Economics & Philosophy (2021), **37**, 441–461
doi:10.1017/S0266267120000437

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Causal effects and counterfactual conditionals: contrasting Rubin, Lewis and Pearl

Keith A. Markus

Department of Psychology, John Jay College of Criminal Justice, CUNY, 524 West 59th Street, New York, NY 10019, USA
Email: kmarkus@jjay.cuny.edu

(Received 24 September 2017; revised 29 July 2020; accepted 12 November 2020; first published online 11 February 2021)

Abstract
Rubin and Pearl offered approaches to causal effect estimation and Lewis and Pearl offered theories of counterfactual conditionals. Arguments offered by Pearl and his collaborators support a weak form of equivalence such that notation from the rival theory can be re-purposed to express Pearl's theory in a way that is equivalent to Pearl's theory expressed in its native notation. Nonetheless, the many fundamental differences between the theories rule out any stronger form of equivalence. A renewed emphasis on comparative research can help to guide applications, further develop each theory, and better understand their relative strengths and weaknesses.

Keywords: Causation; counterfactual conditionals; comparative research

> [T]hese three formalisms of counterfactuals [of Lewis, Rubin and Pearl] . . . are shown to be logically equivalent; a problem solved in one framework would yield the same solution in another. (Pearl *et al.* 2016: 126)

1. Introduction

In this article, I seek to encourage comparative research by exploring differences between both Rubin's Causal Model (RCM) and Lewis' Theory of Counterfactual Conditionals (LTC) on the one hand, and Pearl's Structural Causal Model (SCM) on the other. Rubin (1974) developed a theory of estimation of causal effect sizes based on potential outcomes (alternatively, potential responses), the values of an outcome variable that individuals would have taken on if exposed to a different treatment, taken as ontological primitives. He then worked out methods for estimating treatment effects in a population based on observations of different individuals exposed to different treatments. Pearl (2009*a*) defined counterfactual quantities, which may occur in counterfactual conditionals, as derived quantities in terms of the operations on causal models. He then developed methods for estimating causal effects based on an assumed qualitative causal structure.

Lewis analysed counterfactual conditionals in terms of comparative possibility (1973*b*, 1973*c*; broadly characterized as a similarity theory). Roughly, the counterfactual conditional *If A then B* holds true if and only if *B* holds true in all the most similar possible worlds in which *A* holds true. Lewis analysed causation in terms of counterfactual conditionals (1973*a*, 2004). *A* causally influences *B* if counterfactual dependence holds between the right sorts of events based on the right sorts of counterfactual conditionals. Causation is the ancestral of causal influence. Pearl (2009*a*) analysed a subset of counterfactual conditionals in terms of manipulations of causal graphs or systems of causal equations. Roughly, *If* $A = a$ *then* $B = b$ holds true if suspending the modelled causes of *A* and fixing *A* to *a* in the model results in the equation for *B* yielding the value *b*. SCM represents this as $A_{\mathrm{Do}(B=b)}$. One can interpret SCM as offering a non-reductive theory of causation or as taking type-level causation as an unanalysed primitive. (The online supplementary materials to this article contain a brief introduction to RCM, LTC and SCM.)

These broad differences may appear to take equivalence off the table but equivalence claims occur and hold the potential to mislead researchers.[1] These statements trace back to Galles and Pearl who concluded 'for recursive models, the causal model framework does not add any restrictions to counterfactual statements beyond those imposed by Lewis's framework' (1998: 169). Galles and Pearl (1998: 174) further asserted that because one can characterize certain counterfactual conditionals using recursive structural causal models, 'This establishes, in essence, the formal equivalence of structural equation modeling [meaning SCM] … and the potential-response framework'.[2] Rubin's, Lewis' and

[1]It is illuminating to compare Andrew Gelman's characterization of Morgan and Winship's (2007) book to that of Pearl. Gelman: 'The book is unusual in overlaying several different statistical approaches to causal inference …'. Pearl: 'The book is unique in recognizing the equivalence between the counterfactual and graphical approaches to causal analysis …', where 'counterfactual approach' plausibly refers to RCM and 'graphical approach' to SCM (both quotations from the back cover).

[2]'[W]e can safely conclude that, in adopting the causal interpretation of counteractuals … we are not introducing any restriction on the set of counterfactual statements that are valid relative to recursive systems' (Pearl 2009*a*: 242). The 'formal equivalence between the structural and potential-outcome frameworks covers issues of semantics and expressiveness' (Pearl 2009*a*: 244). '[T]he two frameworks [RCM and SCM] have been proven to be logically equivalent, differing only in the language in which researchers are permitted to express assumptions' (Pearl 2012: 79). 'A systematic analysis of the syntax and semantics of the two notational systems reveals that they are logically equivalent (Galles and Pearl 1998; Halpern 1998); a theorem in one is a theorem in the other, and an assumption in one has a parallel interpretation in the other' (Bollen and Pearl 2014: 314, referring to potential outcomes and structural causal models which the authors refer to as structural equation models; The chapter cited as Halpern 1998 was republished as Halpern 2000.) 'Rubin's framework, known as "potential outcomes," differs from the structural account only in the language in which problems are defined' (Pearl *et al.* 2016: 126). 'We evaluate expressions like "had X been x" in the same way that we handled interventions do(X = x), … In this respect, structural counterfactuals are compatible with Lewis's idea of the most similar possible world' (Pearl and Mackenzie 2018: 268; Pearl and Mackenzie here conflate Stalnaker's theory with Lewis' theory). Pearl and Mackenzie (2018: 260) present 'potential outcomes' and 'the Neyman–Rubin Causal Model' (i.e. RCM) as synonyms. They then present SCM and RCM as addressing the same counterfactual conditionals and causal inferences and present SCM as putting

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

Table 1. Different topics addressed by different authors

Topic	Author: Lewis	Author: Rubin	Author: Pearl
Theory of counterfactual conditionals	Addressed*		Addressed*
Theory of causation	Addressed		(Unclear)
Methodology for estimating causal effects		Addressed*	Addressed*

*Equivalence claim between elements in the same row.

Pearl's work present three distinct bundles of theory (Table 1). Rubin offered a theory of causal estimation but not of counterfactual conditionals or of causation. So, I consider only a posited equivalence between Rubin's and Pearl's theories of causal estimation. Likewise, Lewis (1973*a*) analysed causation in terms of counterfactual conditionals whereas Pearl (2018; Pearl and Mackenzie 2018) contrasts his account with reductive theories of causation. I therefore consider only posited equivalence between Pearl's and Lewis' theory of counterfactual conditionals to the exclusion of their theories of causation.

