmatrix} \end{array}. \tag{9}$$

Thus from the system of equations of the stock-flow consistent model – as given in Table 1 – we have constructed a directed graph which is represented by the adjacency matrix $\mathbf{A}$. This directed graph is shown in Fig. 2a. The graph is plotted in a hierarchical manner. It shows in a binary manner how each system variable causes a changes in the value of other system variables.

To illustrate this, consider the amortization funds AF. At the start of a time period AF is equal to a fraction δ of the stock of capital from the previous time period, K_{-1}. Through the directed graph it can be observed that these amortization funds affect its two children nodes, the demanded loans L_d and the demanded wage bill WB_d. The two branches which emerge from AF differ in an important way. One branch is a chain: L_d affects L_s which subsequently affects M_s before the chain stops. Along this chain the value of any two variables are independent once the value of the intermediary nodes are known. For example the bank deposits M_s are independent of AF once the value of L_s is known.

On the other hand, the second branch emerging from AF through WB_d becomes part of a loop in which we cannot infer such independencies between nodes. WB_d affects its child node W which affects WB_s and this effect passes through the loop before eventually returning and causing an effect in WB_d itself. Thus, the value of intermediary nodes between two nodes does not rule the two nodes independent as the effect of the second node will pass through the loop to eventually cause a change in the first node. Even though the causal effect of WB_d on WB_s is fully explained by WB_s's parent node W, WB_d and W_s are not independent even if we know the value of W because W_s will subsequently lead to a feedback effect in W_b.

The nodes in the closed loop collectively form a strongly connected component. This is illustrated in the directed graph of Fig. 2a by shading each of these nodes. The strongly connected component can be replaced by a metanode to form a DAG – this is shown in Fig. 2b. From the DAG we can



(a) *BMW* directed graph (b) *BMW* DAG

Figure 2: Directed graph (a) and DAG (b) representations of the *BMW* model. Nodes corresponds to the variables of the model while directed links represent the causal dependancies between variables. White nodes do not belong to any loops while grey nodes are part of a strongly connected component.

infer direct causality because the graph is acyclic, i.e., there are no loops. The funds flowing through a node depend directly on its parent nodes, nodes that have links directed to that node. Once we know the value of the parent nodes, the value of a node is independent of all nodes further up the hierarchy from the parent nodes. Similarly, the funds flowing through a node cause changes in the nodes to which it points, its children nodes. However, given the value of its children nodes, the value of nodes further down the hierarchy from its children nodes are independent of the node. The directed graph and DAG give not only a visual representation of the flow of funds through the system but also the causal mechanisms that exist within the system.

4 Conclusion

The objective of this paper was to rigorously show that for every stock flow consistent (SFC) macroeconomic model there is a corresponding directed acyclic graph (DAG) which is unique.

We construct the DAG of any stock flow consistent model in the following manner. Firstly, we use Jacobian methods on the system of equations of the stock flow consistent model to form a directed graph. This directed graph is then decomposed into its corresponding condensation graph by replacing the strongly connected components with single metanodes (Fig. 1). Since every cycle in the directed graph is contained in exactly one strongly connected component, the condensation graph is acyclic and so we have formed a DAG.

We illustrate the theory with an example, the *BMW* model developed by Godley and Lavoie (2007), and provide details of a computational tool that gives the directed acyclic graphical representation of any stock flow consistent model. We have used this package to generate the DAGs of almost every model within Godley and Lavoie (2007) and these are available online.[5]

This formal linkage of two rather disparate fields is important. The DAG makes it much easier to visualise large macroeconomic models. This is useful not only for analysis but also for efficiently solving the model computationally. The topological ordering of the system (Fig. 2) – especially the isolation of cycles – can lead to great improvements in computational speed, which decreases as the model increases in size.

This linkage is important as directed acyclic graphs allow us to use well understood techniques for system-wide causal discovery (Pearl, 2000). Graph-theoretic search methods have not, typically, been used for time series data, the data type we normally use in macroeconomic models.

