



The specter of irreparable ignorance: counterfactuals and causality in economics

George F. DeMartino¹ 

Accepted: 15 November 2020 / Published online: 8 January 2021
© European Association for Evolutionary Political Economy 2021

Abstract

Those economists who have emphasized true uncertainty have tended to draw an epistemic distinction between an ascertainable past and an unknowable future. But in one critical respect—in extracting causal relationships—that epistemic distinction is not warranted. Whether they are situated in the past or future, causal arguments in economics depend equally on counterfactual reasoning. Counterfactualizing entails the construction of fictitious narratives—narratives about worlds that do not exist. Unfortunately, there is no dependable method for ascertaining the uniquely correct counterfactual. This implies that causal claims in economics, too, are irreducibly fictitious. The chief value of counterfactuals, then, is not to prove causation but to help scholars and practitioners confront an inscrutable world—to imagine and prepare for unknowable possible futures. In this endeavor, economic pluralism, which expands the range of plausible counterfactuals, is to be taken as a virtue rather than a curse.

Keywords Uncertainty · Economics · Counterfactuals · Causality · Irreparable ignorance · Pluralism

Nothing can ever happen twice.
In consequence, the sorry fact is
that we arrive here improvised
and leave without the chance to practice.
– Wislawa Szymborska (1998), *Nothing Twice*

Science is based on counterfactuals. Human knowledge is based on constructing counterfactuals. Blind empiricism leads nowhere.
– James J. Heckman 2001

✉ George F. DeMartino
George.DeMartino@du.edu

¹ Josef Korbel School of International Studies, University of Denver, Denver, CO, USA

1 Introduction—*time-travel machines*

A central aspect of economic practice involves modeling future impacts of policy measures. Economists recognize of course that the social world is nondeterministic and/or should be theorized as such given epistemic limitations, which complicates forecasting. But economists typically presume nonetheless that they can map the probability distribution of policy effects which permits dependable estimates of expected utility and other magnitudes of interest. This belief is central to the twentieth century modernist economic project: to devise means to know the future so as to exert control in hopes of promoting social betterment.

Orthodox micro- and macroeconomic theory has pursued various stratagems since the late nineteenth century to repress the problem of an opaque future. Stratagems include the assumptions of rationality and perfect information, partial equilibrium analysis with a liberal reliance on Marshallian *ceteris paribus*, the Walrasian framework with its omniscient auctioneer that yields general equilibrium only by disallowing trading at non-equilibrium prices (Walras' *tâtonnement* process), and comparative statics that represses complex dynamics as the economy presumably evolves between period t_0 to t_j . Later on the new classical macroeconomists theorized economic actors as forming rational expectations of the future. Each, in its own way, sought to domesticate the future, providing economists with a *time-travel machine* that seemed to allow them to see tomorrow, today.

But knowing future economic events is a fraught exercise. Just why this is so has been probed extensively over the past century by important epistemic iconoclasts, and I will not reproduce their arguments here.¹ The dissenters are right to perceive a *specter of irreparable ignorance* haunting economic analysis of the future. That insight has never been well received in the profession. It is disturbing to accept that there are tight constraints not just on what we know now and will know tomorrow but on what we can in principle ever know. It is equally disturbing to acknowledge that these epistemic limits crisscross the terrain of what both theoretical and applied economists would want to know and would need to know in order to do much of what they are presently doing.

Those economists who have emphasized irreparable ignorance have tended to draw a sharp epistemic distinction between the future and the past. In a famous passage, John Maynard Keynes emphasizes our ignorance of the future, not the past. Speaking about events like “the prospect of a European war...the price of copper and the rate of interest twenty years hence...” he writes “About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know...” (Keynes 1937, 213–214).² Knight (2014[1921], 199), too, was preoccupied with our ignorance of future events: “We live only by knowing something about the future;

¹ In addition to those cited in the text the list of epistemic skeptics in economics must include Robbins (1927), Hayek (1974), Hirschman (2013 [1970]), Resnick and Wolff (1987), Nelson (2004), Colander (2005), Easterly (2006), Rodrik (2017), and Crotty (2019). See DeMartino (2011a, b, 2013).