As an analogy, suppose that a theory of brotherhood states that x is a brother to y iff x is a male sibling of y and a theory of sisterhood states that x is a sister of y iff x is a female sibling of y. I shall designate such theories as sharing *weak equivalence* in the sense that one can re-interpret 'brother' to mean sister and 'male' to mean female to state the sisterhood theory using the notation of the brotherhood theory. However, no-one would take this to mean that the brothers and sisters are the same thing. I shall reserve *strong equivalence* to describe two theories that say the same thing in different ways, such as the brotherhood theory expressed in English and expressed in French. The positive arguments offered for the equivalence of RCM to SCM and of LTC to SCM (Galles and Pearl 1998; Pearl 2009*a*) follow the same form. They express SCM in the notation of the other theory, and then show that SCM in borrowed notation is equivalent to SCM in native notation.[3] As such, they demonstrate at most weak equivalence but not strong equivalence. This analysis of the arguments highlights an important methodological point. Both arguments rely on SCM to represent its relation to other theories. This strategy builds the limitations of SCM into the resulting comparison and likewise filters out aspects of the rival theory that do not readily translate to SCM. Successful analysis of the relationships between rival theories instead requires an external vantage point. Expressibility in one of two rival theories under comparison should never serve as either an admissibility criterion or intelligibility criterion for assertions in a discussion comparing them. Why is this important? Presenting SCM as a successor to RCM and LTC invites an interpretation claiming equivalence in the sense that SCM supersedes RCM and LTC (compare Pearl 2009*a*: Ch. 7 and Pearl and Mackenzie 2018: Ch. 8). This supersession narrative refers to the substantive theories themselves and not merely to the formal language or notation in which they are expressed. A mere

structural causal models to the same use as RCM's 'abstract mathematical objects that are managed by algebraic machinery but not derived from a model' (2018: 280).

[3]The arguments mistakenly present SCM expressed in the notation of the rival theory as the rival theory itself, leading to the illicit conclusion of equivalence.

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

equivalence of languages would not suggest the obsolescence of theories expressed using them if the theories differ in content. Equivalence claims warrant attention because of their implications for comparative research.

Upon presenting a theory, it is customary to consider some form of comparative analysis with respect to the theory's rivals. This can involve comparing the relative scope of application of the theories. In the case of linguistic or semantic theories it can involve a comparison of fitting established usage and handling difficult cases. Increasingly, philosophers are citing empirical results to supplement linguistic intuitions. In the case of theories regarding scientific methodology, it can involve the relative performance of research procedures recommended by each theory or the comparative scope of application of two methods. Yet, over two decades and many publications, such systematic comparative analyses have not emerged in the literature on SCM in relation to RCM or LTC. Hausman (1998) and Woodward (2003) clearly distinguished LTC from their contributions to the metaphysics of causation underlying SCM and the related work of Spirtes *et al.* (2000) and Briggs (2012: 140) argued that 'the divergence between the two approaches is larger than commonly recognized'. Imbens (2019) gently addressed Pearl's criticisms of RCM. However, on the whole, methodological literature tends to either treat them as interchangeable or focus on one to the exclusion of the other. The equivalence claims arguably function rhetorically to quell such expectations for comparative research. However, justifying this lack of attention would require strong equivalence, not just weak equivalence.

It is worth stepping back to note that even if these two rival theories were strongly equivalent to SCM, the need for comparative research would remain in place. LTC differs from Stalnaker's (1968) theory and Bennett's (2003) theory, among others. Likewise, a theory of effect estimation rests on a theory of causation and a myriad of approaches to analysing causation differ from Pearl's non-reductive approach and from the manipulation theory suggested by RCM (Schmaltz 2014). Hoover's (2001) emphasis on dispositions differs fundamentally from the Humeanism implicit in both RCM and SCM. Indeed, Lewis' (1973*a*) theory of causation, as distinct from LTC, would still differ from Pearl's notion of causation even if they shared the same theory of counterfactual conditionals. Moreover, modifications of RCM have rendered Steyer's (2005) approach to effect estimation sufficiently different from RCM that the two are not equivalent. It is not possible for SCM to simultaneously retain equivalence with multiple non-equivalent alternatives. As such, even if the equivalence claims held for RCM and LTC, they still would not warrant a moratorium on comparative research. However, the next two sections will show that RCM and LTC do differ from SCM.

2. SCM and RCM on causal effect estimation

Much of what has been written about the relationship between RCM and SCM from the SCM perspective is ambiguous because it conflates RCM with potential outcomes notation. This distinction is crucial to the distinction between strong and weak equivalence. Showing that potential outcomes notation can express both theories does not suffice to show strong equivalence between the theories.

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

2.1. Strong equivalence between RCM and SCM

The following expression lies at the centre of the relationship between RCM and SCM (Pearl 2012: 75; cf. Pearl 2009*a*: eqn 3.51): $Y_x(u) \triangleq Y_{Mx}(u)$.[4] The left-hand side represents in Holland's (1986) notation the response by participant u measured on variable Y under treatment condition x. The right-hand side represents the value of $Y \mid Do(X = x)$ relative to model M. The interpretation of u in SCM varies between an equivalence set of participants having the same values on all disturbance terms, or simply one such respondent. Although the former appears to be the strict interpretation under SCM, the latter is more charitable toward claims of equivalence. Even allowing some flexibility regarding u, however, the lack of equivalence is clearly evident in the presence of M on the SCM side but not the RCM side. The RCM interpretation of the left-hand side refers to the world being studied and the SCM interpretation of the right-hand side refers to a model rather than the world being modelled. This reflects the fact that potential outcomes denote different things in RCM than they do in SCM (Rubin 1974; Pearl 2009*a*). Neither RCM nor SCM internally distinguishes these two kinds of potential outcomes. As such, either the equality in the above expression fails or the left-hand side fails to correspond to potential outcomes in RCM. Either way, RCM and SCM cannot bear strong equivalence because shared referring terms refer to different things in each.

One argument that might be offered to support equivalence is that because SCM has been shown to be complete, it excludes any other interpretation of causal discourse (Pearl 2009*a*). SCM is only complete with respect to the limited domain of discourse that it seeks to characterize and this completeness does not exclude other forms of causal discourse that it fails to characterize (Markus 2011). The scope of the semantic completeness of SCM is limited by its representational incompleteness. The remainder of this section unpacks various aspects of this non-equivalence, involving model semantics, causal interpretation and research practice.

2.1.1. Model semantics

Both model misspecification and implicit accessibility relations distinguish RCM from SCM.

2.1.1.1. Model misspecification. Pearl's tendency to assume a correctly specified model in his exposition obscures an important corollary of the different ways of understanding potential outcomes. Pearl's expression for the consistency rule, $X(u) = x \supset Y(u) = Y_x(u)$, describes a model. Potential outcomes in RCM describe the reality being studied, these determine the causal effect one wishes to estimate. Suppose for a moment that the model incorrectly represents reality. Let $Y_x(u)$ represent the actual potential outcome and let $Y^*{}_x(u)$ represent the counterfactual quantity predicted by the model. If $X = x$, then the consistency

[4] For consistency throughout, $\sim$, $\wedge$, $\vee$ and $\supset$ denote negation, conjunction, disjunction and the material conditional. $\triangleq$ denotes equality by definition and $\Box\!\!\rightarrow$ denotes Lewis' conditional. ω denotes a possible world.

