1

Introduction: Causal Representation

1.1 The causal framework

The notion of cause and effect is deeply engrained in our understanding of the world. It also has profound and sometimes controversial resonances in philosophy, which will be side-stepped in this book. From a philosophical standpoint the attitude here is that causal reasoning and causal representation help us to understand and analyse many topics in macroeconomics. In other words, causality is conceived here as a pedagogical tool for explication of the matter in hand, both to ourselves and to others.

Economists are generally ambivalent about causal arguments. On the one hand such arguments are used extensively in the classroom where we quite often find a form of homespun causal representation produced on the blackboard. But on the other hand causal arguments are often eschewed in the more formal writings of economists. The reluctance to use explicitly causative arguments is probably based on two different reasons, one stemming from the analytical side and the other from the empirical side of the subject.

On the analytical side, economists are most confident about statements to do with equilibrium configurations of the economy but are generally less confident about transitions between equilibria, or the mechanisms by which the equilibria are established. On the empirical side, analysis often falls into the realm of econometrics, which has historically been heir to the skepticism about causality to be found in its other parent discipline, statistics.

The standard methodology of comparative statics in economic analysis illustrates the power of reasoning based on notions of equilibrium. The archetypal example is of supply and demand in some market, in which equilibrium is synonymous with equality between supply and demand. Allow some external condition to change and we get a new equilibrium which can be compared with the original one. For example "bad weather spoiled the harvest and reduced supply so that there is a higher price and lower quantity transacted in the new equilibrium". This is fine, but wouldn't a causal explanation require us to know exactly how the new equilibrium was established? On this topic economists often appear ill at ease, perhaps justifiably.

The mechanism by which the price was driven up may be obscure without knowledge of the institutional and other conditions governing behaviour in this market. The precise details of disequilibrium behaviour depend on circumstances, and as they are often unknown it may be thought better to reason without recourse to such arguments: whatever the circumstances, the equilibrium can be described, so it is safer to keep to that. Nevertheless, it is clear where the causal chain began—with the bad weather—and where it ended, with the rise in price. So although we might not be able to fill in the details, at least the status of these two variables, as cause and effect, is not in doubt despite the many conceivable routes between them. To arrive at this overarching causality, it is sufficient to assert that price rises when demand exceeds supply.

The empirical source of the unease with causal reasoning stems from statistics and is crudely summarised by the true statement that correlation does not imply causation. This, combined with the fact that economists are seldom in a position to conduct controlled experiments, inclines them to follow the statisticians in the tradition begun by Karl Pearson and assume a skeptical stance with regard to causation[1]. The correlation/causation point is often sharpened for economists by the "identification problem" which heavily qualifies the conditions under

---

[1] see J. Pearl, *Causality* (Cambridge University Press, 2000) and K. D. Hoover, *Causality in Macroeconomics* (Cambridge University Press, 2001) for an extensive discussions.

which an empirical association between jointly determined endogenous variables can be ascribed to one mechanism of that joint determination or another. Nevertheless, econometricians as much as anybody are keen to explore causal relationships, and have developed sophisticated methods, both for testing the direction of causality in time series data and for coherently estimating components of simultaneous structural models.

In the classroom, where the need for causal argumentation is irresistible, we often find an appealing but incomplete causal representation in a kind of visual shorthand. To illustrate, consider what might be seen on the blackboard after a class on monetary policy:

$$Ms\uparrow \rightarrow i\downarrow \rightarrow r\downarrow \rightarrow I\uparrow \rightarrow Y\uparrow \rightarrow Md\uparrow \rightarrow i\uparrow$$

This is a wonderfully compact summary of the argument that an increase in the money supply induces a fall in the rate of interest which induces a rise in investment which raises total expenditure and income, which in turn raises the demand for money which pushes the interest rate up. A horizontal arrow could stand for "implies", but more naturally means "causes", while the vertical arrows show the direction of change of the associated variable. But the point here is that the (horizontal) arrow schema is a primitive representation of "causality". Primitive, both in the sense of being crude and unrefined, and also in the sense of being a foundation building block. It does the job of identifying cause and effect, and indeed of displaying a causal chain. However, its shortcomings in this regard are also easily apparent: the illustrated causal chain reveals a contradiction—or at least an incompleteness—in that it shows the rate of interest initially falling but then rising. Then what is the eventual effect of increased money supply on the rate of interest? And furthermore, since the chain could obviously be extended by repeating its last four links, *ad infinitum,* how can any conclusion be reached?

Now we come to the point of this book. Its aim is to present some models in macroeconomics, making use of the vivid "homespun" causal representation of the classroom blackboard, but doing so in a systematic and coherent manner. The vehicle we use to achieve this aim is known in some branches of engineering as a "signal-flow graph", but here it is simply referred to as a "flowgraph".

