CrossMark

The ontological status of shocks and trends in macroeconomics

Kevin D. Hoover

Abstract Modern empirical macroeconomic models, known as structural autoregressions (SVARs) are dynamic models that typically claim to represent a causal order among contemporaneously valued variables and to merely represent non-structural (reduced-form) co-occurence between lagged variables and contemporaneous variables. The strategy is held to meet the minimal requirements for identifying the residual errors in particular equations in the model with independent, though otherwise not directly observable, exogenous causes ("shocks") that ultimately account for change in the model. In nonstationary models, such shocks accumulate so that variables have discernible trends. Econometricians have conceived of variables that trend in sympathy with each other (so-called "cointegrated variables") as sharing one or more of these unobserved trends as a common cause. It is possible for estimates of the values of both the otherwise unobservable individual shocks and the otherwise unobservable common trends to be backed-out of cointegrated systems of equations. The issue addressed in this paper is whether and in what circumstances these values can be regarded as observations of real entities rather than merely artifacts of the representation of variables in the model. The issue is related, on the one hand, to practical methodological problems in the use of SVARs for policy analysis—e.g., does it make sense to estimate of shocks or trends in one model and then use them as measures of variables in a conceptually distinct model? The issue is also related to debates in the philosophical analysis of causation—particularly, whether we are entitled, as assumed by the developers of Bayes-net approaches, to rely on the causal Markov condition (a generalization of Reichenbach's common-cause condition) or whether cointegration

K. D. Hoover (✉)
Department of Economics and Department of Philosophy, Duke University, Box 90097,
Durham, NC 27708-0097, USA
e-mail: kd.hoover@duke.edu

🕉 Springer

generates a practical example of Nancy Cartwright's "byproducts" objection to the causal Markov condition.

1 Shocks and trends

In a preamble to a criticism of the use of the causal Markov condition in causal modeling, Nancy Cartwright tells us:

> Metaphysics and methodology should go hand in hand. Metaphysics tells what something is; methodology, how to find out about it. Our methods must be justified by showing that they are indeed a good way to find out about the thing under study, given what it is. Conversely, if our metaphysical account does not tie in with our best methods for finding out, we should be suspicious of our metaphysics. (Cartwright 2007, p. 132)

While I do not share her views of the causal Markov condition, the idea that metaphysics and methodology must go hand in hand seems right. My aim here will be to explore the interaction of the metaphysics and methodology of macroeconometrics in a specific case.

Economics is a data-rich, and economic data display complex interrelationships both contemporaneously among the values of different variables and intertemporally between values of the same or different variables at different times. These relationships may be captured mathematically in a variety of representations that equally "save the phenomena" (van Fraassen 1980, ch. 3). Yet it is widely agreed that economics—especially since it frequently aspires to be a source of normative guidance or policy advice—cannot be indifferent among these phenomenologically equivalent representations but should privilege representations that support counterfactual analysis and provide a map of effective interventions in the economy. Consequently, economists are committed implicitly to an ontology of causes. Indeed, some economists maintain that the distinction between *econometrics* and *statistics* applied to economic data is the distinction between methods suitable for identifying and measuring *causes* as opposed to methods suitable only for identifying and measuring *associations* (Heckman 2000, p. 45; Hoover 2004).

Econometric models that aim to capture causal relations in support of counterfactual analysis are referred to as *structural* and center on a decomposition into a deterministic structure, which represents the causal connections among variables and unobservable stochastic terms that are the entry point for probabilistic considerations. At least implicitly, economists are concerned with systems of such equations, conceived of in a way that maps fairly cleanly onto to so-called Bayes-net or graph-theoretic analyses of causation (Spirtes et al. 2000; Pearl 2000).[1]

---

[1]   Most discussions of Bayes nets focus on directed acyclical graphs, in which causation is strictly one-way. But this is not essential, and it is characteristic of structural econometrics—going back to its foundations

Epistemological issues complicate economists' understanding of causal ontology. The fundamental vision of many economists can be traced to Walras's (1874[1954]) formalization of general equilibrium—i.e., of the systemic interdependence of economic agents such that directly or indirectly every economic quantity depends on every other economic quantity. The general-equilibrium vision inspires the Walrasian fantasy of a complete agent-by-agent representation of the economy, not dissimilar to Laplace's fantasy of the evolution of the physical universe as the unfolding of a set of initial conditions through a complete set of deterministic laws (Hoover 2001, p. 102).

There is a tension between this fantasy and the need to keep actual economic analysis manageable. A practicable structural equation may relate an effect to its causes and a residual. For example,

$$z = \alpha x + \beta y + \varepsilon, \tag{1}$$

where x, y, and z are variables, α and β are parameters, and ε is an unobservable residual; causes are written on the right-hand side of the equal sign; effects on the left.[2] The residual ε may represent at least two distinct types.[3] The first type is *error*; and there are at least two types of error: measurement error and specification error. Measurement error is straightforward. An example will clarify the idea of specification error: The functional form of (1) is linear. But if this is only an approximation to a true nonlinear structure, then ε will capture the approximation error. Both measurement error and specification error are artifacts of representation that raise practical issues for econometrics, but I wish to set them to one side. The residual may also represent a second type, namely *a genuine substantive causal feature of the world*—also in at least two ways. First, variables that are omitted from the structural representation may, in effect, be impounded in the residual. Such an interpretation is invoked when it is said that the residual reflects causes that are not the primary focus of an investigation. Second, the residual may, in fact, simply reflect fundamental indeterminism—the manner in which the probabilistic character of the causal relationship is represented. In the first case, ε is a portmanteau term for omitted or latent variables that are real causes of z. In the second case, ε is part of the fundamental characterization of z itself.

In setting aside the first type of issue, I am by no means wishing to minimize the practical and epistemological problems involved with residuals representing error. In the infancy of econometrics, Frisch (1938) recognized both that residuals in practice mixed error with substantive causal factors and that practical techniques for separat-

---

Footnote 1 continued
in Haavelmo's (1943, 1944) seminal work—to be concerned with simultaneous (mutual) or cyclical causal relations.

[2] "Causes on the right" is a common, but not always respected, convention in econometrics. It is sometimes helpful to represent the causal relationship more explicitly: Cartwright (2007, p. 13) suggests "$\overset{c}{=}$" and Hoover (2001, p. 40) suggests "$\Leftarrow$" as directed *causal* equalities.

[3] The interpretation of the residual as error or as a substantive causal factor has a long history in econometrics, going back, at least, to Frisch (1938). See Duarte and Hoover (2012); also Qin and Gilbert (2001).

ing the two types were essential. Frisch's student, Trgyve Haavelmo in his famous foundational monograph, *The Probability Approach in Econometrics* (1944), laid the groundwork for those practical techniques, pointing to the need for models that are both statistically adequate to control error and theoretically identified (to reflect an accurate causal articulation). [4] The first type of issue is real and important, but not the target of my current concern. Hoover (2009) argues that causal inference from observational data requires a two-step procedure of first securing a statistically adequate model from which probabilities may inferred and then using those probabilities to support causal inference.[5] No practical procedure for causal inference can successfully implement the second step without having first successfully implemented the first step. Nonetheless, henceforth, I take it for granted that inference is conducted within a framework of statistically adequate models and that the issues to hand arise in the context of the second step, the problem of inferring causal relationships, and not in the context of the first step, the problem of sound inference of probabilities from statistics.

With respect to the residual interpreted as a substantive causal feature of the world, there may be no practical, operational distinction between the first, epistemological interpretation and the second, ontological interpretation. It would only be in the impossible situation in which we knew *all* of the distinct causes of z that we could know which of the two applied in a particular case. Economists refer to residuals under either interpretation as *shocks*. Shocks are not noncausal artifacts, but causally significant, unobserved or unobservable real features of the world. In practice, residuals could be a composite of shocks (omitted causes + fundamental indeterminism) and (measurement + specification) error. Our interest here, however, is only with shocks.