² For Keynes (1936[1973], 152–153) our knowledge of the more immediate future depends on agents' acting on the basis of conventions, such as “that the existing state of affairs will continue indefinitely, except in so far as we have reason to expect a change.” Keynes recognized, however, that this and other conventions are “precarious” and that their unpredictable collapse alters agents' behavior and the course of even near term events. It is at least implicit in this argument that agents do apprehend the past and present (“the existing state of affairs”) with relative certainty. I thank an anonymous referee for making this point.

while the problems of life, or of conduct at least, arise from the fact that we know so little” (cf McCloskey 1990). Shackle (1992 [1972], xi) is most explicit about the distinction between the epistemic status of the past and future. Speaking of Myrdal and Lindahl’s conception of *ex ante* and *ex post* analysis, Shackle contrasts “the forward look into the *void of unknowledge*, and the backward look into the past with its *ascertainable history*” (emphasis added).

The distinction between an *ascertainable past* and an *unknowable future* is intuitively attractive and has been readily accepted by many economists and others who probe the limits to knowledge. As Fischhoff (1982, 335) puts it, “Whatever the question we are asking, it is generally assumed that the past will readily reveal the answers it holds. Of hindsight and foresight, the latter appears as the troublesome perspective.” As concerns the past social world, we have a good bit of information, imperfect though it is, about what in fact happened. By “what happened” I mean the occurrence and sequence of certain events and the approximate values of many variables at regular intervals. As concerns the future, on the other hand, it seems that we cannot know in advance of its arrival the occurrence or order of many events, or the values of key variables.

Consider a scenario in which economists are debating whether raising the minimum wage in New York will increase unemployment 1 year, hence, and if so, by how much. We expect that their predictions might conflict, but we might at least expect the economists to be able to agree on and be certain about the effects of *past* increases in the minimum wage. This question, we might think, submits to certain, objective analysis, relying on well-established theoretical models and empirical techniques.

The intuition, however, is incorrect. Policy design, evaluation, and implementation are ultimately about control in two respects: by changing, say, this or that law, regulation or policy setting X , one can cause a predictable effect Y in line with one’s goals. The control of X enables the control of Y . Hence, it requires an understanding of causal mechanisms—what causes what, or by how much the magnitude of one variable affects another, and why (Heckman 2008). And when it comes to *that* kind of knowledge—knowledge of causal mechanisms and pathways—the basis for a sharp epistemic distinction between the past and the future evaporates. The reason for this, as will be discussed below, is that causal arguments in economics, whether they are situated in the past or the future, depend on counterfactual reasoning. But counterfactual arguments are by their nature *fictitious*. This implies that causal arguments in economics, too, are fictitious. As concerns causality,³ then, and as I will endeavor to show, the past is no more epistemically available than is the future. Causal knowledge of the past eludes us in part because of cognitive biases that I will mention only in passing. I will probe here the more fundamental epistemic problem associated with counterfactualizing that would not be resolved even in the absence of cognitive biases.

To anticipate a likely semantic concern, to say that social science counterfactuals are fictitious does not imply that they lack all reason or empirical grounding. As will become clearer below, I mean that they are statements about alternative worlds—recognizable worlds—which do not exist other than in our imaginations. As a practical matter, the correct standards of judgment, then, do not include their

³ I use the terms causality and causation interchangeably throughout.

truth status or the truthfulness of the causal claims they sustain. But they can nonetheless be judged. The standard is their *productivity*—the degree to which they help us confront the unknowable worlds we inhabit and prepare for unpredictable futures. This is just how they are used by those in the field of “decision-making under deep uncertainty,” discussed below, where practitioners who acknowledge severe epistemic insufficiency look to discover *robust, adaptive* policy that promises to generate good outcomes over a very wide range of possible futures. Recognizing the fictitious nature of causal claims, then, promises to improve rather than obstruct good decision-making.⁴

2 Causation and counterfactuals

A useful definition of counterfactuals is that they are “subjunctive conditionals in which the antecedent is known or supposed for purposes of argument to be false (Skyrms 1980)” (Tetlock and Belkin 1996, 4). Consider the simple statement, “Had central banks pursued policy associated with Modern Monetary Theory, the Great Recession could have been avoided.” “Had central banks pursued policy associated with Modern Monetary Theory” is the false antecedent; “the Great Recession could have been avoided” is the subjunctive conditional.

Counterfactual reasoning is inescapable in daily life, and in the social sciences. I emphasize the point here so that my intervention will not be taken as a critique or, even worse, a rejection of counterfactualizing. According to Judea Pearl (Pearl and Mackenzie 2018), a leading theorist today on formal structures of counterfactualizing in science (Reiss and Cartwright 2003), the ability to counterfactualize distinguishes human reasoning from the cognitive capacities of other species and represents the most formidable challenge to the realization of artificial intelligence. Taleb (2010, 189) speculates that “the ability to project conjectures into the future and play the counterfactual game” may be the “most potent use of our brain.”