1.2 Flowgraphs

The representation of an engineering system, with its various transforming devices and mechanisms, as a graph is quite normal. The graph abstracts and displays the essential structure of the system being modelled. Such a graph may often appear similar to an electric circuit diagram, with its representation of the sundry connected physical elements. When the graph represents a more general system it is known as a "block diagram", in which each block represents some mechanism which transforms input variables into output variables. Textbooks on control engineering and systems engineering are replete with such block diagrams. Variables are conceived as flowing along the connecting lines between the transforming blocks just as electric current might be thought to flow along the wires of an electric circuit, transformed en route by resistors, inductors and so on. The block diagram representation also appears in economics textbooks, showing the circular flow of income. But there is another graphical representation in which the blocks, or "nodes" now, are the variables and the connecting lines between them represent the transformations. This "signal-flow graph", or more simply "flowgraph", easily and naturally represents linear equations, in which case the transformations on the connecting lines are coefficients or simple multipliers.

Flowgraphs and block diagrams are alternative, dual, representations of the same real system. A visual advantage of the block diagram is its physical resemblance to the system being modelled, whereas a visual advantage of the flowgraph is its transparent causal interpretation. In what follows there will be many applications of flowgraphs to macroeconomic models, but first we must find out how they are constructed, manipulated and analysed.

1.3 Building flowgraphs: basic operations

Suppose that variable X causes variable Y, and that we write down the relationship thus:

$$Y = f(X) \tag{1}$$

This representation reads "Y is a function of X", seeming to imply that Y is determined by X, or caused by X. But of course the algebraic representation does not itself imply a direction of causation. The relationship could just as easily be written

$$X = f^{-1}(Y) \tag{2}$$

where the function $f^{-1}(\ )$ is the inverse of $f(\ )$. The mathematical representation is agnostic about the direction of causality. But if we know that X causes Y, then equation (1) seems more meaningful than equation (2).

Let us be more specific and suppose that Y is proportional to X, so that:

$$Y = aX \tag{1a}$$

where a is a constant. Of course, this is algebraically equivalent to:

$$X = Y/a \tag{2a}$$

and the equation representation remains agnostic as to causal direction. But consider now writing the linear proportional relationship (1a) in manner which does show the direction of causality.

$$X \xrightarrow{a} Y \tag{1b}$$

This reads "X causes Y with transmittance a". Instead of an equation, (1b) depicts the relationship between Y and X as an "elementary graph". The arrowed line is a "directed edge", and a directed graph is made up of a set of such directed edges connecting various nodes. Note that we also represent the transmittance—or "gain" in engineering parlance—between X and Y as a. This type of graph is known to systems engineers as a "signal flow graph", or simply a "flowgraph".

Observe that the graph incorporates more information than an equation because it displays the direction of causation, which is crucial for a cognitive interpretation. But, it might be thought, equations have the advantage that they can be manipulated and solved. In fact, graphs too can be manipulated and solved. Moreover, graphical manipulations themselves convey more insight into the system they represent than do the corresponding algebraic manipulations, and solving for the effect of one variable on another is much simpler for a flowgraph than for the corresponding equations. This is particularly true for qualitative solutions.

Because graphs contain more information than equations, the relationship between them is not one-to-one. Many graphs may be implied by one equation, but there will only be one graph that is consistent with a particular causal interpretation of that equation. We now examine graphs involving more than one directed edge or arc, and how they can be simplified. The five parts of Figure 1.1 illustrate different features of graph construction and simplification, and the corresponding equations.

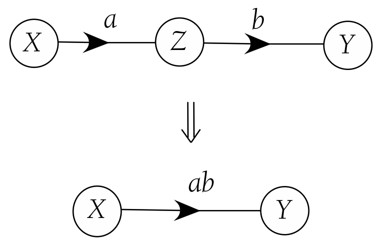
In part (i) we see the basic property of path transmittances: that the overall transmittance is the product of the individual arcs comprising the path. Part (ii) shows that the total effect on a node is the sum of the effects from all its inflowing arcs. These two properties are combined in part (iii) to show that the overall transmittance from X to Y is the sum of the two distinct paths between these nodes. It is evident from this flowgraph representation of "arcs in parallel" that there is just one exogenous variable in this system, namely X, whereas that fact is somewhat opaque in the equations even if the left hand side variable of each equation is designated the "dependent variable".