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

rule asserts that $Y(u) = Y^{*}_{x}(u)$ whereas RCM asserts that $Y(u) = Y_x(u)$. By hypothesis, let the inaccuracy in the model yield $Y^{*}_{x}(u) \neq Y_x(u)$. RCM and SCM then produce distinct claims. Significantly, one needs to extend the notation of SCM to express this difference.

2.1.1.2. Implicit accessibility relations. The contrast between the characterization of treatments in each approach illustrates ways that SCM can include possibilities excluded by RCM. Conversely, the contrast between the characterization of stochastic independence assumptions illustrates the opposite. Consider an example of a compound intervention involving X and Z. The treatment condition includes $X = 1$ and $Z = 1$ whereas the control condition includes $X = 0$ and $Z = 0$. Within RCM, the clear description of the treatments plays a fundamental role in determining the range of possibilities considered (Imbens and Rubin 2015). Drawing on the language of modal logic, one can frame this as the scope of possible worlds under consideration (i.e. deemed accessible). One can represent the compound intervention with the following algebraic model.

$$Y = b1(X) + b2(Z) + U_Y$$
$$X = U_X$$
$$Z = X$$
$$\mathrm{Cov}(U_X, U_Y) = 0$$

For illustration, and without loss of generality, suppose that the actual world is an $X = 0 \wedge Z = 0$ world. Within RCM, the space of possible worlds partitions into two types: $X = 0 \wedge Z = 0$ worlds and $X = 1 \wedge Z = 1$ worlds. Worlds with mismatched values for X and Z fall outside the scope of possible worlds (or are inaccessible worlds) because they correspond neither to the treatment or the control (cf. Rubin 1974: 689). However, SCM assumes modularity as part of the definition of manipulations as modifications of structural causal models. Thus within SCM, the worlds in which X and Z differ remain accessible by assumption. In the context of a specific study, 'Necessarily $X = Z$' holds true from the perspective of RCM but not SCM (cf. Woodward 2016).

If we allow collapsing two variables into one, we lose with one hand what we gain with the other. With two variables, SCM can express the non-identity (distinctness) of the two variables but not their necessary numeric equality. If we allow such replacements and reduce the SCM representation to one variable, then SCM can express necessary numeric equality (in a sense) but cannot express non-identity. As such, we arrive at an assertion that SCM cannot express by forming the conjunction of these two claims: X and Z are non-identical but necessarily numerically equal. This is precisely the kind of theoretical claim that practicing researchers might like to make. Indeed, the Stable Unit Treatment Value Assumption (Rubin 1980) can be interpreted as entailing precisely this sort of claim: If the treatment involves taking aspirin and not taking a placebo researchers plausibly wish to assert that in the experiment participants either took aspirin but not the placebo or took the placebo but not aspirin without being committed to the implausible claim that taking aspirin is identical to not taking the placebo. This is because outside the context of the study, plenty of people take neither. This difference reflects the difference between a narrow representation of treatment in a study in RCM and the broader aim to represent causal generalizations in SCM (see section 2.1.3.1).

Conversely, consider the assumption that $\mathrm{Cov}(U_X, U_Y) = 0$. SCM takes such assumptions as fixed givens. If sufficient independences hold, they provide the grounds for causal inference within SCM. However, as a more thorough Bayesian, Rubin wants to incorporate uncertainty about such assumptions into the analysis. A Bayesian might express this uncertainty in the form of a prior distribution for the covariance that might centre around zero but still show some dispersion around it. Thus, in SCM the modal assertion 'necessarily $\mathrm{Cov}(U_X, U_Y) = 0$' holds true whereas in RCM it does not and instead its negation, 'possibly $\mathrm{Cov}(U_X, U_Y) \neq 0$' holds true. To be clear, nothing prevents SCM from expressing $\mathrm{Cov}(U_X, U_Y) \neq 0$ as a different model.[5] Rather, the point is that RCM and SCM commit themselves to different modal claims in accepting the above model. That is, the same model is interpreted differently within RCM and SCM.

Consider two possible worlds other than the actual world. In ω1, $X = 0$, $Z = 1$ and $\mathrm{Cov}(U_X, U_Y) = 0$. In ω2, $X = 1$, $Z = 1$ and $\mathrm{Cov}(U_X, U_Y) \neq 0$. In SCM, ω1 but not ω2 remains accessible from the actual world, and indeed ω2 is accessible from no worlds accessible to the actual world. Conversely, in RCM, ω2 but not ω1 remains accessible from the actual world, and ω1 is accessible from no worlds accessible to the actual world. As such, the two interpretations of the model yield different modal assertions. Because the range of possibilities varies between RCM and SCM, assertions that hold true in one may hold false in the other. Jointly, these differences demonstrate that neither RCM nor SCM constitutes a special case of the other.

2.1.2. Causal interpretation

Treatment interventions, manipulation and casual effects further distinguish RCM from SCM.

2.1.2.1. Treatment interventions. RCM and SCM differ in a number of inter-related ways involving the conceptualization and representation of treatment interventions. RCM does not assume that the researcher has full knowledge of what may be a plethora of antecedent variables that may moderate the treatment effect. Inference is thus limited to the treatment under study in combination with the pre-treatment histories of the study participants (Rubin 1978). In contrast, SCM assumes prior knowledge of at least the qualitative causal structure of all relevant variables, typically represented as being relatively few in number, exemplifying relatively simple and isolated causal systems. Whereas RCM treats generalization of causal effects as a complex empirical undertaking, SCM reduces it to a

[5]One could interpret the inclusion of a parameter for this covariance with a parameter space that includes zero as an expression of the modal claim that they possibly co-vary. This interpretation, however, would run afoul of a convention in structural equation modelling of not including a parameter unless one hypothesizes a non-zero value.

relatively straightforward logical exercise based on this assumed qualitative causal knowledge (Pearl and Bareinboim 2014).

Correspondingly, RCM emphasizes detailed description of real-world interventions whereas SCM emphasizes interventions as simplified abstractions in a formal calculus. For example, Rubin (1978) placed comparable value on non-reactive double-blind treatments and non-double-blind reactive treatments, cautioning researchers not to assume that they have the same effects. In contrast, Pearl (2009*a*) considered effects of treatments on other variables, but generally assumes that treatments are non-reactive in the sense that the intervention itself only affects the manipulated treatment variable. This non-reactivity is implicit in the way that SCM separates interventions from the model by treating them as operations on the model itself. However, if one were to add an intervention variable to the model, reactivity would violate the consistency rule. For example, consider economic behaviour that differs depending on whether it is freely chosen or imposed by regulatory intervention. RCM encourages the study of such phenomena whereas SCM excludes them, and is thus less general in this respect.

2.1.2.2. Causation and manipulation. A related difference involves the incorporation into RCM of the slogan 'No causation without manipulation' (Holland 1986) but rejected by SCM (Bollen and Pearl 2014; Pearl 2018, 2019). Perhaps under the influence of the equivalence claims, Bollen and Pearl assumed that the slogan must mean the same thing in the context of RCM as in SCM and on that basis they asserted that Rubin and Holland were wrong to make this assertion. Pearl (2018) characterized manipulation theories as intolerant, which is only intelligible if one assumes that they exclude other notions of causation. It seems much more plausible to instead take this difference as further evidence against equivalence and allow that the slogan can hold true in RCM and false in SCM without contradiction because it means different things in each context. In this case, it suggests that RCM and SCM assume different forms of causation. Holland's slogan fits comfortably with traditional manipulation theories which make manipulation essential to causation whereas interventionist theories typically differentiate themselves from manipulation theories on precisely this point (e.g. Woodward 2003).