Counterfactual reasoning applies equally to postdiction and prediction of events and to consideration of our past actions and the actions we contemplate in the present (Lebow 2010, 276). But most causal accounts of the past and future in economics do not refer explicitly to subjunctive conditionals.⁵ Economists are less inclined to frame their causal arguments in terms of counterfactual reasoning than are political scientists, historians, and other scholars who routinely probe individual cases, such as the causes of WWI, or comparativists who explore small numbers of cases. Relative to training in political science and history, economic training underemphasizes counterfactual methods. But economic explanation is just as dependent on counterfactual analysis as are political and historical explanations. Why is this so?

The answer stems from the structure of standard causal claims that are routinely encountered in the social sciences, including economics, and that underpin much of the theoretical work on counterfactualizing (e.g., Lewis 1973, 556). In a passage that

⁴ If the embrace of epistemic insufficiency risks nihilism, as many fear, it is of the “sensible” sort advocated by Samuels (1993) that urges humility but certainly does not impede decision-making.

⁵ The most notable exception is Fogel (1964). See also Goldin (1973).

appears frequently in the literature on counterfactuals, Hume (2007 [1748], section VII) famously defines causality as follows:

We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. Or in other words where, if the first object had not been, the second never had existed.

As is widely recognized, Hume's statement conflates two distinct arguments concerning causality. The first sentence defines causality in terms of constant conjunction (Goertz and Levy 2007; Lewis 1973), which suggests a relation of sufficiency. The second sentence defines causality instead in terms of necessity.⁶

When economists claim that X is causally related to outcome Y , they are typically making one of several kinds of claims, such as (a) through (c):

- (a) that X is sufficient for Y ($X \rightarrow Y$)
- (b) that X is necessary for Y ($\sim X \rightarrow \sim Y$, where \sim stands for "not")
- (c) that X is necessary and sufficient for Y ($X \rightarrow Y$ and $\sim X \rightarrow \sim Y$)

(a) is consistent with Hume's constant conjunction definition of causality; (b) is Hume's necessity definition. Establishing the strongest causal connection between X and Y requires presumption (c) since (a) implies that had X not intervened, Y might have occurred anyway owing to some other causal factors (the case of overdetermination), while (b) is satisfied even when X serves as a necessary background condition that permitted but did not cause Y in the useful sense of "cause." Goodman (1947) illustrates the point with the example of the presence of oxygen for the lighting of a match. Under normal circumstances, in answering the question "Why did the match light?", oxygen is treated as necessary for but not the cause of a match lighting.

Whether in any particular case, we claim (a), (b), or (c) is in fact immaterial for the argument advanced here. All three claims entail the exploitation of a counterfactual in which X does not occur. (b) and (c) imply that but for X , Y would not have occurred (see Goertz and Levy 2007, 15; Lewis 1973, 557). The sufficient but not necessary account given by (a) is trickier. Although Y might have occurred even in the absence of a particular cause X , (a) implies that the absence of X matters.⁷ Consider a multivariate explanation of a single case, such as the causes of WWI. Here the claim that some X is sufficient but not necessary entails a counterfactual claim that had X not happened, Y could have happened nonetheless as a consequence of some other causal pathway not involving X (Pearl and Mackenzie 2018). That claim is explicitly counterfactual, since it describes a world in which X did not occur.⁸

⁶ In citing Hume's approach to causality, I follow the established tradition in the counterfactual literature. Mill (1843[2002], ch. VIII) offers a more sophisticated treatment of inference and experimental inquiry, as summarized in his five canons of causality. Mill's approach to causality, too, implies counterfactual reasoning. For alternative accounts of causality, see Mackie (1965), Cartwright (2004, 2007), and Fullbrook (2007).

⁷ Elster (1978, 176) argues that Hume's sufficient, constant conjunction statement "implies a counterfactual statement, even if the theory does not accept that causation should be defined in terms of counterfactuals."

⁸ Where instead a causal claim of sufficiency is made in the large- n multivariate context, the absence of a sufficient cause X (where $X = 0$) (a) reduces the probability of occurrence of Y by some magnitude where Y is a binomial variable; or (b) alters the expected value of a continuous variable Y by eliminating one of multiple causal pathways that bear on Y .