Equation (1) is a contemporaneous or static relationship in which we assume that variables are not time dependent. In contrast, an example of a time-dependent or dynamic relationship is provided by the stochastic difference equation:

$$z_t = \alpha x_t + \beta x_{t-1} + \gamma z_{t-1} + \varepsilon_t, \tag{2}$$

where γ is a new parameter and the subscripts indicate time periods $t = 0, 1, 2, \ldots$ If there are no issues of approximation or specification error and (2) is causally apt, then the residual ε_t represents a shock. The cumulation over time of such shocks forms a (stochastic) *trend*.[6]

One version of scientific realism is characterized by the assertion that the theoretical posits of our best scientific explanations are real. Some critics of scientific realism ground their criticism in the unobservability of some of these entities (van Fraassen 1980). There is something question-begging in grounding the criticism in a stipulation

---

[4] On the connection to Frisch and Haavelmo, see particularly Spanos (1989, p. 408, and section 5) and Spanos (1990, section 3).

[5] On practical techniques for securing statistical adequacy, see Spanos (1990, 1995), Mayo and Spanos (2013), Hendry (1995), Hoover et al. (2008) and Hoover (2013).

[6] There can also be deterministic trends in which the value of a variable depends noncausally and deterministically on time. Our concern here is only with stochastic trends.

that observation is restricted to cases involving the unaided senses. Not only is that a hard line to draw (do I cease to observe when I put on my glasses, though I observe precious little without them), it also defies the common usage of ordinary people and scientists who include in "observation" knowledge gained through acquaintance mediated through various instruments. Observations may be obtained through shorter or longer chains of instrumentation and with greater or lesser cooperation from other theories and models. At a rather ill-defined stage, indirect observation is perhaps better termed "inference". Yet, neither observation nor indirect inference makes sense without something real to be observed or inferred. A fundamental question for econometric analysis is whether shocks and trends are real and, if so, whether and how they can be observed or inferred.

2 Recovering shocks

2.1 Causal sufficiency and the identification of shocks

Early econometrics typically treated shocks as a nuisance and focused on the problem of estimating the parameters of structural equations.[7] To keep the analysis practically manageable variables were divided into *endogenous* variables to be "explained" by the econometric model and *exogenous* variables that were regarded as having been determined outside the model. Counterfactual analysis was mainly thought of as assessing the paths of endogenous variables given alternative paths of exogenous variables—either on the assumption that the shocks would average out to zero or that their variability would be translated into the probability distribution of the endogenous variables.

Given the dominant Walrasian worldview, the distinction between endogenous and exogenous variables was a best pragmatic. In principle, since everything depended on everything else, all variables should be endogenous. Once exogenous variables are abandoned, counterfactual analysis must come down to either working out the consequences of alternative choices of parameters or tracing through the effects of shocks. Shocks can no longer be regarded as a nuisance but take pride of place as the fundamental drivers of the variables within a *regime* defined by a constant set of parameters. But shocks are epistemically opaque.

The problem can be easily understood by considering Eq. (1) as a single equation embedded in a *directed acyclical graph*, the mainstay of the Bayes-net approach. Think about z as sitting at the bottom of the causal order: x and y cause z, but z does not cause x or y. Econometricians typically have observations on the variables and have to estimate the parameters. Typical estimation methods amount to probabilistic conditioning, so that

$$z = \hat{\alpha}x + \hat{\beta}y + \hat{\varepsilon}, \tag{3}$$

---

[7] Haavelmo (1944), Koopmans (1950), Hood and Koopmans (1953). See Duarte and Hoover (2012) for a history of the role of shocks in econometrics).

where the "hat" over the parameters and the shock indicate that these are estimates formed by conditioning, and the shock is governed by a probability distribution.[8] The key point is that the shock ε is not directly observable, but can be indirectly inferred—provided that (1) is causally apt—by backing an estimate of it out of (3):

$$\hat{\varepsilon} = z - \hat{\alpha}x - \hat{\beta}y. \tag{4}$$

The epistemic opaqueness arises from the fact that the data alone may not provide enough information to guarantee that (1) was a causally apt characterization of the structure in the first place. Suppose that z were determined by (5) instead of by (1), in which now x is still causally upstream but y is causally downstream from z:

$$z = \alpha x + \varepsilon. \tag{5}$$

An estimate of (5) based on the same set of data as used in (4) yields

$$\tilde{\varepsilon} = z - \tilde{\alpha}x, \tag{6}$$

where the tildes indicate estimates. The critical point is that in general $\hat{\alpha} \neq \tilde{\alpha}$ and $\hat{\tilde{\varepsilon}} \neq \tilde{\varepsilon}$. We require additional knowledge about which structure (1) or (5) is correct. Where is that to come from?

In practice, the most common modern macroeconometric models, known as *structural autoregressions* (or *SVARs*), involve systems of equations like (2) in which all variables are endogenously determined. The key issues at this point do not involve time dependence or dynamics, so it is easier to think of systems of equations like (1):

$$x = \alpha_{xy}y + \alpha_{xz}z + \varepsilon^x, \tag{7a}$$
$$y = \alpha_{yx}x + \alpha_{yz}z + \varepsilon^y, \tag{7b}$$
$$z = \alpha_{zx}x + \alpha_{zy}y + \varepsilon^z. \tag{7c}$$

System (7) can be written more compactly in matrix notation as

$$\mathbf{AX} = \begin{bmatrix} 1 & -\alpha_{xy} & -\alpha_{xz} \\ -\alpha_{yx} & 1 & -\alpha_{yz} \\ -\alpha_{zy} & -\alpha_{zy} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \varepsilon^x \\ \varepsilon^y \\ \varepsilon^z \end{bmatrix} = \mathbf{E}. \tag{8}$$

Each of the rows of $\mathbf{A}$, which is the first matrix after the first equal sign, corresponds to the parameters of one of the Eq. (7a–c).

The epistemic opacity of system (7) or (8) is reflected in part in the fact that the variables do not provide enough independent information to estimate all of the parameters

---

[8] The distribution is typically assumed to have mean of zero, which can be guaranteed by including a constant in the conditioning set, which is equivalent to an innocuous normalization of the measurements of the variables; and the distribution is often assumed to be normal or Gaussian, which can be justified in many circumstances by considerations that need not concern us here.

Fig. 1 A causal graph
recoverable using Bayes-net
algorithms
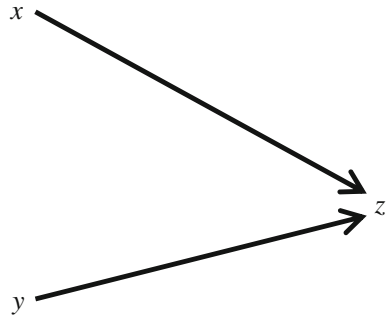
of the system. The classic approach of the Cowles Commission is to *identify* the para-
meters by restricting their number at least to be no greater than the available distinct
sources of information provided by the data (Koopmans 1950; Hood and Koopmans
1953). This amounts to assuming a priori that some of the parameters (α_{ij}) equal zero
and/or some of the correlations among the shocks vanish.[9] The approach is rational-
istic in that it is supposed to be grounded in theoretical considerations independent of
the available empirical evidence.

Bayes-net approaches (e.g., the PC algorithm) provide an empirical approach to
identification—essentially placing zeroes according to the evidence of various condi-
tional independence relationships discoverable from the data themselves (Spirtes et al.
2000, ch. 5, Cooper 1999). The best case for the application of Bayes-net algorithms
occurs when the data set is causally sufficient. Data are *causally sufficient* when every
variable that causes more than one variable in the data set is included in the data set.
Causal sufficiency guarantees that, if the system of equations excludes any variable
from explicit consideration, it can be impounded in *one* of the shocks and that the
shocks are not probabilistically dependent on each other. Causal sufficiency is a strin-
gent condition. But even when causal sufficiency obtains, if the unobservable real
underlying structure does not contain zero parameters ($\alpha_{ij} = 0$ for some variables
$i \neq j$) or does not place the zero parameters in the right places, a search algorithm
may be unable in principle to recover the structure completely from the data.