2.1.2.3. Causal effects. One prominent difference between RCM and SCM involves how each defines causal effects. For RCM, a causal effect involves a difference in expected value of the potential outcome between two or more treatments (Rubin 1974).[6] In contrast, SCM defines causal effects as differences in probability distributions. These two characterizations come apart when we consider two probability distributions that differ in some way other than expected value. One example occurs where one treatment results in greater variation in outcomes than another, even though the expected value remains the same across treatments. In such cases, adopting RCM would lead to the conclusion of no causal effect where SCM would lead to the opposite conclusion even though the facts of the matter are not in dispute. The difference in conclusions simply

[6]Rosenbaum and Rubin (1983) also considered a ratio of potential outcomes, but the formulation as a difference has become the standard means of exposition.

reflects non-equivalent criteria used to define the estimand. In this respect SCM is more general than RCM.

2.1.3. Research practice

Finally, modelling method versus theory, research design practice, and m-bias distinguish RCM from SCM with respect to research practice.

2.1.3.1. Modelling method versus theory. Many of the above differences relate to a basic difference in representation. SCM seeks to encapsulate general scientific knowledge represented in multi-purpose causal models and use them to guide estimation of various causal effects included in the model. In contrast, RCM instead emphasizes the representation of specific events in the context of a specific study.[7] Whereas SCM might encourage researchers to re-use the same model in different studies estimating different effects, RCM encourages study specific models to guide the methodological choices specific to each study. Relatedly, RCM seeks to separate the model of treatment from the model of the outcome so that methodological decisions do not depend upon knowledge about outcomes. In contrast, SCM combines treatment variables and outcome variables into one model considered together. Thus, when Pearl criticized RCM for failing to express ignorability assumptions in the language of the causal model (e.g. 2009*a*, 2012), he framed as a failing something that RCM claims as a central virtue.

In part, the differences between the two approaches to effect estimation may reflect focus on different types of causal systems. SCM tends to focus on examples of stable causal systems whereas RCM tends to focus on examples of systems that develop over time. The canonical example for SCM is a sprinkler set-up in which the effects of various variables in the system remain constant over time, depending at most on the states of other variables in the system. In contrast, the canonical example for RCM involves taking aspirin to alleviate a headache. There is a limited window over the course of the headache in which the causal effect is present: taking aspirin will not prevent the headache and eventually the headache will go away on its own without aspirin. Expositions of causal inference using SCM tend to focus on timeless models that can be applied to any intervention on any of the variables. Rubin (1978) emphasized the times of treatments and of effects, denying that the same causal model can be reapplied at a later time to allow the same individual to receive contrasting treatments. This is why the Fundamental Problem of Causal Inference, that one cannot observe more than one potential outcome for the same case, plays a more prominent role in RCM than in SCM. Ignoring development over time in a canonical SCM model could lead to the violation of various axioms such as the definition of a model in terms of fixed functional relations, the Causal Markov Condition, Faithfulness, Consistency, or Effectiveness. For example, developmental processes could create co-variances between variables not represented in the synchronic causal structure. Likewise, spontaneous changes in variables due to developmental processes can be taken as violating Effectiveness, $X_{xw}(u) = x$,

[7]Hiddleston (2005) drew a similar contrast between *event variables* and *property variables* such that a given individual can take on a given property more than once, but a given event constituted by an individual taking on a property can only occur once.

because a manipulated variable will not necessarily maintain its manipulated value. One could express a longitudinal time-specific model in SCM-like notation. However, so doing would take one away from the standard SCM methodology for effect estimation and in the direction of RCM. Expressing one methodology in the notation of the other does not alter the substantive differences between them.

2.1.3.2. Research design and scientific practice. Another cluster of differences between RCM and SCM involves the conceptualization of the variables used to formulate potential outcomes. Whereas RCM focuses on designing studies with appropriate variables to answer a fixed question, SCM focuses on finding questions that can be answered with a fixed set of variables. As such, Pearl tends to take a set of variables serving as nodes in a causal graph as absolute givens (contrary to Woodward 2016) whereas Rubin advises the researcher to consider modifying these to optimize the extent to which the study meets assumptions required for causal inference. For example, Imbens and Rubin (2015) suggest modifying the granularity of the outcome variable (Y) to maximize the plausibility of assumptions required for causal inference. This difference can be contextualized as an element of scientific realism in RCM that does not translate well to SCM: Rubin clearly views the functional relations between units, treatments and potential outcomes as specific to the particular measures chosen while recognizing that different measures can support inferences about the same underlying entities. In contrast, SCM tends to accept variables in a specific form as the fundamental objects of causal description. In this respect, different RCM studies can focus on different RCM potential outcomes to answer questions about the same topic whereas for SCM it is more natural to consider studies using different variables as answering questions about different topics.[8]

Finally, standard presentations frame RCM in terms of a process of collective cognition bringing together disputants regarding some issue (e.g. health risks associated with a product such as cigarettes) and providing a process by which they can agree on a design to settle an issue before knowing the result. In contrast, standard presentations frame SCM as representing the cognitive processes of an individual researcher making causal inferences from data in isolation from others. In this respect, SCM comes much closer to classical evaluation practices whereas RCM comes much closer to the emphasis on stakeholder buy-in and utilization found in contemporary evaluation practices (Shadish *et al.* 1991).

2.1.3.3. M-bias. The controversy surrounding M-bias sheds additional light on differences between RCM and SCM. Shrier (2008) suggested that M-bias offered a counter-example to Rubin's advocacy of conditionalization when estimating causal effects. The M-bias example involves a treatment (W) solely a function of a disturbance (U_1), an outcome (Y) solely a function of W and a disturbance (U_2), and a single covariate (X), a function of U_1 and U_2. From the SCM perspective, U_1 and U_2 exhibit unconditional independence and thus there exists

[8] I do not mean to suggest a broader contrast of realism in RCM with antirealism in SCM. Prima facie, there are ample influences of each in both and working out the full mosaic of such influences in each would constitute a complex and subtle undertaking.

no bias in the unconditional estimate of the effect of W on Y. However, conditioning on X introduces stochastic dependence between U_1 and U_2. So, the conditional estimate contains bias (see also Pearl 2009*b*; Sjolander 2009). Rubin (2009) responded that attempting to estimate the causal effect this way relied on highly improbable happenstance (see also Imbens 2019). Pearl (2009*c*: 3) responded that '*M* -bias is not a phenomenon that depends on finely-tuned, "hoped-for compensating imbalances" as caricatured by Rubin but, rather, a structural property, persisting no matter what parameters are assigned to the various associations in the model'.