3 Counterfactuals: past and present thinking

Various political scientists, historians, modal logicians, computer scientists, and economists have defined causality in terms of counterfactuals or argued that causal claims necessarily require counterfactualizing.⁹ Tetlock and Belkin (1996, 3-4) make the point directly:

We can avoid counterfactuals only if we eschew all causal inference and limit ourselves to strictly noncausal narratives of what actually happened (no smuggling in causal claims under the guise of verbs such as ‘influenced,’ ‘responded,’ ‘triggered,’ ‘precipitated,’ and the like).

Among economists, Jevons (1882, 240) embraces a counterfactual approach to causation: “By the Cause of an event we mean the circumstances which must have preceded in order that the event should happen... A necessary or indispensable antecedent is however identical with a cause, being that without which the effect would not take place.” Hicks (1980, 7-8) explicitly centers his account of causality on counterfactualizing:

What then do we mean, in terms of the New Causality, if we assert that A caused B? ... We will take it that we have satisfactory evidence that both of these events did actually occur. For causality, we must be maintaining that if A had not existed, B would not have existed; if not-A, then not-B. But not-A and not-B are not events which have happened; they are events which have not happened. (In recent discussions among historians they are described as ‘counterfactual’).¹⁰

The most prominent advocate of the centrality of counterfactuals to causality in economics today is the econometrician and microeconomist J.J. Heckman. In his view (2001, 17),

A causal relationship is only well defined if a theory of potential counterfactuals is articulated and a mechanism generating variation in the causes is clearly specified. Both steps are pre-statistical and require thought experiments involving counterfactuals in imaginary worlds.

I do not take a position here on whether *all* causal claims do in fact require counterfactualizing (though I suspect they do). I restrict my goal to demonstrating that counterfactual reasoning is central to the standard forms of causal claims in economics, and to teasing out important epistemic implications.¹¹ And to achieve those goals we

⁹ In addition to the collections and essays explored here, see also Simon (1952); Abbott (1974); Sugden (2000); Morgan and Winship (2007); Nathan (2017); and Mahoney and Barrenechea (2017).

¹⁰ Hicks (1980) explores Adam Smith’s use of counterfactual reasoning in his analysis of the role of navigable waterways in economic development. Schliesser (2017, ch. 5) finds that Smith also relies on counterfactualizing in his treatment of perfect sympathy, natural price, and optimal institutions.

¹¹ My interest is the epistemic problem associated with what Reiss and Cartwright (2003) call “genuine” questions: questions that arise in the context of actual (past or present) or potential (future) economic policy interventions. I therefore sidestep the distinct semantic issues at the heart of the philosophical literature on causality and counterfactuals which focuses on the meaning of causal and counterfactual statements. Reiss and Cartwright (2003) do examine the semantic problem and illuminate how difficult the problem is. In particular, they challenge the semantics of both Pearl and Heckman’s respective approaches.

must take account of the challenge that economists must attend to, in one way or another, when specifying causation—Holland’s (1986, 947) “Fundamental Problem of Causal Inference”:

Causal inference is ultimately concerned with the effects of causes on specific units, that is, with ascertaining the value of the causal effect in (1) $[Y_t(u) - Y_c(u)]$, which gives the difference in the value Y of unit u under a treatment (t) and non-treatment (c). It is frustrated by an inherent fact of observational life that I call the Fundamental Problem of Causal Inference.

Fundamental Problem of Causal Inference. It is impossible to observe the value of $Y_t(u)$ and $Y_c(u)$ on the same unit and, therefore, it is impossible to observe the effect of t on u .

The problem arises because “we cannot rerun history at the same time and the same place with different values of our explanatory variable each time” (King et al. 1994, 91). Hence the need for counterfactualizing machinery to mentally simulate the rerunning of history. But the veracity of the counterfactual scenario—say, where X does not occur—cannot ever be confirmed or disconfirmed empirically *because that historical path was precluded when X , in fact, occurred*. When establishing a causal explanation, then, we must generate a *fictitious narrative*, a narrative about *what would have been* in an alternative world that was identical to the existing world in every respect but for the fact that X did not occur.¹² Counterfactualizing requires generating a hypothesized account of an imagined world. At the heart of objective, causal claims, then, is the construction of a *fiction* without which the objective causal claim cannot be sustained (cf. Beckett 2016, 51ff).¹³

Social *science* grounded in *fiction*: the claim rings false. But a number of other theorists, including economists, have emphasized the role of fiction in social science. Edward Leamer goes to the heart of the matter at the outset of his macroeconomic textbook:

You may want to substitute the more familiar *scientific* words ‘theory and evidence’ for ‘patterns and stories.’ Do not do that. With the phrase ‘theory

¹² It may be that the nonoccurrence of X requires, for the sake of plausibility, many other changes in the alternative world. But how are we to know what other changes must be made to accommodate the counterfactual antecedent? See below.