When a system includes a feedback process of some description, its flowgraph will contain a loop, as shown in part (iv) of Figure 1.1. Here Y depends on X and Z, but Z in turn depends on Y. Algebraically this is treated as a system of "simultaneous equations", without further ado, though the very notion of simultaneity begs questions about causality and dynamics in the system it represents. A loop occurs when some path emanating from a node (*i.e.* a variable) eventually returns to that same node. Here there is a "loop transmittance" of bc. It is quite possible for a node to be on several distinct loops, which may or may not have an arc or arcs in common. Distinct loops which include the same nodes are said to be "touching", as in part (v). As will be seen, it is important to know whether loops are touching or non-touching when it comes to solving them for some overall transmittance between two nodes. Here there are two loops which have nodes Y and Z in common, and are therefore touching. It can be seen that if the lower loop were simplified by absorbing the W node there would then be two simple feedback paths in parallel from Z to Y which could be consolidated to produce a flowgraph with just one feedback loop having a loop transmittance equal to $b(c+ed)$.

i. Arcs in cascade

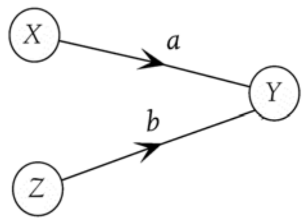
$$\left.\begin{array}{l} Z = aX \\ Y = bZ \end{array}\right\}$$

$$\Rightarrow \quad Y = abX$$



ii. Summation

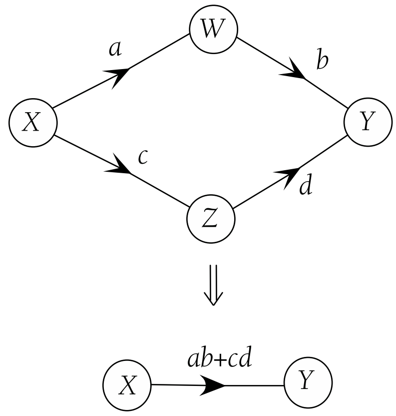
$$Y = aX + bZ$$



iii. Arcs in parallel

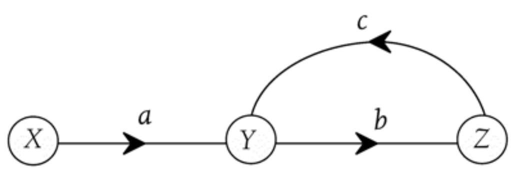
$$W = aX$$
$$Z = cX$$
$$Y = bW + dZ$$

$$\Rightarrow Y = (ab + cd)X$$



iv. Loop

$$Y = aX + cZ$$
$$Z = bY$$



v. Touching loops

$$Y = aX + cZ + dW$$
$$Z = bY$$
$$W = eZ$$

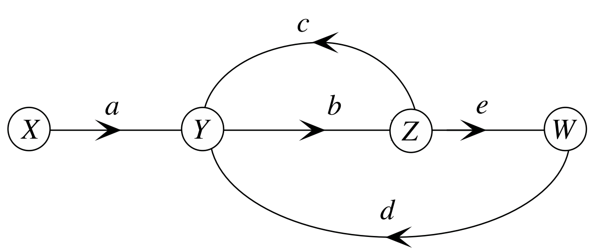


Figure 1.1 Equations and graphs

1.4 Condensing a graph

The process of condensing a graph is equivalent to the algebraic process of substituting from one equation into other equations. The basic operations involve the combination of arcs in parallel (see above), the elimination of intermediate nodes with arcs in cascade, the reduction of loops involving several nodes to “self loops” on one of those nodes, and the absorption of self loops.

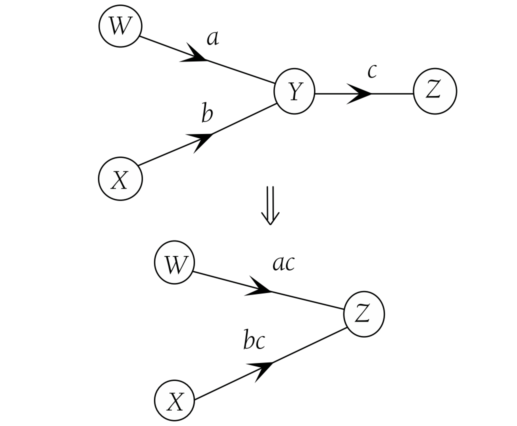
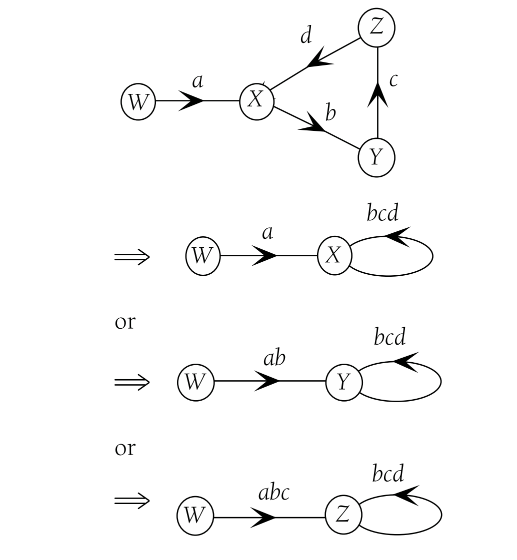
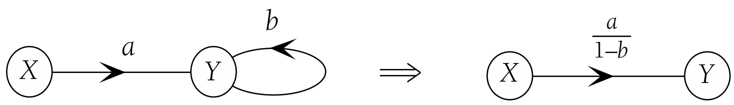
i. Absorbing a node $Z = cY$ $Y = aW + bX$ $\Rightarrow Z = acW + bcX$	
ii. Consolidating a loop $X = aW + dZ$ $Z = cY$ $Y = bX$ $\Rightarrow \quad X = aW + bcdX$ or $\Rightarrow \; Y = abW + bcdY$ or $\Rightarrow \; Z = abcW + bcdZ$	