Two of the differences considered above shed light on the exchange. When Pearl spoke of independence persisting no matter the values of the parameters in the model, he described the scope of possible worlds defined by the parameter space of the model. Because Cov(U_1, U_2) is not included in the model as a parameter, it necessarily has the value of zero for SCM. However, from the RCM perspective, a researcher adopting a scientific approach should exercise caution by acknowledging some uncertainty about the assumption of stochastic independence between U_1 and U_2. Whereas the SCM strategy uses assumptions such as that about U_1 and U_2 to minimize conditionalization, the RCM strategy does just the opposite, relying on conditionalization to ease the reliance on strong assumptions.

Similarly, the example illustrates SCM's estimation strategy of seeking questions that can be answered given a fixed set of variables. From the RCM perspective, however, the example presents a poor choice of variables with which to answer the stated question. X offers a poor choice of co-variate because it has no causal effect on W but itself causally results from variables that do cause W. Good research design, from the RCM perspective, would not simply accept a set of variables as given but would rather design a study to include variables that determine W (Rubin 2007). Each author appears to have correctly understood M-bias within the context of his own assumptions, but the example draws into relief the differences between the assumptions of each author.

2.2. Weak equivalence between RCM and SCM

The conclusion that SCM can be expressed in potential outcomes notation in a way that renders it equivalent to SCM expressed in its native notation seems unexceptional. The fact that RCM and SCM can be expressed in potential outcomes notation rests on using the notation to express different things in each case. This underscores the importance of not conflating potential outcomes notation with a particular theory expressed in that notation such as under the ambiguous rubric of a potential outcomes framework.

2.3. Section summary

This section has enumerated differences between RCM and SCM involving semantics, causation and research practice (Table 2). It seems very plausible that such different methodologies have different strengths and weaknesses and different applications. Greater focus on areas of divergence in comparative research can help bring these to light and help to refine methodological advice to researchers.

Table 2. Contrasting two methodologies for effect estimation

RCM	SCM
$Y_x(u)$ as a model-independent feature of the world	$Y_x(u)$ as a model-dependent feature of a model
Effects attributed to combination of pre-intervention treatment history and treatments	Effects attributed to treatments only
Emphasis on thick description of real-world interventions	Emphasis on abstract, simplified interventions as formal calculus
No causation without (worldly) manipulation	Causation without (worldly) manipulation
Emphasis on representation of specific events in the context of a specific study	Emphasis on representation of generalized scientific knowledge
Emphasis on separating treatment model from outcome model	Emphasis on combining treatment model with outcome model
Emphasis on developing systems	Emphasis on stable systems
Emphasis on seeking best data to answer a fixed question	Emphasis on seeking questions that can be answered with fixed data
Emphasis on modifying variables to optimize assumptions	Emphasis on taking a set of variables as fixed givens
Emphasis on designing data to support desired analysis	Emphasis on designing analysis of predetermined data
Emphasis on social cognition and social processes in applied research	Emphasis on inference by an isolated individual researcher

3. SCM and LTC on counterfactual conditionals

I will first focus on strong equivalence with LTC and then with the broader family of similarity theories before addressing weak equivalence.

3.1. Strong equivalence with LTC

Three fundamental differences between LTC and SCM stand out as particularly salient: ontological priority, truth makers and vagueness.

3.1.1. Ontological priority

Consider the relations that might hold between the rates set by the Federal Reserve, the level of investment and the growth of the economy. Such relations include counterfactual dependence, type-level causal relations that hold across instances, and token-level causal relations between specific instances. SCM describes type-level causal relations as ontologically basic and counterfactual dependence and token-level causal relations as either derived from type-level causal relations (Halpern and Hitchcock 2015) or as mutually dependent (Woodward 2003). In contrast, Lewis takes counterfactual dependence as basic, token-level causal relations derived from counterfactual dependence, and type-level causal relations derived from token-level causal relations. Consequently, LTC views scientific laws as non-causal generalizations about matters of fact whereas SCM views scientific laws as general causal statements. Because laws impact similarity between worlds, this yields a further difference. For LTC, causation has no bearing on similarity, whereas when SCM is expressed in similarity terms, similarity is based on causal relationships represented in causal models (Galles and Pearl 1998). LTC can describe a world with counterfactual dependence but no causation but not vice versa whereas SCM, at least on one interpretation, can do precisely the opposite. LTC is consistent with a world in which all counterfactual dependencies hold between the wrong sorts of events to support causation. Although one might avoid circularity by treating token-level causal relations as dependent on counterfactual dependence and counterfactual dependence as dependent on type-level causal relations (Bennett 2003; Halpern and Hitchcock 2015), Lewis scrupulously avoided a causal theory of conditionals (Lewis 1986*b* pace Galles and Pearl 1998; cf. Lewis 1999: Ch. 15; Hall 2004; Briggs 2012) to safeguard against circularity in his counterfactual theory of causation. In contrast, Pearl dismisses counterfactual theories of causation as circular based on his causal theory of counterfactuals (Pearl and Mackenzie 2018). One interpretation of SCM might allow causal structures that produce no counterfactual dependence. A more natural interpretation might require that causation and counterfactual dependence always occur together. Either interpretation differs from LTC.

Counterfactual dependence is not limited to causal dependence but can also derive from logical dependence, means-end relations, relational properties, constituent properties (Kim 1974), supervenience (Woodward 2015; Baumgartner and Gebharter 2016), grounding (Schaffer 2016) and constitutional dependence (Starr 2019). LTC accommodates non-causal counterfactual dependence (Lewis 1986*b*: Ch. 23). If we take SCM at face value (e.g. the Causal Markov Condition is causal), then it denies any non-causal counterfactual dependence. However, it is not obvious that

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

structural models capture only causal dependence as opposed to dependence in general, in which case SCM is not limited to causation as advertised. However one resolves this structural ambiguity in SCM, it differs from how LTC handles non-causal dependence.

Ontological differences might not matter to policy analysts if both theories deemed the same conditionals as true, but they do not. Examples involving non-causal laws produce different results. Consider 'If this patch of ground had not contained talc it would not have contained asbestos' in a world where the two co-occur as a matter of non-causal law. Because violations of law produce less similarity than small local violations of fact, LTC judges the conditional true. Because there is no causal link between the presence of talc and the presence of asbestos, SCM judges the conditional false, or simply lacks the ability to represent the problem.

3.1.2. Truth makers

A second but related difference involves truth makers for counterfactual claims. Although it took time for readers to absorb the claim, Lewis held that every possible world existed in just the same way as our own (Lewis 1986*a*) and that while no information can flow between worlds, we can have knowledge of what is possible or necessary as modal expressions understood in our world (cf. Bennett 2003: Ch. 10). Pearl has tended to focus on correctly specified models, making it unclear whether the truth makers of counterfactual conditionals are the model itself or what it represents (Markus 2011; Cartwright 2017). However, even if we take the latter interpretation, he clearly rejects the idea of modal expressions depending upon other worlds for their truth values (e.g. Pearl and Mackenzie 2018: Ch. 8) presenting SCM as a more favourable alternative. One might reconstruct LTC to adjust it to Pearl's ontology in this respect but, again, the result would no longer be LTC but rather an alternative theory. Moreover, Lewis did not adopt this view lightly but did so in response to arguments against the view that our actual world exists in a manner not shared by other possible worlds. Lewis argued that for any recursively defined language, there will be more propositions required to fully describe the world than sentences available to express them (Lewis 1973*b*: Ch. 4). This poses a challenge for the idea that one could construct a structural causal model 'of the entire world' (Briggs 2012: 147) to bring SCM into alignment with LTC. Moreover, such a modification would conflict with typical uses of structural causal models applied to small sets of variables.