¹³ Researchers who engage the problem hope that it can be managed with two kinds of ontological assumptions about the relationship between causal variables and the nature of their causal effects, which are not always made explicit: (1) that the world can be usefully partitioned into the realm of necessity (denoting what is often defined in terms of structural relations, or laws) and the realm of contingency (denoting non-structural, chance, or law-defying relations); and (2) unit homogeneity or, minimally, “constant causal effect” on dependent variables (see King et al. 1994, 91; Holland 1986, 948). On the first presumption, see DeMartino (1993a, b); Russo et al. (2008); and Hay (2020). Those who emphasize ontological necessity in explanatory accounts are drawn to “powderkeg/spark” counterfactuals (see Thompson 2007), while those who emphasize the salience of contingency are drawn to “catalyst” counterfactuals (see Lebow 2010). See Hay’s (2020) rebuke to Hobsbawm, where Hay emphasizes the ontological (as opposed to the epistemological) nature of contingency.

and evidence' come hidden stow-away after-the-fact myths about how we learn and how much we can learn. The words 'theory and evidence' suggest an incessant march toward a level of scientific certitude that cannot be attained in the study of complex, self-organizing human system that we call the economy. *The words patterns and stories much more accurately convey our level of knowledge, now, and in the future as well. It is literature, not science.* (Leamer 2009, 3, emphasis added)

Scholars who share this insight tend to offer demarcation criteria to distinguish scientific from unscientific counterfactuals. For instance, Morgan (2013, 235) carefully probes the interpenetration of fact and fiction in economic models that permit *what-if* analysis, advancing the claim that “*Imaginative stories* prompted by the what-if questions that economists like to ask about their model worlds are where we see *economists playing their games of make-believe*” (emphasis added). In her view, what separates economic science from science fiction is the practice of “keying” or matching aspects of a fictitious model to facets of the actual world, so that inferences from models can tell us something useful about the world. She urges us to think of models as *artifacts*—as tools for social investigation that nonetheless incorporate fictitious claims.

Morgan's approach is insightful and promising. It puts to rest the worry that in the domain of social science counterfactualizing, anything goes. The problem, however, which Morgan herself acknowledges, is that good science fiction (and other genres of fiction, including fantasy) operates in much the same way. Good fiction often teaches us about our world by probing some imagined world that is very much like our own in most respects (keyed to our world), but unlike ours in one or more salient respects (see Beckert 2016, 63–64). As Lebow (2010, 278) puts it, “The conspicuous violation of real-world conditions, limitations or social practices in such fiction ironically requires the emphatic continuity of other real-world givens (e.g., gender, bodies, eating, death, politics, ethics).” Fiction that is too remote from us may not hold our interest, let alone teach us about ourselves and our world. In good fiction, a narrative unfolds that yields up insights we might not otherwise achieve. My own inclination, formed in part by engaging Morgan, is not to expect or look to provide unbending epistemic demarcation criteria for separating social science counterfactual fictions from other fictions. Instead, we should recognize the fictitious qualities of good social science (cf McCloskey 1998; Lebow 2010). Lebow (2010, 259) warns political scientists that a failure to do so is harmful: “One of the more interesting—and ironic—findings in this regard is the degree to which a commitment to science and scientific methods by international relations scholars can constitute a major impediment to their practice of science.” The same is certainly true of economists.

4 Economic research methods and the pursuit of the “right” counterfactual

The problem of the fictitious nature of counterfactuals would register as a benign curiosity rather than a pressing epistemic problem were there to be just one or a small set of plausible counterfactuals that generate closely aligned causal

inferences. As concerns complex social phenomena of the sort economists investigate, however, there are in principle innumerable plausible counterfactual fictions leading to conflicting causal inferences. The range of viable economic counterfactuals is limited only by economists' imagination and purpose. In economics (and other social sciences), the reach of human imagination and the diversity of purposes is evidenced by the proliferation of contending theoretical frameworks and by the range of research methods.