If the data are causally sufficient and if the causal structure of the system is known—
either a priori or empirically—then the shocks can be easily be recovered from the
data just as in Eq. (4). Figure 1 shows a causal graph that is in fact recoverable with
Bayes-net algorithms for acyclical graphs. It corresponds to the estimated version of
(8) with possible but missing causal linkages indicated by zeroes in the estimated **A**
matrix:

$$\hat{\mathbf{A}}X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\hat{\alpha}_{zy} & -\hat{\alpha}_{zy} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \hat{\varepsilon}^x \\ \hat{\varepsilon}^y \\ \hat{\varepsilon}^z \end{bmatrix} = \hat{\mathbf{E}}, \qquad (9)$$

where the "hats" once again indicate estimated values.

---

[9] Other sorts of restrictions may also be identifying so long as they reduce the dimensionality of the
parameter space to be no greater than the dimensionality of the data space.

2.2 Macroeconomics and the reality of shocks

The central problem is to understand how we should think about the estimated shocks: Are they real? Or are they simply artifacts of the representation and estimation procedure? These are not merely philosophers' questions, but are in fact questions that economists themselves raise—often only implicitly, but sometimes explicitly. A good example of addressing the questions explicitly is found in Glenn Rudebusch's "Do Measures of Monetary Policy in a VAR Make Sense?" (1998a).[10]

For Rudebusch (1998a, p. 907) monetary policy analysis—that is counterfactual analysis—would be easy in the context of a complete structural model. The SVAR model is one that focuses on the causal articulation of only the contemporaneous variables, as in system (7) or Eq. (8), but allows the dynamic structure (omitted from our examples) to reflect only the associations or correlations among variables rather than the complete causal articulation. Focusing on a single equation in the SVAR for the variable that represents the monetary policy instrument (usually the overnight interbank borrowing rate of interest known as the Federal funds rate), Rudebusch (p. 908) asks two questions: First, does the equation correctly represent the reaction of monetary policymakers—that is, does it contain the correct causal variables and estimate the correct parameters? Second, do the residuals plausibly represent the shocks to monetary policy? Rudebusch argues that SVARs are useful guides to monetary policy only if both questions can be answered, yes. He is, in fact, skeptical.

One of Rudebusch's (p. 908) concerns is that parameter estimates in the analogue to (7) and (8) may not be stable over time. Though he does not analyze this issue further, there are two quite distinct possibilities. First, the system may be an accurate causal articulation, but through a policy action or a "natural experiment," the underlying parameter has actually changed. An estimate of the system that failed to account for such a change would generate parameters that were not structural, but artifacts that combined the information of distinct regimes. The issues are, however, not the ontological ones of the correctness of the causal structure or the reality of the shocks, but are epistemological or methodological. On the epistemological side, the values assigned to the parameters and shocks are questionable. On the methodological side, we are pointed to the necessity of properly accounting for regime change.[11]

The second possibility is different. If the system fails to be causally sufficient, the actual estimates of parameters and residuals are really just complex conditional associations without straightforward causal interpretations. Worse still, such estimates need not be stable, as changes in variables omitted from the causal representation may induce changes in these conditional associations. The residuals cannot legitimately be regarded as shocks, because they lack the causal articulation that is the hallmark of a real variable. The observed instability of the estimates is, in this case, a signal of their unreality. These two possibilities are conceptually distinct. One of the central difficulties of econometrics is that they are hard to distinguish in practice.

---

[10] Sims (1998) comments on the original paper, and Rudebusch (1998b) replies.

[11] In fact, Hoover (2001, chs. 8–10) investigates methods of using such changes of regime within a stable causal structure to extract information about the causal structure itself.

The second source of instability, as it is grounded in the failure of causal sufficiency, is related to Rudebusch's (p. 916) more general concern with the omission of structure. It is possible that a simpler system could be nested within a more complex system and yet both be causally sufficient. But such a nesting relationship is very special, and a general worry is that any overly simplified system would not be causally sufficient and that, as a result, the residuals could not be legitimately taken to represent shocks. It is puzzling and inconsistent that Rudebusch (p. 926) argues that the failure of different SVAR specifications to agree amongst themselves on the actual values for monetary-policy shocks "would seem to be a critical failing". The notion that estimates are more reliable if they are robust to alternative specification is an appealing one to many economists and to some philosophers, but it is hardly defensible if the estimates are meant to characterize real causal parameters and shocks; for alternative specifications will rarely all satisfy causal sufficiency (see Cartwright 1991; Hoover 1995; Hoover and Perez 2000). Rudebusch (p. 928) comes round to endorsing this view indirectly when he observes that the failure of different SVARs to agree on the shocks "may reflect the fact that different authors have mined the data in different ways…"

What could be evidence for the success of an SVAR at recovering real shocks? Rudebusch argues that consistency of SVAR evidence with evidence from a completely different source would add weight to the identification. The suggestion sounds similar to the robustness standard that we just rejected, but is really quite different. It is more closely related to the variety-of-evidence thesis (see Claveau 2013) and is exemplified by the well-worn example, beloved of philosophers of science, in which it is claimed that the reality of molecules is underwritten by the multitude of different methods that produce consistent estimates of Avogadro's number (Perrin 1913[1990]). Rudebusch's deployment of the standard does not lead to the same happy outcome: he claims that methods of estimating shocks based on financial markets and the historical record of monetary policy actions by the Federal Reserve fail to conform to the shocks typically recovered from actual SVARs (pp. 920–927).

Related to using a *variety of evidence* as a desideratum of ontological independence, Rudebusch (p. 928) suggest another standard that is both evidence for, and a valuable property of, independent shocks—namely, *transferability*. In particularly, he suggests that if we are successful at identifying ontologically independent shocks, then the estimates of these shocks themselves may be used as instrumental variables in estimating the coefficients of other causal structures. This is an example of Hacking's (1983, pp. 22–24) criterion "if you can spray them, then they are real," used as evidence for the ontological independence of unobservable posits of theory. The key idea is that an estimated value may be theory-ladened (or in Rudebusch's case, model-ladened), but its ontological independence is established by using it (the "spraying") in an independent context.

Sims (1999) provides an example in which the ontological independence of the shocks is assumed. He poses a counterfactual question: Had the Federal Reserve managed monetary policy the same way that the modern Federal Reserve under Alan Greenspan's chairmanship managed monetary policy, would the Great Depression have been avoided? His approach is to estimate an SVAR for both a period that includes the Great Depression as its last subperiod and for the Greenspan era (1987-end of the sample). In each, he identifies the equation for a short-term rate of interest as the

monetary-policy rule. He substitutes the rule for the Greenspan era into the system for the Great Depression. Finally, using the shocks recovered from the original SVAR for the Great Depression, he simulates the paths for important economic variables to see how they would have behaved facing the same external conditions (i.e., the shocks) but confronting them with a latter-day monetary policy. The counterfactual experiment is problematic in that there are good reasons to question the legitimacy of the assumption that the rest of the SVAR for the Great Depression would remain stable in the face of the changed monetary-policy rule (see Hoover and Jordà 2001; Hoover 2011). Nonetheless, the main point for our purpose is the implicit assumption that the estimated shocks for the Great Depression can be legitimately treated as observed variables to drive the counterfactual simulation.

3 The ontology of latent variables

Sims's use of shocks in counterfactual analysis presupposes that the SVAR from which they were estimated was a causally apt articulation of a causally sufficient set of variables. Causal sufficiency, however, is a stringent condition that may be hard to fulfill and that may not be readily verifiable. Little effort is, in fact, made to verify it in typical econometric applications. Causal sufficiency can easily fail if a common cause is omitted from the SVAR—either a variable that could be easily measured or one that is truly *latent* (i.e., one that is causally important but for which there is no independent method of measurement). Reichenbach's (1956, p. 157) *principle of the common cause* states that if two variables are probabilistically dependent either one causes the other or there is a common cause of both. The omission of a common cause, therefore, may result in treating variables as directly causally connected that are, in fact, probabilistically independent conditional on their common cause.

Bayes-net causal search algorithms based on the causal Markov condition, a generalization of Reichenbach's principle, may discover the traces of omitted variables in the patterns of probabilistic dependence and independence among variables (Spirtes et al. 2000, ch. 10). What is important at this juncture is not the details, but the fact that this may present an opportunity for recovering a latent common cause. For example, suppose that as in Fig. 2, C is a common cause of A and B, then Reichenbach's principle supposes that conditioning on C would render A and B probabilistically independent. C is said to *screen off* A from B (and vice versa).