3.1.3. Vagueness

Lewis accepted vagueness as an inherent part of counterfactual conditionals in natural language use and attempted to match the appropriate level of vagueness in his truth conditions based on similarity (1973*b*: Ch. 4). In contrast, Pearl sought to replace vagueness with precision (Pearl 2009*a*; Pearl and Mackenzie 2018). LTC offers a descriptive theory constrained by natural language whereas SCM offers a prescriptive theory that seeks to improve upon natural language. As such, the two theories disagree regarding not just specific truth conditions but also on the general character of those truth conditions.

3.2. Strong equivalence with other similarity theories

LTC dramatically differs from SCM regarding counterfactual conditionals. Starr (2019) groups similarity theories with strict conditional theories as placing greater emphasis on logical truths than determining the truth of specific counterfactual assertions but he groups causal model approaches with two other types of theories that have the opposite emphasis. Other similarity theories have more in common with SCM by endorsing causal laws (Jackson 1977; Bennett 2003) or the Conditional Law of Excluded Middle (Stalnaker 1981) or alternative accounts of possible worlds (Stalnaker 1976). Nonetheless, such similarities do not negate the differences, nor do they reduce the value of comparative research.

3.3. Weak equivalence

As with RCM, it seems natural that one can re-express SCM in notation borrowed from LTC. This does not diminish the value of comparative research given that the theories are distinct even when expressed in the same notation.

Despite the fact that the constituent terms in LTC and SCM differ in meaning and reference, it still holds interest to compare their formal structure. Lewis (1973*b*: Ch. 5) considered whether his logic of comparative possibility might have applications in other areas without suggesting that these areas would thereby reduce to counterfactual conditionals. They simply involve distinct domains that may share a logical isomorphism similar to that between the brotherhood theory and the sisterhood theory from the Introduction. We now know that neither the logical system defined by Lewis' axioms for comparative possibility nor an axiomatic representation of SCM constitutes a special case of the other (Halpern 2013; Zhang *et al.* 2013). Setting aside hyperbolic paraphrases, Galles and Pearl only claimed that SCM imposes no further restrictions on LTC when restricted to recursive models. This claim is analogous to stating that propositional logic imposes no further constraints on predicate logic when applied to formulae constructed only by the conjunction, disjunction and negation of atomic propositions. This merely shows that the restrictions imposed on the comparison by recursive models already impose the restrictions inherent to SCM. LTC is not limited to causal models and thus does not share these restrictions.

Some of the divergence between the two can be seen in the following examples. $(A \mathbin{\Box\!\!\rightarrow} B) \vee (A \mathbin{\Box\!\!\rightarrow} {\sim}B)$, the Conditional Law of Excluded Middle: This is not implied by LTC but is implied by the Existence axiom in SCM, and by the grammar of causal models. Lewis (e.g. 1973*b*) offered arguments against it based on ties or an infinite series of more similar worlds. $(A \,\&\, {\sim}A) \mathbin{\Box\!\!\rightarrow} A$: LTC judges this unambiguously true whereas it violates the grammar of causal models and possibly the Effectiveness axiom of SCM. It thus falls into a truth value gap in SCM. $A \mathbin{\Box\!\!\rightarrow} \Diamond A$: This is a common inference pattern in empirical sciences where an experiment proves something possible. However, nothing causes itself to be possible, so the conditional must rely on some form of non-causal dependency (Kim 1974) excluded by SCM. Moreover, this may fall into a truth value gap because SCM lacks translation rules for modal operators. $(A \mathbin{\Box\!\!\rightarrow} (B \mathbin{\Box\!\!\rightarrow} C)) \equiv (A \wedge B) \mathbin{\Box\!\!\rightarrow} C)$, Import-Export: Natural extensions of SCM to handle iterated conditionals imply this

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

whereas LTC does not (Briggs 2012; Fisher 2017). McGee (1985) discussed a substitution instance that greatly illuminates the difference in intuitions here: $A \mathrel{\Box\!\!\to} (B \mathrel{\Box\!\!\to} \sim A)$. Any interventionist theory such as SCM would see this as a violation of Effectiveness because the initial antecedent manipulates A to true and thus it must remain so, contradicting the embedded consequent. However, from the perspective of LTC, the initial A is just a temporary context from which the embedding conditional is interpreted, but that need not carry over to the embedded conditional. Modifying an earlier example, imagine A = talc present and B = asbestos present. From the LTC perspective, the actual world is a $\sim A \wedge \sim B$ world and the nearest world is an $A \wedge B$ world. These two worlds are closer to one another than to other worlds because they both conform to the law that talc and asbestos occur together. To evaluate the conditional $A \mathrel{\Box\!\!\to} (\sim B \mathrel{\Box\!\!\to} \sim A)$, we first move to the nearest world to satisfy A, then, to evaluate the nested conditional, we move back to the actual world to satisfy $\sim B$. The conditional comes out true because the actual world is a $\sim A$ world. In contrast, SCM would combine the antecedents to take us directly to the nearest $A \wedge \sim B$ world where $\sim A$ holds false, and thus the conditional comes out false. (Figure 1; Further details in online supplement.)

3.4. Section summary

LTC and SCM differ substantially (Table 3). They have different implications for reasoning about economic policy and other conditionals. There remain many unanswered questions for comparative research applying these and other theories of counterfactual conditionals to economic research. To appreciate the differences, it is important that such comparative research consider the content of the theories and not just uninterpreted formal calculi. Identical formal expressions can mean very different things in each theory.

4. Anticipated objections[9]

One might object that SCM could be modified or extended to make either LTC or RCM equivalent to it. Such objections frequently focus on the extension of notation, which is why it is important to remember that the notation does not determine the content expressed with it. Such objections can also run afoul of parity arguments if they are offered in favour of one theory replacing another when either theory can be symmetrically extended to encompass the other. If modification is unconstrained, then any formal system can be modified to make it equivalent to any other. For the present article, I limit SCM to what Pearl and associates have presented it as and limit equivalence claims to actual equivalence to SCM without modification.

One might anticipate the objection that once a formal proof has been provided, objections to the conclusion can only represent failures to understand the proof. This objection rests on a misunderstanding of formal methods. The power of formal methods rests not in their implanting a Cartesian tree of knowledge that

[9]In anticipating the below objections, I implicitly draw on both SCM related publications and also more informal statements that have helped clarify the thinking behind the published statements.