Owing to its self-conception as *science*, and one that seeks to generate interventions in the world that depend on causal knowledge, economists have worked to circumscribe the range of appropriate counterfactual reasoning. Along with sanctioned theoretical frameworks, approved research methods provide this service. I would suggest that research methods in economics are not principally what they seem to be. They are not principally guides to help us see correctly *what is* in the world. They are better theorized as dictates concerning how to construct the uniquely correct account of *what is not* so as to generate and sustain uniquely correct causal inferences. This claim requires justification. Here I survey some of the chief methods by which economists hope to overcome the pesky problem of the counterfactual to sustain causal claims.¹⁴

4.1 Randomized controlled trials

Economic randomized controlled trials (RCTs) today are contesting established research methods in economics for generating causal claims (Deaton 2009; Deaton and Cartwright 2018). RCTs appear to circumvent the Fundamental Problem of Causal Inference (see King et al. 1994, 94) and to displace counterfactualizing. That view is incorrect and is not held by leading RCT researchers.

RCTs are typically pursued to infer causal connections at the level of micro policy interventions, often in the context of economic development. The approach requires careful stratification and assignment in the construction of treatment and control groups. Finding a statistically significant difference between the average outcome of a treatment and no treatment permits the inference that the treatment caused the observed outcome. As Deaton and Cartwright (2018, 2) put it, "RCTs are perceived to yield causal inferences and estimates of average treatment effects (ATEs) that are more reliable and more credible than those from any other empirical method." It is useful to continue the quotation, since it bears directly on the question as to why RCTs are thought to generate dependable causal inference:

[RCTs] are taken to be largely exempt from the myriad problems that characterize observational studies, to require minimal substantive assumptions, little or no prior information, and to be largely independent of 'expert' knowledge that is often regarded as manipulable, politically biased, or otherwise suspect.

¹⁴ The focus in this paper is on standard or mainstream economics, not heterodox political economy. The problem of counterfactualizing emerges in heterodox causal accounts as well. I note, however, that critical realism rejects the role of counterfactualizing in causal explanations (e.g., see Lawson 2009). In contrast, "anti-essentialist" Marxism solves the problem by displacing causality in favor of the relation of "overdetermination." See Ruccio's (2005) critique of Lawson for a concise introduction. Ruccio and Amariglio (2003) and Resnick and Wolff (1987) present comprehensive explorations of the approach.

At best the assignment to treatment and control groups manages the Fundamental Problem—it cannot be avoided. It does so by engaging counterfactual reasoning in two ways. First, we are urged to presume that had the treatment group not received the treatment the counterfactual outcome would be identical to the outcome observed in the control group and vice versa. In the words of Duflo et al. (2006, 8), “Since the treatment has been randomly assigned, individuals assigned to the treatment and control groups differ in expectation only through their exposure to the treatment. Had neither received the treatment, their outcomes would have been in expectation the same.”

RCTs are an attempt to approach this ideal condition. But doing so requires that all the salient observable and unobservable features of the two populations, and the environments they inhabit, are identical in all respects other than the fact that one did and one did not receive the experimental treatment. That condition is approached to varying degrees depending on the nature of the problem under study but is never in fact achieved, owing to the inherent complexity of individuals, social relations, and the systems in which they are embedded; the simultaneous operation of multiple social and natural systems that impact outcomes; and myriad other contextual factors. Equally important, specifying the degree to which the condition is realized in any particular case requires judgments about what are the salient features that must be identical or sufficiently similar. The problem is that making those judgments (salience, sufficient similarity) requires the introduction of “expert knowledge.” Only theory can instruct researchers as to what must be included and what can be excluded in judgments about whether the treatment and control group are sufficiently alike. Until social science develops the means to run history multiple times, however, we must accept the fact that the counterfactual yielded by even the most impeccable empirical methods is irreducibly fictitious since it depends on claims about alternative worlds in which the treatment and control groups are in fact identical.

RCTs also engage counterfactual reasoning in the application of the findings of field studies in one context to other contexts, which depends on ascertaining the “external validity” of a study’s findings. RCTs are intended to acquire knowledge about a treatment for subsequent use in other populations (including perhaps the research population at a later point in time). In converting research results to treatment, we must be prepared to presume that if the same treatment had been applied to other sets of sufficiently similar individuals, then the result would have been causally identical (or at least similar enough) to the trial results to warrant the treatment. To presume otherwise—to presume that the findings from one trial do not generate insights applicable elsewhere—would be to recognize that policy interventions entail *just one damn experiment after another*, that the moment of “treatment” is continually deferred. There is much to be said for that way of thinking about policy, as Hirschman understood (see Grabel 2017; DeMartino and Grabel 2020). But it is not how RCT’s are theorized (see Duflo et al. 2006). They are intended to yield dependable causal connections grounded in the uniquely correct counterfactuals that can be exploited elsewhere in the design of treatment. Counterfactual reasoning all the way down. Unfortunately, as Reiss and Cartwright