Figure 1.2 Effects of node absorptions

Figure 1.2(i) illustrates the effect of node absorptions. The procedure is simply to remove the node from the flowgraph, but in doing so to retain all the path connections between the remaining variables along with their path transmittances.

All loops can be reduced to self-loops on one of the variables within the loop by a sequence of node eliminations, as illustrated in Figure 1.2(ii). Any variable within the loop can act as the self-loop node. Note that the loop transmittance is not affected by the choice of node variable upon which to focus the loop. Thus in the example shown, the three-node loop can be reduced by three different sequences of node absorptions. The loop transmittance is always the same, and the path transmittances between the remaining variables are unaltered.

By these means, all loops can be reduced to self loops. Moreover, all self loops can be absorbed into the direct paths that lead into a node, as shown in Figure 1.3.



$$Y = aX + bY \quad \Rightarrow \quad Y = \frac{a}{1-b}X$$

Figure 1.3 Absorption of self loop into inflowing arc

The limiting reduction of a graph produces the *reduced form* of the model, in which all the paths are direct connections from source nodes (exogenous variables) to sink nodes (endogenous variables). The reduced form has an important role in econometrics, where it contrasts with the structural form of a model which expresses our understanding about how the model works in its detailed components. The reduced form hides those details but reveals the overall cause and effect roles of the variables—the magnitudes of those effects are the transmittances associated with the arcs.

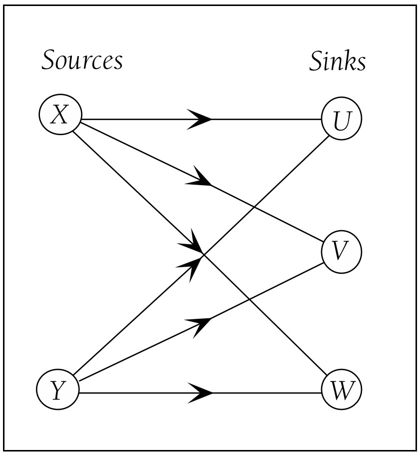


Figure 1.4 Reduced form

1.5 Evaluation of transmittances

i. Direct paths

We often wish to calculate the effect of a change in an exogenous variable on one or several endogenous variables. In a flowgraph, an exogenous variable is recognised by the fact that it only has outflow arcs emanating from it. Any variable with an inflow arc is endogenous, even if it also has outflows. Where there are no loops in the flowgraph it is straightforward to calculate the overall transmittance from an exogenous to an endogenous variable. All that needs to be done is to identify all the paths connecting the two nodes, calculate the transmittance along each of them and add them up. Consider the flowgraph in Figure 1.5.

There are two exogenous variables, U and V, and all the other variables are endogenous. Note that there is no path connecting V to X, so V cannot affect X and thus the transmittance between them is zero. Consider the transmittance between V and Z. It is simply egd, since only arcs in cascade are involved. Now consider the transmittance between U and Z. There are three direct paths, and the overall transmittance is, accordingly, $a + bcd + bfgd$.