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

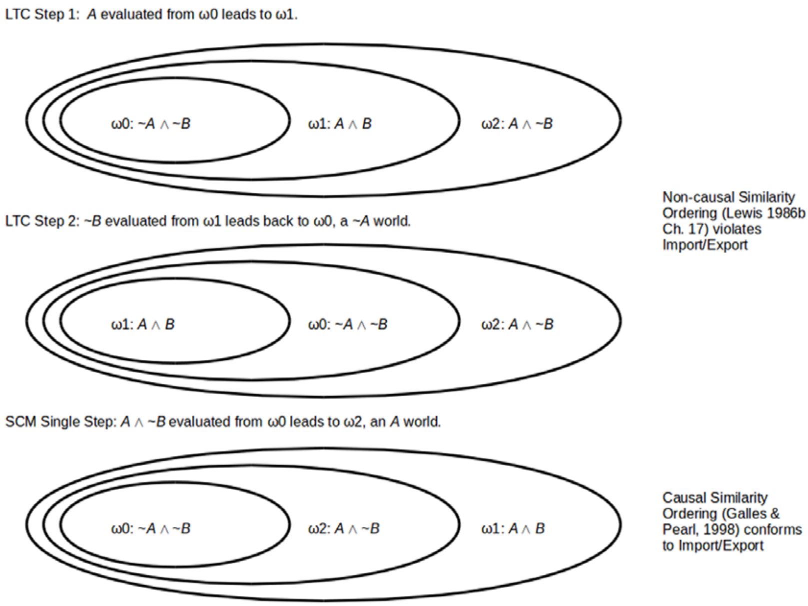

Figure 1. Contrasting Evaluations of the Nested Counterfactual Conditional A □→ ($\sim B$ □→ $\sim A$).

places existing results beyond criticism or revision. To the contrary, their power rests in the ability of formal methods to accelerate progress through theory revision by making it easier to discover and correct flaws and limitations (Pollock 1990). Moreover, agreement on the interpretation of the symbols constitutes a fundamental precondition for all formal methods. In the case of the equivalence claims, such agreement was lacking and premature formalism impeded progress by masking the differences.

One might also anticipate an objection that SCM represents the sole modern terminology for discussing causation and counterfactual conditionals, antiquating earlier approaches except to the extent that they are absorbed into SCM. This entails that the only interesting sense of equivalence consists of one internal to SCM that evaluates conformity of another theory to SCM. The criticism relies on circular reasoning given that it assumes what the critique of equivalence claims draws into question. More broadly, interpreting everything from within one perspective and systematically discounting anything external to that perspective impedes progress. Adopting such a strategy insulates the target theory from effectively responding to criticism rather than maximizing the benefits of such criticism. Science and philosophy both depend upon the effective use of critical feedback as a correction mechanism for theories.

A defender of equivalence might concede that LTC and RCM have interpretations that differ from SCM and still argue that one should dismiss such interpretations as inferior to interpreting them in a way that conforms with SCM. The objection asserts equivalence as normative rather than descriptive. Even if, as one might expect, SCM fulfils Pearl's goals better than RCM or LTC,

https://doi.org/10.1017/S0266267120000437 Published online by Cambridge University Press

Table 3. Contrasting two theories of counterfactual conditionals

LTC	SCM
Non-causal dependence	All dependence is causal
Non-causal counterfactual conditionals	All counterfactual conditionals are causal
Counterfactual theory of causation	Non-reductive theory of causation or no theory
All possible worlds exist in the same way	Only the actual world exists
Possible worlds are complete worlds	Possible worlds are partial world descriptions
Similarity is based on facts and laws	Similarity is based on causation
Laws are not causal	Laws are causal
Humean supervenience: Local matters of fact are ontologically basic	Type-level causal relationships are ontologically basic

the previous sections give reason to believe that there exist other goals, outside of SCM, that RCM and LTC fulfil better. As such, one finds no systematic argument that the goals of SCM exhaust the goals that might motivate interpretations of RCM and LTC outside of SCM.

An objection might draw from the various criticisms Pearl offers of LTC and RCM for not sharing one or another feature of SCM, or not having addressed some problem addressed by SCM (e.g. Pearl 2009*a*, 2012; Bollen and Pearl 2012; Pearl and Mackenzie 2018). Such objections again fail to maintain parity by making SCM the arbiter between itself and its rivals. Once one recognizes LTC and RCM as distinct programmes with their own distinct objectives, such arguments do not provide a reason to abandon those objectives in favour of SCM. Rather than framing them as competitors in a zero-sum game, it seems much more productive to approach them as distinct streams of research that can strengthen and benefit from one another (Markus 2013, 2016). A similar reply applies to any such objection constructed through a double-standard, dismissing LTC or RCM for reasons that also apply to SCM.

Finally, one can imagine the critique of equivalence claims being mischaracterized as hostile toward SCM. This objection rests on circular reasoning by presupposing that the equivalence claims play an essential part in the value of SCM. It also bears emphasis that the suppression of comparative research resulting from equivalence claims potentially hinders effective development of SCM as much as it does LTC and RCM. The critique of equivalence claims thus represents as much an effort to protect SCM from the harmful effects of equivalence claims as to protect its rivals.

5. Conclusion

SCM differs from both RCM and LTC in ways that impact prescriptive methodology for research practice. Comparative research can certainly incorporate analysis of commonalities between the theories. Given the substantial differences between SCM and each of the two rival theories, however, researchers should avoid mixing methodological advice between rival theories and methodologists should avoid mixing theorems between rival theories. Even when the same formal expression holds across theories, its interpretation may differ. As such, I hope that critical evaluation of the equivalence claims can re-open the door to greater exploration of the comparative strengths and weaknesses of the rival theories.

Acknowledgements. An earlier version of this article was presented at the International Meeting of the Psychometric Society, 18 August 2017, in Zurich, Switzerland. I am grateful to Rachel Perlin for comments on an earlier version of this article.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S0266267120000437

References

Baumgartner M. and A. Gebharter 2016. Constitutive relevance, mutual manipulability, and fat-handedness. *British Journal of Philosophy of Science* **67**, 731–756.

Bennett J. 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Bollen K.A. and J. Pearl 2014. Eight myths about causality and structural equation models. In *Handbook of Causal Analysis for Social Research*, ed. S.L. Morgan, 301–328. Dordrecht: Springer.

Briggs R. 2012. Interventionist counterfactuals. *Philosophical Studies* **160**, 139–166.

Cartwright N. 2017. Can structural equations explain how mechanisms explain? In *Making a Difference: Essays on the Philosophy of Causation*, ed. H. Beebee, C. Hitchcock and H. Price, 132–152. Oxford: Oxford University Press.

Fisher T. 2017. Causal counterfactuals are not interventionist counterfactuals. *Synthese* **194**, 4935–4957.

Galles, D. and J. Pearl 1998. An axiomatic characterization of causal counterfactuals. *Foundations of Science* **3**, 151–182.

Hall N. 2004. Rescued from the rubbish bin: Lewis on causation. *Philosophy of Science* **71**, 1107–1114.

Halpern J.Y. 2000. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research* **12**, 317–337.

Halpern J.Y. 2013. From causal models to counterfactual structures. *Review of Symbolic Logic* **6**, 305–322.