Cartwright (1993, section 3) objects that we have no right to assume that C will act as a screen. In her particular example, she treats A and B as the deterministic byproducts
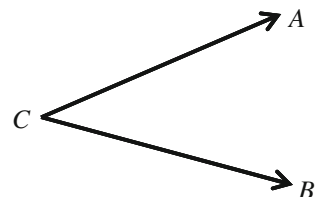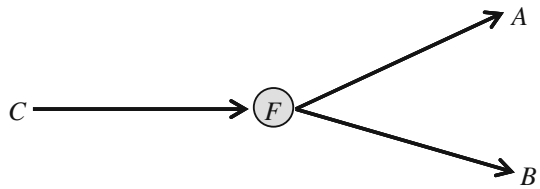
Fig. 2 *A causal graph in which C screens off A from B*

Fig. 3 Screening-off fails when the causal connection of *C* to *A* and *B* is indeterministic (*F* indicates *C fires/does not fire*)

of a process with indeterministic causal connections.[12] In the Bayes-net literature, to take a well known case, causes and effects may themselves be probabilistic, but the causal connection (what is represented by the arrow itself in a causal graph) is regarded as deterministic. In Cartwright's example, when *C* is *off*, neither *A* nor *B* is produced; when *C* is *on*, *A* and *B* are produced only with some probability less than one, and *A* is produced if, and only if, *B* is produced—that is, there is a deterministic relationship between *A* and *B* and, independent of the randomness of *C* itself, there is an additional source of randomness (namely, the causal stimulus (*C* being on) is transmitted only with a probability less than one). In effect the causal connection is regarded as indeterministic. In this case, *C* does not screen off *A* and *B*, because *C* being *on* provides only probabilistic knowledge that *A* or *B* is produced; while the additional knowledge that *A* is produced makes it certain that *B* is produced.

Cartwright's assumption that *A* is produced if, and only if, *B* is produced is inessential to her point. It is enough that *A* and *B* be correlated. The essence of the point is that she envisages the causal linkage itself to be indeterministic, and in such as case, the probability of *B* conditional on *C* and *A* will be greater than the probability of *B* conditional on *C* alone, even if the former probability is less than one. Cartwright challenges the ontological assumptions at the heart of the econometrician's and the Bayes-net practitioner's partitioning of the world into structure (i.e., the deterministic relationships among variables) and shocks (the bearer of probabilistic variation). They see the arrows in Fig. 2 as deterministic linkages; Cartwright sees the arrows themselves as probabilistic.

Papineau (1991, p. 406) analyzes such situations differently. He argues that there is an omitted variable (*F*) that indicates whether or not the causal linkage is operative (i.e., when the cause has "fired"). So, Fig. 3 replaces Fig. 2, and the arrows correspond to deterministic links. If *F* is itself stochastic (i.e., if it is driven by shocks not shown in the figure), then naturally *C* will not, just as Cartwright suggests, screen off *A* from *B*. But *F* will act as a screen. Cartwright replies simply that no *F* may be available to us: wanting it does not make it so.

Surely, Cartwright is correct that we may not have any independent method to observe the necessary *F*. But could it be a latent variable? If we are willing to assume that the causal Markov condition applies, then we could easily calculate what *F* would have had to be in order to serve as a screen. In other words, we could use the relationship between *A* and *B* as a device to measure or indirectly observe *F* (cf. Boumans 2005, ch. 5, appendix). How should we regard the output of such a

---

[12] See Hoover (2001, pp. 92–97) for a detailed discussion of Cartwright's case; see also the related discussion on pp. 53–54 on the meaning of "probabilistic causality"; cf. Humphrey's (1989, pp. 46–47).
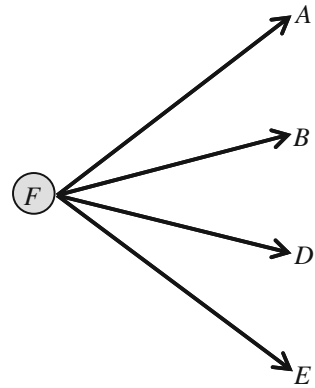
measurement? One possibility is that F is just an artifact: we can always calculate it ex post, but it has no ontological independence of its own. Another possibility is that we can show that the F calculated in one context does work in another. For example, suppose that we recover F from A and B and that we then find that our estimated F screens off the probabilistic dependence of two different variables D and E. This might suggest that, ignoring C, the data were related as in Fig. 4. Let $\hat{F}_{AB}$ be the estimate of F recovered as the screen of A from B and analogously for other variables. The ontological independence of F could receive substantial support if it turns out that $\hat{F}_{AB} \approx \hat{F}_{AD} \approx \hat{F}_{AE} \approx \hat{F}_{BD} \approx \hat{F}_{BE} \approx \hat{F}_{DE}$. This could be seen as an example of the evidential variety or transferability that were regarded as hallmarks of ontological independence in Sect. 2.2. Such a finding would go a long way to undermine the salience of Cartwright's skepticism of the implicit causal ontology shared by structural econometrics and Bayes-net approaches.

The logic of extracting a latent variable as an implicit common cause is essentially one of finding what a variable would have to be to turn a causally insufficient set of variables into a causally sufficient set. It can work only when, in addition, we are lucky enough that the causal structure (or at least the relevant part of it) is recoverable. Bayes-net methods, for instance, sometimes are able to recover only an equivalence class of causal structures, rather than a unique structure.[13] The approach requires a good deal of systemic knowledge and relies on the delicate statistical tests for conditional independence, which can mislead when the departures from causal sufficiency are too great. Recovering ontologically independent shocks from SVARs thus poses a significant methodological challenge for causal modeling. The matter may, as we will suggest presently, be more tractable if we turn from recovering shocks to recovering trends. Since trends are accumulations of shocks, we must first examine the dynamics that we had previously set to one side.

---

[13] Acyclic graphs that have the same skeletons (i.e., the same variables are connected by causal arrows without attention to their direction) and the same unshielded colliders are observationally equivalent (Pearl 2000, theorem 1.2.8, p. 19). An unshielded collider is a pattern in which two variables that are not connected directly are the source of causal arrows that point into the same third variable. Figure 1 is an example of an unshielded collider.

4 Dynamic processes

4.1 Causal ambiguity in dynamic models

Dynamic processes are often complex, but the main points can be made with some simple examples. A very simple univariate dynamic process is one in which the value of a variable (x) in one period (t) depends on its value the period before plus a shock (ε):

$$x_t = \rho x_{t-1} + \varepsilon_t, \tag{10}$$

where ρ is a parameter. If $|\rho| < 1$, then x may be represented equivalently as a weighted sum of all the past shocks where the weights decline geometrically:[14]

$$x_t = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}. \tag{11}$$

In principle, (11) shows that any shock has a permanent effect on the value of x; but the fact that the weights decline, mean that practically the "memory" of the shock is wiped out relatively quickly.

Roughly speaking a *stationary* time series is one for which the moments (e.g., the mean and variance) are finite and independent of time, and for which the covariance (or correlation) between its values at different times depends on how far apart those times are but not on the absolute time (e.g, the calendar date). (See Hendry 1995, p. 42 for a more formal definition). If the shock ε_t is drawn from an identically, independent normal distribution, then it fulfills the definition of a stationary time series. But in fact, so does x_t in (10) or (11): the expressions for the mean, variance, and covariances will be more complicated, but they fulfill the criteria of stationarity nonetheless.

Consider a dynamic process involving two variables:

$$x_t = \beta y_t + u_t, \tag{12a}$$
$$u_t = \rho u_{t-1} + \varepsilon_t. \tag{12b}$$

The value of the variables in (12a) can be given an equivalent representation:[15]

$$x_t = \rho x_{t-1} + \beta y_t + \rho \beta y_{t-1} + \varepsilon_t. \tag{13}$$

If we define $\Pi_1 = \rho$, $\Pi_2 = \beta$, and $\Pi_3 = \rho\beta$, (13) may be rewritten as:

$$x_t = \Pi_1 x_{t-1} + \Pi_2 y_t + \Pi_3 y_{t-1} + \varepsilon_t, \tag{14}$$

subject to the restriction that $\Pi_1 \Pi_2 = -\Pi_3$.