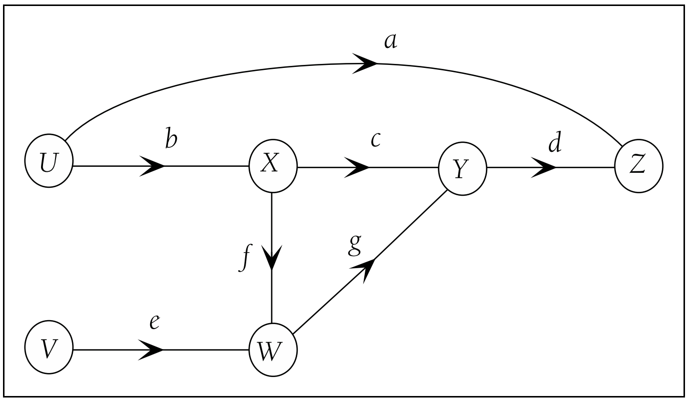


Figure 1.5 Direct paths

It is perhaps instructive to consider the same model written out as equations:

$$\begin{aligned} X &= bU \\ W &= eV + fX \\ Y &= cX + gW \\ Z &= aU + dY \end{aligned}$$

These equations have been written down in a particular order to show the structure of the model. It is evident that if we are given the values of the exogenous variables, U and V, the system can be solved from top to bottom by substitution, *i.e.* the system is "recursive" in the sense given to that term by the Swedish economist Herman Wold. The same point can also be made by expressing the model in matrix form, with the endogenous variables on the left hand side and the exogenous variables on the right hand side:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -f & 1 & 0 & 0 \\ -c & -g & 1 & 0 \\ 0 & 0 & -d & 1 \end{pmatrix} \begin{pmatrix} X \\ W \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} b & 0 \\ 0 & e \\ 0 & 0 \\ a & 0 \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix}$$

where the thing to note is the lower triangular structure of the matrix of coefficients of the endogenous variables. When a flowgraph has no loops, the matrix of coefficients on the endogenous variables can always be written in lower triangular form.

ii. Paths with loops

The advantages of flowgraphs are also evident in models that contain feed-back mechanisms—that is, flowgraph models with loops. If we wish to evaluate the transmittance between an exogenous and an endogenous variable in such a flowgraph, there are two ways to go about it. Firstly we could use the methods of graph reduction outlined above, until the graph shows a single arc connecting the variables we are interested in. Alternatively we can use a formula known as Mason's rule.

To use Mason's rule we need to extract the following information from the flowgraph: (i) all the direct transmittances from the exogenous variable to the endogenous variable, *i.e.* all the T_i, where i indexes the set of paths connecting the variables, and (ii) all the loop transmittances in the system, L_j, indexed by j. Let the overall transmittance we seek be labelled T, then Mason's rule is:

$$T = \frac{1}{\Delta}\sum_i T_i \Delta_i$$

where Δ, which is the "system determinant", and the "cofactor" terms Δ_i are evaluated in terms of the various loop transmittances. They are given by:

$$\Delta = 1 - \sum_j L_j + \left[\sum_{k,l} L_k L_l - \sum_{k,l,m} L_k L_l L_m + ...\right]_{nontouching\ loops}$$

and Δ_i is the value of Δ excluding all terms which involve loops which touch the ith forward path (*i.e.* set the value of such touching loop transmittances to zero).

The term $\sum L_k L_l$ is the sum of the products of all pairs of non-touching loop transmittances, and similarly the term $\sum L_k L_l L_m$ is the sum of the products of all loop transmittances for all possible combinations of three non-touching loops. The term in square brackets goes on to include similar terms for combinations of four, five *etc*. non-touching loops until the possible combinations are exhausted. This usually occurs well before we reach combinations of four non-touching loops, thankfully! Consider the example shown in Figure 1.6.

In this flowgraph there are two forward paths from U to Z, with transmittances T_1=*abcde* and T_2=*afe* respectively. There are three loops, with transmittances L_1=*cg*, L_2=*feh*, and L_3=*bcdeh*, and the loops L_1 and L_2 are non-touching, so the system determinant is:

$$\begin{aligned}\Delta &= 1-(L_1+L_2+L_3)+L_1L_2 \\ &= 1-cg-feh-bcdeh+cgfeh\end{aligned}$$

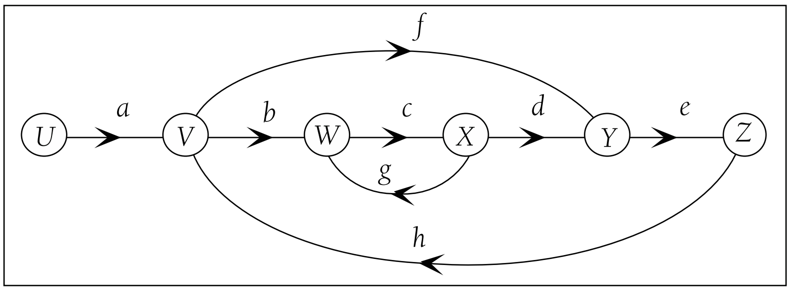


Figure 1.6 Flowgraph with loops

Because all three loops touch path T_1, setting $L_1=L_2=L_3=0$ in the expression for Δ gives $\Delta_1=1$. Because loops L_2 and L_3 touch path T_2, setting $L_2=L_3=0$ in the expression for Δ gives Δ_2=1− L_1 =1− cg. These results can now be substituted into Mason's formula to get the overall transmittance from U to Z as:

$$\left\langle \frac{Z}{U} \right\rangle = \frac{T_1\Delta_1 + T_2\Delta_2}{\Delta} = \frac{abcde + afe(1-cg)}{1-cg-feh-bcdeh+cgfeh}$$

where the left hand side has been expressed as a ratio in angle brackets to signify "transmittance". The context determines how to interpret transmittance.