Halpern J.Y. and C. Hitchcock 2015. Graded causation and defaults. *British Journal for the Philosophy of Science* **66**, 413–457.

Hausman D.M. 1998. *Causal Asymmetries*. Cambridge: Cambridge University Press.

Hiddleston E. 2005. Causal powers. *British Journal for the Philosophy of Science* **56**, 27–59.

Holland P. 1986. Statistics and causal inference. *Journal of the American Statistical Association*, **81**, 945–970.

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

Imbens G.W. 2019. *Potential outcome and directed acyclic graph approaches to causality: relevance for empirical practice in economics*. National Bureau of Economic Research working paper no. 26104. doi: 10.3386/w2614.

Imbens G.W. and D.B. Rubin 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, NY: Cambridge University Press.

Jackson F. 1977. A causal theory of counterfactuals. *Australasian Journal of Philosophy* **55**, 3–21.

Kim J. 1974. Noncausal connections. *Noûs* **8**, 41–52.

Lewis D. 1973a. Causation. *Journal of Philosophy* **70**, 556–567.

Lewis D. 1973b. *Counterfactuals*. Malden, MA: Blackwell.

Lewis D. 1973c. Counterfactuals and comparative possibility. *Journal of Philosophical Logic* **2**, 418–446.

Lewis D. 1986a. *On the Plurality of Worlds*. Malden, MA: Blackwell.

Lewis D. 1986b. *Philosophical Papers*, Vol. **2**. Oxford: Oxford University Press.

Lewis D. 1999. *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.

Lewis D. 2004. Causation as influence. In *Causation and Counterfactuals*, ed. J. Collins, N. Hall and L.A. Paul, 75–106. Cambridge, MA: MIT Press.

Markus K.A. 2011. Real causes and ideal manipulations: Pearl's theory of causal inference from the point of view of psychological research methods. In *Causality in the Sciences*, ed. P. McKay Illari, F. Russo and J. Williamson, 240–269. Oxford: Oxford University Press.

Markus K.A. 2013. An incremental approach to causal inference in the behavioral sciences. *Synthese* **191**, 2089–2113.

Markus K.A. 2016. Consistent treatment of variables and causation poses a challenge for behavioral research methods: a commentary on Nesselroade and Molenaar (2016). *Multivariate Behavioral Research* **51**, 413–418.

McGee V. 1985. A counterexample to modus ponens. *Journal of Philosophy* **82**, 462–471.

Morgan S.L. and C. Winship 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Pearl J. 2009a. *Causality: Models, reasoning and inference*, 2nd edn. Cambridge: Cambridge University Press.

Pearl J. 2009b. Remarks on the method of propensity score. *Statistics in Medicine* **29**, 1415–1416.

Pearl J. 2009c. *Myth, Confusion, and Science in Causal Analysis.* Unpublished. http://ftp.cs.ucla.edu/pub/stat_ser/r348-warning.pdf.

Pearl J. 2012. The causal foundations of structural equation modeling. In *Handbook of Structural Equation Modeling*, ed. R.H. Hoyle, 68–91. New York, NY: Guilford Press.

Pearl J. 2018. Does obesity shorten life? Or is it the soda? On non-manipulable causes. *Journal of Causal Inference* **6**. doi: 10.1515/jci-2018-2001.

Pearl J. 2019. On the interpretation of do(x). *Journal of Causal Inference* **7**. doi: 10.1515/jci-2019-2002.

Pearl J. and E. Bareinboim 2014. External validity: from Do-calculus to transportability across populations. *Statistical Science* **29**, 579–595.

Pearl J. and D. Mackenzie 2018. *The Book of Why: The New Science of Cause and Effect*. New York, NY: Basic Books.

Pearl J., M. Glymour and N.P. Jewell 2016. *Causal Inference in Statistics: A Primer*. Chichester: Wiley.

Pollock J.L. 1990. *Nomic Probability and the Foundations of Induction*. New York, NY: Oxford University Press.

Rosenbaum P.R. and D.B. Rubin 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55.

Rubin D.B. 1974. Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* **66**, 688–701.

Rubin D.B. 1978. Bayesian inference for causal effects: the role of randomization. *Annals of Statistics* **6**, 34–58.

Rubin D.B. 1980. Randomization analysis of experimental data: the Fisher randomization test comment. *Journal of the American Statistical Association* **75**, 591–593.

Rubin D.B. 2007. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in Medicine* **26**, 20–36.

Rubin D.B. 2009. Should observational studies be designed to allow lack of balance in covariate distributions across treatment groups? *Statistics in Medicine* **29**, 1420–1423.

Schaffer J. 2016. Grounding in the image of causation. *Philosophical Studies* **173**, 49–100.

Schmaltz T.M. 2014. *Efficient Causation: A History*. Oxford: Oxford University Press.

Shadish W.R., T.D. Cook and L.C. Leviton 1991. *Foundations of Program Evaluation: Theories and Practice*. Newbury Park, CA: Sage.

Shrier I. 2008. Letter to the Editor. *Statistics in Medicine* **27**, 2740–2741.

Sjolander A. 2009. Propensity scores and M-structures. *Statistics in Medicine* **29**, 1416–1420.

Spirtes P., C. Glymour and R. Scheines 2000. *Causation, Prediction and Search*, 2nd edn. Cambridge, MA: MIT Press.

Stalnaker R. 1968. A theory of conditionals. In *Studies in Logical Theory*, ed. N. Rescher, 98–112. Oxford: Blackwell.

Stalnaker R. 1976. Possible worlds. *Noûs* **10**, 65–75.

Stalnaker R. 1981. A defense of conditional excluded middle. In *Ifs*, ed. W.L. Harper, R. Stalnaker and G. Pearce, 87–104. Dordrecht: D. Reidel.

Starr W. 2019. Counterfactuals. In *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), ed. E.N. Zalta. <https://plato.stanford.edu/archives/fall2019/entries/counterfactuals/>.

Steyer R. 2005. Analyzing individual and average causal effects via structural equation models. *Methodology* **1**, 39–54.

Woodward J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward J. 2015. Interventionism and causal exclusion. *Philosophy and Phenomenological Research* **91**, 303–347.

Woodward J. 2016. The problem of variable choice. *Synthese* **193**, 1047–1072.

Zhang J., W.-L. Lam and R. De Clercq 2013. A peculiarity in Pearl's logic of interventionist counterfactuals. *Journal of Philosophical Logic* **42**, 783–794.

Keith A. Markus is Professor of Psychology at John Jay College of Criminal Justice of The City University of New York and also a member of the faculty of the Graduate Center of The City University of New York. He is coauthor with Denny Borsboom of *Frontiers of Test Validity Theory: Measurement, Causation and Meaning*. His research focuses on causal explanation, causal inference, and test validity. URL: http://jjcweb.jjay.cuny.edu/kmarkus/.

Cite this article: Markus KA (2021). Causal effects and counterfactual conditionals: contrasting Rubin, Lewis and Pearl. *Economics and Philosophy* **37**, 441–461. https://doi.org/10.1017/S0266267120000437