---

[14] Equation (11) follows from (10) by lagging (10) and substituting the expression into the right-hand side of (10) to eliminate for x_{t-1}. This introduces a term $\rho^2 x_{t-2}$, which is eliminated in the same way. After n such substitutions, we are left with an expression involving the shocks and $\rho^n x_{t-n}$. As $n \to \infty$, this term approaches zero, and the expression on the right-hand side of (11) remains.

[15] To derive (13), lag (12a) one period, multiply by both sides ρ, subtract the result from the original (12a); then substitute from (12b) and rearrange to get (13).

Both the system (12) and Eq. (14) reproduce the same path of values for x. They are phenomenally equivalent. But they are not ontologically equivalent. In (12a) u is a real variable—perhaps omitted from explicit analysis or latent, but a genuine cause with its own dynamic history nonetheless. In (14) u has disappeared. Equation (14) is equivalent to (12a) but only in the very special case that the restriction on the Π_j's is fulfilled. If, for example, the values of the Π_j's were drawn at random, then the restriction could be fulfilled only on a set of Lebesgue measure zero. That is, not however, grounds to reject (14) as a legitimate representation; for there may be good economic reasons for the relevant restrictions to be fulfilled.[16] The key point is that we can always back u_t out of (12a); but what is its ontological status? If system (12) is an apt articulation of the causal structure, then this is a way to measure an ontologically independent latent variable. If Eq. (14) is the apt articulation of causal structure, then u_t is simply an artifact of the representation.

The problem posed by u_t will serve as an analogy for the more complex problem of trends. But before we can get there, we need to consider some other aspects of dynamics.

4.2 Nonstationary processes

Consider Eq. (10) with the modification that $\rho = 1$, so that

$$x_t = x_{t-1} + \varepsilon_t. \tag{15}$$

Then (11) would be replaced by:

$$x_t = x_{t-n-1} + \sum_{j=0}^{n} \varepsilon_{t-j}, \tag{16}$$

where x_{t-n} is an initial condition n periods earlier, which does not vanish as $n \to \infty$, since there is no weight [the equivalent to ρ^j in (11)] that gets smaller as n gets bigger—i.e., $\rho^j = 1$ for all j. Unlike in (11) a shock in (16) transmits it full effect permanently to x—it has perfect "memory."

A variable such as x_t is an example of a *nonstationary* time series. Its mean and variance are both functions of time (Hendry 1995, section 3.8). Yet, if we subtract x_{t-1} from both sides of (15) to yield

$$\Delta x_t = x_t - x_{t-1} = \varepsilon_t, \tag{17}$$

we see, since ε_t is a stationary random shock, that the first-difference of this nonstationary time series is itself stationary.

A time series that becomes stationary after being differenced once is said to be *integrated of order 1*—often written I(1). The order of integration is the number of times a time series must be differenced to achieve stationarity. Equation (15) is a

---

[16] See Hoover (2001, pp. 166–170) for an analogous case.

particularly simple case, as all that remains after differencing is a single random shock. But cases in which what remains is a more complex, yet still stationary expression, such as the right-hand side of Eq. (11) are equally cases of integration of order 1. A time series that is already stationary does not need to be differenced to achieve stationarity and so it integrated of order zero—that is, I(0).

A time series such as (15) is called a *random walk*. If x is the distance between an aimlessly wandering drunk and his front door, then (15) says that, at any step t, he is as likely to walk closer as further away. The best predictor of where he will be on the next step is just wherever he happens to be standing right now.

Two I(1) time series may be related. Imagine that the aimless drunk is followed by a deeply concerned friend. The friend follows a path dictated in part by his desire to walk safely, in part by his desire to keep out of sight, and in part by the overriding consideration not to lose sight of the drunk. The two friends may get closer or further apart, but they will never get too far apart. The path of the friend will differ in detail from the path of the drunk, but over the longer run, it is the path of the drunk that determines the main line of the path of the friend. If the distance between them increases too much, the friend accelerates his pace toward the drunk.

We can think of the path of the drunk as a *stochastic trend*: it is just the cumulation of his lurches and stumbles—the shocks to his position. The path of the friend can be thought of as stationary deviations from the trend. These must be stationary, because if they were nonstationary, the variance would not be finite and the friend would at some point, contrary to our initial assumption, drift arbitrarily far from the drunk. For the same reason, the distance between the drunk and the friend must be stationary. As a result the path of each friend is dominated by the single stochastic trend: the drunk generates the trend; the friend adjusts to it. And this is a general property. Imagine that the younger brother of the friend wonders just what is up and follows his brother (not the drunk), but keeps his distance.[17] Then his path will also be dominated by the stochastic trend of the drunk. There are now three paths—one trend and two adjusting. In any system of such variables, the number of trends plus the number of adjusting relationships adds up to the total number of variables.

Let x_t be the position of the drunk and y_t the position of the friend. The time series are said to be *cointegrated* if both are individually I(1) and a linear combination of them is stationary: $x_t - \beta y_t = $ a stationary or I(0) process. And if their positions are cointegrated, they will not drift too far apart.

Cointegration shares some features with correlation. Both are species of probabilistic dependence. When two time series are correlated, though not perfectly correlated, they move roughly together. Similarly, when two series are cointegrated they move roughly together, although this may be evident only if we take a longer time horizon. Thus, if we track the drunk and his friend for five minutes, we may not see a pattern to the friend's path; yet if we track them for an hour, we see that the friend never actually loses sight of the drunk.

---

[17] After having tweaked the commonplace example of cointegration as man following a drunk with the addition of a brother following the man, I discovered the structurally identical explication in Murray (1994), in which the drunk is followed by two dogs rather than two brothers.

The common statistics used to analyze correlation and many related notions require stationary data. Thus, Pearson's correlation coefficient—a mainstay of statistical analysis—is an exact analogue to the parameter that governs probabilistic dependence in a normal distribution, which is a stationary distribution. There is no similar analogy with a nonstationary distribution, which is the main source in macroeconomics of the phenomenon of "spurious correlation."[18]

Correlation is an intransitive relationship: e.g., if A and B are probabilistically independent and both cause E (a configuration referred to in the Bayes-net literature as an "unshielded collider"), A is correlated with E and E is correlated with B, but A is ex hypothesi not correlated with B. (Figure 1 is an example of an unshielded collider.) In contrast, cointegration is transitive: if A is cointegrated with E and E with B, then A and B must also be cointegrated. If this were not so, then A and B could drift infinitely far apart; but E cannot drift infinitely far from either of them; thus if E tracks A it loses B and if E tracks B it loses A—either way, a contradiction.

With cointegration, there is no analogue to conditional correlation. For example, if A, B, and C are stationary variables of a closed universe and each variable is connected to one, and only one, of the others by a deterministic causal connection, then each pair of variables (directly connected or not) is correlated and there is one of the variables such that conditioning on that variable eliminates the correlation between the other two. In a similar scenario in which the three variables are nonstationary and pairwise cointegrated, nothing similar happens. In fact, if two variables are cointegrated, the fact of their cointegration is robust to the introduction of further variables. A new variable may not be cointegrated with the original pair or it may be, in which case transitivity guarantees that the cointegration relationship is unaffected by the third variable.

It is worth noting that, although cointegration is similar in some respects and different in others from correlation, a variable may be both cointegrated with another and correlated with it. A variable can be decomposed into a stationary and a nonstationary component. It is certainly true that the nonstationary components of two I(1) time series may be correlated. It may also be true, though I have not seen this worked out, that there is a nonstationary analogue to correlation and conditional correlation. It is just that cointegration is not a candidate to serve as that analogue.

---

[18] The notion of spurious or nonsense correlations goes back to Yule (1926). Sober (2001) provides a time-series example. He suggests that two trending series, the price of bread in Britain and sea levels in Venice are correlated because they are both trending upward. This is *spurious correlation* because the series are trending and, therefore, nonstationary; so that ordinary correlation measures that were designed with reference to stationary probability distributions (e.g., Pearson's correlation coefficient) do not bear the same relationship to nonstationary distributions and do not in the nonstationary case indicate probabilistic dependence. See Hoover (2003) for a detailed discussion of Sober's example. Hoover (2009) and especially Hoover (2007) addresses spurious correlations that arise from a variety of sources. The common characteristic of all types of spurious correlation seems to be that they apply the wrong probability model to the data and, therefore, use unwarranted probabilities in the second step of the two-step procedure for causal inference mentioned in Sect. 1—which reinforces the importance of getting the first, statistical step right.