If the model is static and linear then the transmittance represents the ratio of the change in Z to the change in U, *ceteris paribus*—*i.e.* holding the other exogenous variables constant. If the model is static and non-linear, so that the arc transmittances *a,b,c,..etc.* represent partial derivatives, the flowgraph represents a local linear approximation and the overall transmittance can be interpreted as the total derivative dZ/dU, which is strictly only valid for small changes. Finally, if the model is linear and dynamic, so that the system contains dynamic operators, such as lags or time derivatives (represented by the Laplace transform operator), the overall transmittance represents the "transfer function" from U to Z.

Now let us see how our model can be solved by expressing it in matrix notation and using Cramer's rule. From the flowgraph we can write down the following equations:

$$
\begin{aligned}
V &= aU + hZ \\
W &= bV + gX \\
X &= cW \\
Y &= dX + fV \\
Z &= eY
\end{aligned}
$$

which can be expressed in matrix terms, with endogenous variables on the left as:

$$
\begin{pmatrix} 1 & 0 & 0 & 0 & -h \\ -b & 1 & -g & 0 & 0 \\ 0 & -c & 1 & 0 & 0 \\ -f & 0 & -d & 1 & 0 \\ 0 & 0 & 0 & -e & 1 \end{pmatrix} \begin{pmatrix} V \\ W \\ X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} U \end{pmatrix}
$$

so Cramer's rule gives the solution for Z as:

$$
Z = \frac{\begin{vmatrix} 1 & 0 & 0 & 0 & aU \\ -b & 1 & -g & 0 & 0 \\ 0 & -c & 1 & 0 & 0 \\ -f & 0 & -d & 1 & 0 \\ 0 & 0 & 0 & -e & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 0 & 0 & -h \\ -b & 1 & -g & 0 & 0 \\ 0 & -c & 1 & 0 & 0 \\ -f & 0 & -d & 1 & 0 \\ 0 & 0 & 0 & -e & 1 \end{vmatrix}} = \frac{abcde + afe(1-cg)}{1 - cg - feh - bcdeh + cgfeh} U
$$

which is the same result as before, but derived by way of tedious and error-prone algebra.

1.6 Modelling with flowgraphs: further topics

Here we discuss certain topics which arise in the context of modelling with flowgraphs, namely (i) qualitative flowgraphs, (ii) dynamics, (iii) parameter variation and (iv) reversal of causality.

i Qualitative flowgraphs

The "qualitative calculus" is a characteristic of comparative static analysis. It is easily and naturally accommodated within a causal flowgraph representation. Although the flowgraphs set out above represent linear systems with algebraic parameters, those parameters could be taken to be the partial derivatives of non-linear functions for which the linear formulation serves as a local approximation. It is usually the case in economics that our theoretical understanding of the components within the system is "non-parametric", often comprising monotonicity assumptions rather than more precise assumptions about functional form. Then the true information in the flowgraph is of the *signs* of the parameters. For this reason it is useful to maintain a convention that all the algebraic parameters shown on a flowgraph are positive, so that a negative effect is signified by a minus sign in an arc transmittance. This convention enables an immediate inference about the direction of change in an endogenous variable induced by a change in an exogenous variable: the task is merely to identify the signs of all the paths connecting the two variables. This can be seen on inspection in graphs that are not too complicated. Of course in a situation in which two such paths have path transmittances with different signs, the overall impact on the endogenous variable is ambiguous.