Our further work will concentrate on recovering directed acyclic graphs from empirical stock flow consistent models to aid in model selection and to generate a topological ordering within the model structures.

References

Bessler, D. A., J. Yang, and M. Wongcharupan (2003). Price dynamics in the international wheat market: modeling with error correction and directed acyclic graphs. *Journal of Regional Science 43*(1), 1–33.

Caverzasi, E. and A. Godin (2015). Post-keynesian stock-flow-consistent modelling: a survey. *Cambridge Journal of Economics 39*(1), 157–187.

Dos Santos, C. H. and G. Zezza (2008). A simplified,'benchmark', stock-flow consistent post-keynesian growth model. *Metroeconomica 59*(3), 441–478.

[5]The GitHub repository is here.

Godley, W. (1999, April). Money and credit in a keynesian model of income determination. *Cambridge Journal of Economics 23*(2), 393–411.

Godley, W. and M. Lavoie (2007). *Monetary Economics An Integrated Approach to Credit, Money, Income, Production and Wealth.* Palgrave-Macmillan.

Heer, B. and A. Maussner (2009). *Dynamic General Equilibrium Modelling.* Springer.

Hoover, K. D. (2001). *Causality in Macroeconomics.* Cambridge: Cambridge University Press.

Hopcroft, J. E. (1983). *Data structures and algorithms.* Pearson education.

Lauritzen, S. L. (2001). Causal inference from graphical models. *Complex stochastic systems*, 63–107.

Morgan, S. L. and C. Winship (2007). *Counterfactuals and causal inference: Methods and principles for social research.* Cambridge University Press.

Newman, M. (2010). *Networks: an introduction.* Oxford University Press.

Pearl, J. (2000). *Causality: models, reasoning and inference*, Volume 29. Cambridge Univ Press.

Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM journal on computing 1*(2), 146–160.

Tobin, J. (1982). Money and finance in the macroeconomic process. *Journal of Money, Credit and Banking 14*(2), 171–204.

Endogenous Variables		Equations
AF	Amortization funds	$AF = \delta K_{-1}$
C_d	Consumption goods demand	$C_d = \alpha_0 + \alpha_1 YD + \alpha_2 M_{h-1}$
C_s	Consumption goods supply	$C_s = C_d$
DA	Depreciation Allowance	$DA = \delta K_{-1}$
K	Stock of capital	$K = K_{-1} + I_d - DA$
K^T	Target stock of capital	$K^T = \kappa Y_{-1}$
L_d	Demand for bank loans	$L_d = L_{d-1} + I_d - AF$
L_s	Supply of bank loans	$L_s = L_d$
I_d	Demand for investment goods	$I_d = \gamma(K^T - K_{-1}) + DA$
I_s	Supply of investment goods	$I_s = I_d$
M_h	Bank deposits held by households	$M_h = M_{h-1} + YD - C_d$
M_s	Supply of bank deposits	$M_s = M_{s-1} + L_s - L_{s-1}$
N_d	Demand for labour	$N_d = Y/pr$
N_s	Supply of labour	$N_s = N_d$
W	Wage rate	$W = WB_d/N_d$
WB_d	Wage bill – demand	$WB_d = Y - \bar{r}L_{d-1} - AF$
WB_s	Wage bill – supply	$WB_s = N_s W$
Y	National Income (GDP)	$Y = C_s + I_s$
YD	Disposable income	$YD = WB_s + \bar{r}M_{h-1}$

Exogenous Variables and parameters	
α_0	Exogenous component in consumption
α_1	Propensity to consume out of income
α_2	Propensity to consume out of wealth
δ	Depreciation rate
γ	Speed of adjustment of capital to its target value
κ	Capital-output ratio
pr	Labour productivity
$\bar{r}$	Exogenously set rate of interest on bank loans

Table 1: Variables, equations and parameters of the *BMW* model.