5 Recovering trends

5.1 Near decomposability and a hierarchy of causal structures

The decomposition of a nonstationary system of equations into a stationary and nonstationary part is closely related to Herbert Simon's (1996, ch. 8) analysis of hierarchies of complex systems based on the notion of near decomposability. A system is *decomposable* in Simon's account when some parts are tightly related to one another so that they form a unit, and the unit interacts with other units only as a whole. Decomposability is rare, yet near decomposability is more common. A system for Simon is *nearly decomposable* when

> (1) …the short-run behavior of the component subsystems is approximately independent of the short-run behavior of the other components; [and] (2) in the long run the behavior of any one of the components depends in only an aggregate way on the behavior of the other components. (Simon 1996, p. 198)[19]

It is worth quoting Simon's concrete example of a nearly decomposable system in full:

> Consider a building whose outside walls provide perfect thermal insulation from the environment. We shall take these walls as the boundary of our system. The building is divided into a large number of rooms, the walls between them being good, but not perfect, insulators. The walls between rooms are the boundaries of our major subsystems. Each room is divided by partitions into a number of cubicles, but the partitions are poor insulators. A thermometer hangs in each cubicle. Suppose that at the time of our first observation of the system there is a wide variation in temperature from cubicle to cubicle and from room to room – the various cubicles within the building are in a state of thermal disequilibrium. When we take new temperature readings several hours later, what shall we find? There will be very little variation in temperature among the cubicles of each room, but there may still be a large temperature variation *among* the rooms. When we take readings again several days later, we find an almost uniform temperature throughout the building; the temperature differences have virtually disappeared. (Simon 1996, p. 198)

To take an economic example, on a quarter-to-quarter or year-to-year basis, the level of employment in an economy is dominated by a complex interaction among relatively quickly adjusting factors, such as wage rates, interest rates, private and government expenditure. Yet, on a century-to-century basis, the level of employment is dominated by the slowly changing level of overall population, which probably has little to do with the detailed factors that govern employment in the short run, though it may well have to do with the also relatively slowly changing level of technology.

Similarly, the rates of return on various financial assets on a minute-to-minute or day-to-day basis are determined mainly by the rapid process of arbitrage in finan-

---

[19] See Simon and Ando (1961) and Simon and Iwasaki (1988) for a formal treatments.

cial markets, while the aggregate level of all interest rates on a year-to-year basis is determined by the more slowly developing profitability of industry in general.

Simon provides no criterion for distinguishing fast and slow adjustment. My two economic examples, in fact, suggest that it will often be a matter of degree or a matter of how the analysis is contextualized. In the case of simple dynamic processes, the speed of adjustment may be captured by parameters governing time dependence, such as ρ in Eqs. (10)–(13): $|\rho|$ closer to zero corresponds to slow adjustment; $|\rho|$ closer to 1 to fast adjustment. Yet, it is important to note that there is a *sharp* distinction between $|\rho| < 1$ and $|\rho| = 1$. The probability distributions governing process in which $|\rho| = 0$ and $|\rho| < 1$ differ in detail, but belong to a single family of stationary distributions. The probability distributions governing $|\rho| = 1$ belong to an entirely different family of distributions governing Brownian motion (Hendry 1995, ch. 3, section 10). The transition between the two families represents a kind of critical phenomenon familiar in certain physical applications (e.g., see Batterman 2001). There is a sense, then, in which the long-run causal structure emerges from the short-run structure at the point of transition between stationary (I(0), $|\rho| < 1$) processes and nonstationary (I(1), $|\rho| = 1$) processes. The short-run and long-run causal structures can have distinct ontologies and may, in fact, pull in opposite directions.

5.2 Cointegration and stochastic trends

To work with a concrete example, a typical *structural cointegrated vector autoregression* (*structural CVAR*) for three variables takes the form:

$$
\begin{aligned}
&\underbrace{\begin{bmatrix} 1 & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & 1 & \alpha_{yz} \\ \alpha_{zy} & \alpha_{zy} & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix}}_{\Delta\mathbf{X}_t} = \text{stationary dynamic component} \\
&+ \underbrace{\begin{bmatrix} \pi_{xx} & \pi_{xy} & \pi_{xz} \\ \pi_{yx} & \pi_{yy} & \pi_{yz} \\ \pi_{zx} & \pi_{zy} & \pi_{zz} \end{bmatrix}}_{\mathbf{\Pi}} \underbrace{\begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix}}_{\mathbf{X}_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon^x \\ \varepsilon^y \\ \varepsilon^z \end{bmatrix}}_{\mathbf{E}}
\end{aligned}
\tag{18}
$$

The left-hand side of (18) is essentially the same as the left-hand side of (8), except (trivially) the minus sides are omitted in the $\mathbf{A}$ matrix and (significantly) the dependent variables are expressed as first differences. If the variables in $\mathbf{X}$ are nonstationary (I(1)), then those in $\Delta\mathbf{X}_t$ are stationary (I(0)), so that all the issues of causal structure that were raised for Eq. (8) [or system (7)] remain pertinent. In particular, the need for causal sufficiency as a foundation for an apt representation of the causal structure and as a prerequisite to backing out the short-run component of the shocks is unchanged.

Our interest, however, is on the right-hand side, particularly on the term $\mathbf{\Pi X}_{t-1}$, which represents the long-run relationships among the data. If the variables in $\mathbf{X}$ are I(1) and, if they are all cointegrated, then they must share at least one common trend. Here, in fact, they must have either one or two common trends. If they have zero, then, contrary to our assumption, the variables would be stationary; if they have three, then,

again, contrary to assumption, the variables would not be cointegrated. We pointed out previously that for a single dynamic process, such as Eq. (10), $|\rho| = 1$ implies a stochastic trend. The analogue for a system of equations is that there is a stochastic trend for each eigenvalue of $\mathbf{\Pi}$ equal to one. In the vernacular of the time-series econometrician, such a system is said to have a *unit root.* It is also a mathematical fact that the rank of a matrix such as $\mathbf{\Pi}$ equals the number of variables minus the number of unit roots. Since $\mathbf{\Pi X}_{t-1}$ describes how the variables adjust to each other (i.e., the nature of their cointegrating relationships), this mathematical fact confirms the point made informally with respect to the random-walking drunk, his friend, and his brother that the number of cointegrating relationships and the number of stochastic trends add up to the number of variables.

It is possible statistically to separate the stationary from the nonstationary components of (18) (see Juselius 2006, ch. 7, Section 1). The nonstationary component can be written as

$$
\underbrace{\begin{bmatrix} \Delta \tilde{x}_t \\ \Delta \tilde{y}_t \\ \Delta \tilde{z}_t \end{bmatrix}}_{\Delta \tilde{\mathbf{X}}_t} = \underbrace{\begin{bmatrix} \pi_{xx} & \pi_{xy} & \pi_{xz} \\ \pi_{yx} & \pi_{yy} & \pi_{yz} \\ \pi_{zx} & \pi_{zy} & \pi_{zz} \end{bmatrix}}_{\mathbf{\Pi}} \underbrace{\begin{bmatrix} \tilde{x}_{t-1} \\ \tilde{y}_{t-1} \\ \tilde{z}_{t-1} \end{bmatrix}}_{\tilde{\mathbf{X}}_{t-1}} + \underbrace{\begin{bmatrix} \tilde{\varepsilon}^x \\ \tilde{\varepsilon}^y \\ \tilde{\varepsilon}^z \end{bmatrix}}_{\mathbf{E}}, \tag{19}
$$

where the tilde over variables indicates that the influence of the stationary component on the nonstationary has been removed statistically. Equation (19) reflects Simon's idea of near decomposability: "decomposable" because of the division between the stationary and the nonstationary, but only "near" because the *levels* of variables (long run) influence the *changes* in the same variables (short run). In our analogy to the drunk and his friend, Eq. (19) shows how the distance between the drunk and his friend influences the next step of the friend independently of any influences from the friend's own recent movements or from the change in the position of the drunk. For example, if the gap between them widens, the friend will have to pick up the pace just because the gap *is* wide and, not particularly because it has *become* wider after either his or the drunk's last step, even though the fact of the gap being wide is not unconnected with the whole history of their steps.