Systems which contain loops may appear to pose a problem for this simple procedure for the sign calculus. That could indeed be so in some unusual cases, but usually the loops can be ignored for the purpose of signing effects. An important constraint on the ability of loops to upset the simple calculation is that the systems we consider are assumed to be stable. This implies that the system determinant is positive, so the denominator in Mason's rule formula cannot change the sign given in the numerator. By itself this is not sufficient to justify the simple sign calculus based on the path transmittances because a negative path cofactor could reverse the sign of the term in the numerator of Mason's rule. So as a precaution we should identify any loops that do not touch the path concerned (and any such non-touching pairs etc), and mentally check whether the cofactor could conceivably be negative—a warning sign would be non-touching loops with loop transmittances exceeding one in value. If this seems possible, then it would be wise to do the algebra on the numerator in Mason's rule. But in the relatively simple models we shall encounter that is seldom necessary.

ii. Dynamics

There is a close connection between a causal understanding of a system and its dynamic behaviour. Whenever we consider a static model there are dynamics lurking not far below the surface. But they are often uncharted territory for economics and as such best avoided. However from a causal perspective it is often useful to recognise them. For example, the simple market supply and demand model only concerns equilibrium states, but an explanation about how the system gets from one such state to another can hardly avoid bringing in dynamics (e.g. "excess demand stimulates a rise in the price"). A causal understanding of the system hinges crucially on implicit or explicit dynamics. Static models usually represent an equilibrium state of an underlying dynamic system, and the implicit dynamics can come in various guises. Perhaps the most common is the assignment of exogeneity to particular variables. This tells us what changes first, even though algebraically we consider the system to be simultaneous.

The very notion of "simultaneity" means that the time-scale at which our flow variables are represented is much larger than that at which the dynamic process of influence from one variable to another operates. If we could write down the model with sufficiently fine granularity in time, we would find that all the arcs in the model embody lags or accumulations. For most models such detail would be an inessential distraction at best, but at worst would result in complete obfuscation. We would not be able to see the wood for the trees. Sometimes, however, the essence of a model's behaviour is in its dynamics, which therefore must be represented explicitly. This actually helps in the construction of the flowgraph because it is natural to assume that the past causes the future rather than the reverse.

The analytical tools needed for a dynamic model depend on the characterisation of time: as continuous or discrete. This determines whether the equations of the model are represented as differential or difference equations. Either form of dynamic equation can be presented and analysed within a flowgraph representation, but since the stronger tradition in economics is that of discrete time, the dynamic models here are formulated in discrete time as difference equations. In practice this means that the models may contain the lag operator L. Where it is encountered, L should be treated as an algebraic parameter. This means that in the flowgraph representation of dynamic models certain arcs will have an arc transmittance involving L, or quite often $(1–\mathrm{L})^{-1}$ which implies accumulation or time-integration.

For example, consider a model in which the rate of change of p (the logarithm of the price level) is inversely proportional to u (the rate of unemployment): $\Delta p = p_t - p_{t-1} = -au$. With the use of the lag operator, this translates into $p(1–\mathrm{L}) = –au$, or

$$p = \frac{-au}{1-\mathrm{L}},$$

allowing the elementary graph representation depicted in Figure 1.7.

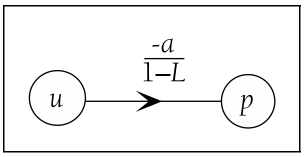


Figure 1.7 Dynamic arc

iii. Parameter variation

It is often of interest to consider the implications of some parameter within an overall transmittance taking a particular value. This is especially so for the lag operator L, because L =0 corresponds to the "shortrun" outcome while $L = 1$ corresponds to the "long-run" or "equilibrium" outcome. Consider the latter case. Often the overall transmittance will be a ratio in which both the numerator and the denominator may contain terms involving $(1–L)^{-1}$. Then it is usually helpful to re-express the transmittance by multiplying both the numerator and the denominator by (1–L) before setting L=0 or L=1. This avoids "infinities".

Zeros and infinities in individual arc transmittances alter the topography of the flowgraph. When an arc transmittance goes to zero the flowgraph is simply modified by removing that arc. But the more interesting and more difficult case is when an arc transmittance goes to infinity. Then the only paths that matter are the loops and transmittances which include that arc because its effect dominates all others. It is often the case that an arc transmittance which goes to infinity is part of a loop. Then we restrict attention to those paths whose transmittances T_i or cofactors Δ_i include that arc, because otherwise the path transmittance would effectively go to zero. Similarly we consider only those loops or non-touching loop pairs that contain the arc. In the example of Figure 1.6 above, let $b \rightarrow \bullet$, then the direct path from U to Z that bypasses node W can be ignored and similarly the two loops, between W and X and between V, Y and Z, can be ignored. Thus Mason's gain formula for the transmittance from U to Z becomes $abcde/(1–bcdeh)$ which simplifies to $–a/h$ on dividing both numerator and denominator by the infinitely large b.

When an arc transmittance goes to infinity the graph changes as follows: the arc is removed, the initial node of the arc is exogenised and the final node of the arc is endogenised through a new route. The last two steps involve a reversal of causality. If there are several possibilities for the new route, then the topography of the derived system is ambiguous.

iv. Reversal of causality

This means that at least one variable which was previously considered to be endogenous is now made exogenous. That is to say, its value is fixed outside the model. For each such variable whose role has changed another, previously exogenous variable, must now be considered endogenous, with its value determined by the system under study. Of course we could just go back to the equations, and manipulate them so that the new set of endogenous variables appears on the left-hand side, and construct a revised graph from the equations. But it is also possible to manipulate the existing graph directly.