(2003, 20) warn us, “There is no systematic answer available to the question of how to establish external validity.”¹⁵

4.2 Econometrics

Much of the methodological literature on counterfactualizing focuses on a single or a small number of cases and implies that counterfactualizing is essential in those contexts but not in large- n research involving structural modeling (e.g., Goertz and Levy 2007, ch. 2; Russo et al. 2008). That view is incorrect. Fearon (1996, 61) provides insight:

in regression analysis and other statistical means of testing causal hypotheses, one assumes that if any particular case in the sample had taken a different value on one of the independent variables, the dependent variable would have differed by a systematic component that is the same across cases plus a random component.

For instance, a simple, linear multivariate model may specify inter alia a causal relation between a change in household income in t_0 (y_0) and a lagged change in consumption in t_1 (c_1). Once we calculate the mean effect of Y on C , we are urged to expect that in any particular randomly selected household h_i in the sample H , a counterfactual change in the actual value of y_i would be likely to have an effect on c_i given by the magnitude of the inferred parameter (plus a random error term).

Heckman argues explicitly that causal inference in the econometric approach depends on counterfactual reasoning. For Heckman (2008, 3), counterfactualizing represents the essential first step:

The econometric approach to causal inference carefully distinguishes three problems: (a) Defining counterfactuals, (b) Identifying causal models from idealized data of population distributions (infinite samples without any sampling variation), and (c) Identifying causal models from actual data, where sampling variability is an issue. (Heckman 2008, 3)

¹⁵ Duflo et al. (2006, Part 8) recognize the problem. They examine the difficulties surrounding establishing external validity and present a range of strategies to confront them. Some of the difficulties are unique to RCTs, while some arise within other approaches to policy research. For present purposes, the researchers note that since a policy’s impact cannot be tested relative to “every single permutation and combination of contexts, we must rely on theories of behavior that can help us decide whether if the program worked in context A and B it is also likely to work in C” (71). The problem also arises with respect to evaluation of similar but not identical programs. They write, “Here again, it is a combination of replications and theory that can help generalize the lessons from a particular program” (see also the discussion in section 8.5). The emphasis on the role of theory in these connections is of course warranted. But theory dependence returns us squarely to the counterfactual problem. How do we know, after all, which theory of behavior, if any, is correct in making any particular assessment of external validity? Sometimes, when the relevant social system is simple, the problem is trivial. Other times, when experimental interventions involve complex systems, the problem is daunting if not irresolvable.

Elsewhere, Heckman writes,

The econometric framework is explicit about models of counterfactuals and how to construct them, sources of interventions, and the sources of stochastic variability and their relationship. The theory of structural estimation is based on this principle. (Heckman 2005, 6)

In the policy context, Heckman notes, reliance on counterfactualizing is particularly clear. Here, structural estimation of models allows us “to forecast the effects of interventions never previously experienced to calculate a variety of policy counterfactuals” (Heckman 2001, 10).

4.3 Axiomatic-deductive reasoning

Causal inference from non-empirical methods, such as axiomatic deductive modeling, likewise depends on counterfactualization. Heckman (2005, 2; 6; cf Schabas 2008) again provides insight:

A model is a set of possible counterfactual worlds constructed under some rules. ... Economists since the time of Haavelmo (1943, 1944) have recognized the need for precise models to construct counterfactuals and to answer “causal” questions and more general policy evaluation questions...

Standard neoclassical models, for instance, join premises to conclusions by way of a utility maximization rule, in which rational agents seek to satisfy their preferences in the context of constraints. By changing one aspect of that context, *ceteris paribus*, we can infer their reactions.¹⁶ Our demonstration that, say, a change in relative prices from their actual values causally induces a certain adjustment in demand behavior is a counterfactual exercise. Here, the antecedent and the consequent are joined by the law of demand—it connects the two in a counterfactual web that allows us to explore innumerable possible worlds (cf. Heckman 2001, 13–14; Morgan 2013).