The trends are not themselves represented in (19). Let us assume that there is only one stochastic trend, then in a move that is analogous to the transformation of Eqs. (13) or (14) into system (12) we can represent the long-run component of the variables as

$$
\begin{bmatrix} \tilde{x}_t \\ \tilde{y}_t \\ \tilde{z}_t \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} \phi_x \\ \phi_y \\ \phi_z \end{bmatrix} \sum_{j=0}^{\infty} \tau_{t-j}, \tag{20}
$$

where the ϕ_i are constant parameters and τ is the shock to the trend, and the variables subscripted 0 are initial conditions. Thus, the trend is

$$
T_t = T_{t-n-1} + \sum_{j=0}^{n} \tau_{t-j}, \tag{21}
$$

where T_0 is an initial condition. By analogy with Eqs. (15) and (16), the trend in (21) may also be expressed as

$$T_t = T_{t-1} + \tau_t. \tag{22}$$

Equation (22) is also analogous to the dynamic process in system (12) with T in the role of u and τ in the role of ε.

It turns out, once we have estimates of the parameters (π_{ij}) and shocks (ε^j) in (19), that the values of T and the τ are uniquely determined. We can, therefore, always back estimates T and the τ out of equations such as (19) that take account only of observable variables. Without going into the details of the calculation, the approach is to find the sequence of values for τ_t that the render the variables cointegrated.

It is possible that one of the variables itself is, in fact, the trend. For example, if z were the trend ($z = T$), then in (19) $\pi_{zx} = \pi_{zx} = 0$ and $\pi_{zz} = 1$ (and in (20) $\phi_z = 1$). These are statistically testable hypotheses on the basis of an estimate of (19).[20]

It need not be the case, that the trend is observable. That case raises one of the main ontological issues: which representation, (19) or (20) and (22) together, is causally apt? One possibility is that (19) reflects the ontologically independent causal structure, so that the matrix **A** having a unit root is just a contingent fact about the parameters of the system (the elements π_{ij} of the **Π** matrix). In that case, the trend T and its shocks τ are just representational artifacts. On the other hand, it may be that the trend is ontologically independent, so that it should be regarded as an unobserved latent variable. In that case, the perfectly tuned values of the π_{ij} are just the phenomenal consequence of omitting the stochastic trend from the representation. The question of which representation is causally apt is precisely analogous to the question of whether system (12) [the analogue to (20) and (22) together] or Eq. (13) [the analogue to (19)] is causally apt.

The evidence that would favor the ontological independence of the stochastic trend T is of the same general type as favored the ontological independence of the common cause F in Sect. 3 above—namely, evidential variety and transferability. Start with two cointegrated variables. Adding a third variable to the system will not alter that fact: the third variable may be cointegrated with the original two or it may add a new stochastic trend. Depending on the exact details of the causal connections among the variables, the new variable may substantially alter the estimates of the parameters of the CVAR such as Eq. (18). A recent paper by Johansen and Juselius (2012) shows, while the precise values of estimates of the common stochastic trend (e.g., T) and the trend shock (τ) will be different, that asymptotically, as the number of observations becomes large, the estimates from the narrower and wider data set will converge. While we cannot rule out for any one set of data that the parameters are perfectly tuned to produce cointegration, it would quite generally require a different tuning to produce the same result in a wider data set. Thus, the stability of the estimates of the trend and its shock across increasingly wider data sets is strong evidence that the trend is, in fact, an ontologically independent, though not directly observed, variable.

---

[20] When the hypothesis is accepted, z is said to be *weakly exogenous* (see Juselius 2006, ch. 11, sections 1 and 2).

The ontology of nonstationary processes conveys an epistemological advantage not enjoyed by stationary processes. The relationship used to identify or measure latent trends is transitive cointegration rather than intransitive conditional probabilistic independence. An important consequence is that knowledge about trends can be built up in a *piecemeal* fashion—we do not have to begin with the notion that everything depends on everything else. Cointegration is a robust property that, once a minimally sufficient set of variables has been identified, does not depend on the details of the causal structure. This is not the case for conditional probabilistic independence, which depends fundamentally on the match between the underlying causal structure and the representation of that structure in the probability model. The consequence is that we can identify and measure latent stationary shocks only when the shocks together with the explicitly observed variables form a causally sufficient set and when we are fortunate enough that the patterns of causal dependence and independence in that set imply a unique causal structure. Our epistemic projects need not be so ontologically blessed. This is no counsel of despair; rather it is a recognition of real methodological challenges.

The advantages of nonstationarity should not be overstated. Although cointegration provides a smoother path to the identification and estimation of stochastic trends, it does not provide any particular advantages with respect to the determination of long-run causal structure. Having transformed trends into observable variables through indirect measurement, we can add them to the variables of a structural CVAR such as (18), but we will not have represented the long-term causal order of those variables until we can select the right set of restrictions on the Π matrix. To put it another way, the structural CVAR is only partly structural. The nonstationary elements remain a purely phenomenal description of the data.

The very features that were helpful in identifying and measuring trends—that cointegration is transitive and that cointegrating relationships are fundamentally deterministic—are unhelpful at uncovering asymmetrical causal relationships. To infer causal structure, we will typically require a set of methodological resources similar to those required in the nonstationary case: causal sufficiency and measures of conditional probabilistic dependence and independence. Indeed, there will a disadvantage in that the statistical methods for analyzing conditional probability are better developed and more familiar for the stationary case.

Both the advantages and disadvantages of nonstationarity arise from the emergence of the long-run structure in keeping with Simon's account of near decomposability. Simon treats that emergence as a case of aggregation. In his rooms/cubicle example, in the long run any one of the thermometers from any one of the cubicles represents the state of the rooms equally well—the detailed information cubicle information is redundant and only the common temperature of the room matters. Choosing one thermometer to represent the long-run state thus is an economical aggregation. Similarly, since all of the cointegrating variables index the same common stochastic trend, they are redundant, and it is possible for some purposes to economize on information. But in a state of long-run equilibrium, the thermometers are also uninformative about the internal structure of the rooms and the building, as they provide no basis for discrimination. Similarly, the commonness of the common stochastic trend is a marker of uninformativeness. The trends can provide useful information about the long-term

causal structure only if they are resituated to display their relationship to short-term changes in the variables.

6 On the interface of ontology and methodology

The discussion so far has been full of details, so it would perhaps be useful to end by drawing the main threads together. The structural vector autoregression is the dominant tool of empirical macroeconomics. It is sometimes taken to be a purely phenomenal description of the data. But economics is a policy science and it is, therefore, in the words of the title of James Woodward's (2003) book, devoted to "making things happen"—or at least to understanding how they happen. As a result economics needs models that can support counterfactual analysis.

In building such models, economists have relied on ontology that partitions economic processes into a deterministic structure and indeterministic shocks. Such a division does not in itself dictate a particular representation of the causal structure nor provide a particular identification or basis for measuring the not-directly-observable shocks. The debate between Rudebusch and Sims (Sect. 2.2) simply reminds us that the issue of how to get the right causal structure and to recover the shocks is not an abstract philosophical concern but one that animates practitioners and is even expressed ontologically: are the measured shocks the *real* shocks?

Within the economists' own ontological presuppositions, answering that question is a matter of ascertaining that the modeled causal structure is the correct one; then the shocks backed out from the model and the actual data will be correct. Cartwright, however, challenges those ontological presuppositions—particularly that the causal connections among variables can be regarded as deterministic. Her challenge has an epistemological grounding: one should not presume—as the strategy for backing out the shocks from a model does presume—that a variable must exist to screen off the probabilistic dependence between two causally unconnected variables unless one can point to the variable.