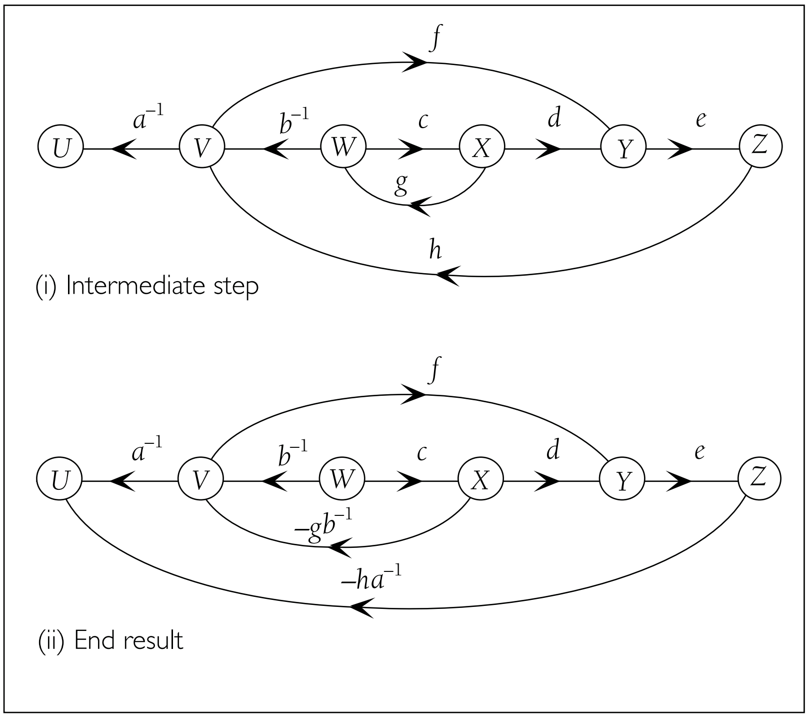


Figure 1.8 Reversal of causality

This is illustrated in Figure 1.8 which reverses the causality between the W and U nodes of the flowgraph of Figure 1.6. The procedure is as follows: first select the two variables whose roles are to be reversed and find a path connecting the two nodes (if there is more than one such path, then a choice must be made and the eventual outcome graph will not be unique); next reverse all the arrows of the arcs along that path and invert all the associated arc transmittances; next change the sign of the transmittance of any arc that converges onto any node in the new path; finally shift the end points of all these converging arcs to the next node in the new direction, preserving the new transmittances to those nodes.

1.7 Brief bibliography

The representation of linear models as flowgraphs was introduced into the world of systems engineering by Samuel Mason in the following papers.

S. J. Mason, "Feedback theory: Some properties of signal flow graphs," *Proc. IRE*, vol. 41, pp. 1144-1156, Sept. 1953.

S. J. Mason, "Feedback theory: Further properties of signal flow graphs," *Proc. IRE*, vol. 44, pp. 920-926, July 1956.

The techniques of flowgraph manipulation and solution are discussed in many textbooks of control systems analysis, including the following as representative examples.

B. C. Kuo, *Automatic Control Systems*, 7th ed., 1995, New York: Prentice Hall.

R. C. Dorf, *Modern Control Systems*, 8th ed., 1998, New York: Addison Wesley.

F. H. Raven, *Automatic Control Engineering*, 5th ed., 1995, New York: McGraw Hill International.

J. J. D'Azzo and C. H. Houpis, *Linear Control Systems Analysis and Design*, 4th ed., 1995, New York: McGraw-Hill

R. T. Stefani, C. J. Savant Jr., B. Shahian and G. H. Hostetter, *Design of Feedback Control Systems*, 3rd ed., 1994, Boston: Saunders College Publishing.

C. S. Lorens, *Flowgraphs for the Modeling and Analysis of Linear Systems*, 1964, New York: McGraw-Hill.

Of these, Kuo's is the most extensive presentation, and in it he emphasises the fact that Mason introduced the signal-flow graph as a cause-and-effect representation. But the connection to a causal interpretation appears to be of minor importance to most modern control systems engineering textbooks, which see the advantages of the flowgraph representation firstly as a more elegant and simpler picture than the block diagram, secondly as a means of solution via Mason's rule as an alternative to step-by-step reduction of block diagrams, and thirdly as a vehicle for the representation of state space methods in systems of differential equations. The short monograph by Lorens is the source of the causality reversal technique outlined in this chapter.