Presumably Cartwright's challenge is not meant to be a generic rejection of claims for the reality of unobserved entities. So, an effective answer would be one that gives us grounds for preferring the assumption of indirectly observed, ontologically independent entities. The criteria of variety of evidence (i.e., independent methods of indirect measurement that provide essentially the same values for the shocks) and transferability (i.e., a shock identified in one context appears to be causally efficacious in another) are frequently held to support claims of ontological independence. The methodological challenge is to support such claims in particular cases. As the Bayes-net analysts have taught us, such support may rely on difficult-to-establish claims of causal sufficiency. Difficult, however, is not impossible.

The fact that economic data are frequently nonstationary both poses its own methodological challenges and opens up the ontological possibility of a distinct hierarchy of causal structures. The cointegrated vector autoregression, which is a preferred technique among time-series econometricians for dealing with nonstationary data, offers an instance of Simon's near decomposability as it divides dynamic processes into distinct short-run and long-run causal structures. Stochastic trends, which are the

nonstationary counterpart to stationary shocks, turn out to be easier to identify and measure because of the ontology of cointegrating relations. Nonetheless, the challenge of the discovery of causal structure remains similar in the nonstationary case to the stationary—a point not much appreciated by economists.

References

Batterman, R. W. (2001). *The devil in the details: Asymptotic reasoning in explanation, reduction and emergence*. Oxford: Oxford University Press.

Boumans, M. (2005). *How economists model the world into numbers*. London: Routledge.

Cartwright, N. (1991). Replicability, reproducibility, and robustness: Comments. *History of Political Economy*, *23*(1), 143–155.

Cartwright, N. (1993). Marks and probabilities: Two ways to find causal structure. In F. Stadler (Ed.), *Scientific philosophy: Origins and development* (pp. 113–119). Dordrecht: Kluwer.

Cartwright, N. (2007). *Hunting causes and using them: Approaches in philosophy and economics*. Cambridge: Cambridge University Press.

Claveau, F. (2013). The independence condition in the variety-of-evidence thesis. *Philosophy of Science*, *80*(1), 94–118.

Cooper, G. F. (1999). An overview of the representation and discovery of causal relationships using Bayesian networks. In C. Glymour & G. F. Cooper (Eds.), *Computation, causation, and discovery* (pp. 3–64). Menlo Park, CA: American Association for Artificial Intelligence and Cambridge, MA: MIT Press.

Frisch, R. (1938). Statistical versus theoretical relations in economic macro-dynamics, reprinted in O. Bjerkholt (Ed.) In *Foundations of modern econometrics: The selected essays of Ragnar Frisch* (Vol. 1). Aldershot: EdwardElgar.

Duarte, P. G., & Hoover, K. D. (2012). Observing shocks. In H. Maas & M. Morgan (Eds). *Histories of observation in economics, supplement to history of political economy, 44*(5), 226–249.

Haavelmo, T. (1943). 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–115.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of the natural sciences*. Cambridge: Cambridge University Press.

Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century perspective. *Quarterly Journal of Economics*, *115*(1), 45–97.

Hendry, D. F. (1995). *Dynamic econometrics*. Oxford: Oxford University Press.

Hood, W. C. & T. Koopmans (Eds.). (1953). *Studies in econometric method, Cowles Commission monograph* (Vol. 14, pp. 1–26). New York: Wiley.

Hoover, K. D. (1995). Comments on Cartwright and Woodward: Causation, estimation, and statistics. In D. Little (Ed.), *On the reliability of economic models: Essays in the philosophy of economics*. Dordrecht: Kluwer.

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

Hoover, K. D. (2003). Nonstationary time series, cointegration, and the principle of the common cause. *British Journal for the Philosophy of Science*, *54*(4), 527–551.

Hoover, K. D. (2004). Lost causes. *Journal of the History of Economic Thought*, *26*(2), 149–164.

Hoover, K. D. (2007) Probability and structure in econometric models, unpublished working paper, Duke University [early version of Hoover (2009)], downloadable from: http://public.econ.duke.edu/~kdh9/Source%20Materials/Research/Probability%20and%20Structure%2019%20September%202007.pdf.

Hoover, K. D. (2009). Probability and structure in econometric models. In *Logic methodology and philosophy of science: Proceedings of the thirteenth international congress* (pp. 497–513). London: College Publications.

Hoover, K. D. (2011). Counterfactuals and causal structure. In P. M. K. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 338–360). Oxford: Oxford University Press.

Hoover, K. D. (2013). The role of hypothesis testing in the molding of econometric models. *Erasmus Journal for the Philosophy of Economics*, *6*(2), 42–65.

Hoover, K. D., & Jordà, Ò. (2001). Measuring systematic monetary policy. *Federal Reserve Bank of St. Louis Review*, *83*(4), 113–137.

Hoover, K. D., & Perez, S. J. (2000). Three attitudes towards data-mining. *Journal of Economic Methodology*, *7*(2), 195–210.

Hoover, K., Johansen, S., & Juselius, K. (2008). Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression. *American Economic Review, 98*(2), 251–255.

Humphreys, P. (1989). *The chances of explanation: Causal explanations in the social, medical, and physical sciences*. Princeton: Princeton University Press.

Johansen, S., & Juselius, K. (2012). *An asymptotic invariance property of the common trends under linear transformation of the data*. Department of Economics, University of Copenhagen, Discussion Paper No. 10-30

Juselius, K. (2006). *The cointegrated VAR: Methodology and Applications*. Oxford: Oxford University Press.

Koopmans, T. C. (Ed.). (1950). *Statistical inference in dynamic economic models, Cowles Commission monograph 10*. New York: Wiley.

Mayo, D. & Spanos, A. (2013). Ontology and methodology in statistical modeling. In *Paper presented at the conference ontology and methodology*, 4–5 May 2013 at Virginia Tech.

Murray, M. P. (1994). A drunk and her dog: An illustration of cointegration and error correction. *The American Statistician*, *48*(1), 37–39.

Papineau, D. (1991). Correlations and causes. *British Journal for the Philosophy of Science*, *42*(3), 397–412.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Perrin, J. (1913[1990]). *Atoms*. Woodbridge, CT: Ox Bow Press.

Qin, D., & Gilbert, C. (2001). The error term in the history of time series econometrics. *Econometric Theory*, *17*(2), 424–450.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Rudebusch, G. D. (1998a). Do measures of monetary policy in a VAR make sense? *International Economic Review*, *39*(4), 907–931.

Rudebusch, G. D. (1998b). Do measures of monetary policy in a VAR make sense? A reply to Christopher A. Sims. *International Economic Review*, *39*(4), 943–948.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.

Simon, H. A., & Ando, A. (1961). Aggregation of variables in dynamic systems. *Econometrica*, *29*(2), 111–138.

Simon, H. A., & Iwasaki, Y. (1988). Causal ordering, comparative statics, and near decomposability. *Journal of Econometrics*, *39*(1/2), 149–173.

Sims, C. A. (1998). Comment on Glenn Rudebusch's 'Do measures of monetary policy in a VAR make sense? *International Economic Review*, *39*(4), 933–941.

Sims, C. A. (1999). The role of interest rate policy in the generation and propagation of business cycles: What has changed since the '30s? In Fuhrer, J. C. & Schuh, S. (Ed.) *Beyond shocks: What causes business cycles. Federal Reserve Bank of Boston Conference Series, No. 42* (pp. 121–60). Boston: Federal Reserve Bank of Boston.

Sober, E. (2001). Venetian sea levels, British bread prices, and the principle of the common cause. *British Journal for the Philosophy of Science*, *52*(2), 331–46.

Spanos, A. (1989). On rereading Haavelmo: A retrospective view of econometric modeling. *Econometric Theory*, *5*(3), 405–429.

Spanos, A. (1990). The simultaneous equations model revisited: Statistical adequacy and identification. *Journal of Econometrics*, *44*(1), 87–108.

Spanos, A. (1995). On theory testing in econometrics: Modeling with nonexperimental data. *Journal of Econometrics*, *67*(2), 189–226.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT Press.

van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.

Walras, L. (1874[1954]). *Elements of pure economics, translated by William Jaffe*. Homewood, IL: Richard Irwin.

Woodward, J. B. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*(1), 1–63.