The insurmountable epistemic problem in this approach is that there are in principle many laws (and innumerable background conditions against which they operate), which generate diverse counterfactuals. In the case of neoclassical theory, for instance, Tetlock and Belkin (1996, 26) are right to argue that its adherents “display impressive confidence in their counterfactual claims” but often fail to recognize that the “calculus of rational action is not, however, the only theoretical logic that we can use to infer what would have happened in this or that counterfactual scenario.” The behavioral economics assault on the standard neoclassical model has substantially undermined faith in the rationality proposition, but not in causality. To be able to advance causal

¹⁶ The demonstration requires the presumption that we can vary inputs in ways that “do not affect the structure of the causal relation but that affect the realized outputs” (Heckman 2001, 14). Reiss and Cartwright (2003, 24) refer to this as a restrictive “Galilean” assumption which yields an “epistemically convenient system” for ascertaining causality, but which may not illuminate most actual causal contributions in the social world. Note that the Marshallian (1961[1890]) *ceteris paribus* assumption, where we ask *what would be the effect of the cause were all else to be held constant*, necessarily transports us into hypothetical, counterfactual worlds. See Marshall (1961).

claims and be of use in the policy domain—to nudge agents to behave in the intended ways—the new behavioralists search for *predictable* patterns of irrationality so as to be able to sustain their alternative counterfactuals (Ariely 2008).

4.4 Agent-based modeling of complex adaptive systems

Computational experiments involving agent-based modeling (ABM) of complex adaptive systems represent a particularly attractive means for re-running “history” through potentially very large numbers of trials. For instance, one can begin each trial with identical sets of agents and agent properties (preferences, strategies, risk tolerance, etc.), while precisely controlling variation in a stochastic environment (enacting Pearl’s “do” operator; see Pearl and Mackenzie 2018). In comparing two scenarios, such as a policy intervention and the status quo, numerous trials can be run under the two scenarios to discover the effect of each on relevant social outcomes. Counterfactualizing is explicit in the ABM exercise. Davis (2018, 5) states the case explicitly:

Simulation methods construct artificial worlds that resemble and imitate real worlds, and consequently they compare what *could* be the case in terms of how that artificial world is microspecified with what *could* be the case in resulting macrostructural terms *were* the real world to closely resemble the artificial one. That is, reasoning in reflexivity terms involves counterfactual reasoning... .

But how dependable is any counterfactual that arises from the use of a complex adaptive model when, say, designing and assessing policy? In complexity economics, agents (and objects) are constituted dynamically, in flux. Perpetual change results from the complex entanglements that bind agents and objects—entanglements that alter both. In turn, discontinuities in the nature of agents and objects produce further systemic changes. For periods of time, economic systems can appear to approximate an equilibrium state—but then, suddenly and without warning, the system can generate severe instability owing to tipping points and phase transitions. As Elsner (2017, 941) explains, “complex adaptive systems (CAS) usually display dynamics with often abrupt transitions between some temporary order and volatile disorder, stasis alternating with turbulence and transition.”

Arthur’s (2015, 2) description of how complexity economics breaks with neoclassical theory is directly relevant to the matter of specifying determinate counterfactuals. Complexity theory, he argues, “gives us a world closer to that of political economy than to neoclassical theory, a world that is organic, evolutionary, and historically-contingent.” He adds:

Complexity economics thus sees the economy as in motion, perpetually ‘computing’ itself—perpetually constructing itself anew. Where equilibrium economics emphasizes order, determinacy, deduction, and stasis, complexity economics emphasizes contingency, indeterminacy, sense-making, and openness to change. (2015, 1)

In the complexity approach, a “solution” (or explanation) takes the form of identifying “a pattern, a set of emergent phenomena, a set of changes that may induce further changes, a set of existing entities creating novel entities.” The goal is not the discovery of universally applicable theorems “but the deep understanding of mechanisms that create these patterns and propagations of change” (2013, 19). The emphasis on non-equilibrium is central. Non-

equilibrium emerges from external shocks but also endogenously, from circumstances and processes that define the economy. Arthur emphasizes two: the deep uncertainty facing economic actors and technological change which disrupts the economy to a much greater extent than even Schumpeter recognized (Arthur 2013, 4–5).¹⁷

The problems associated with generating determinate counterfactuals in the complexity framework should be evident. Even if complexity accounts of an event are taken to be causal in nature, they do not lend themselves to the specification of a definitive counterfactual, or even a manageable set of plausible counterfactuals, since the insertion of a false antecedent into the past or future would have to be recognized as potentially inducing any number of alterations in identities and patterns of behavior, many of which might be novel. The complexity approach warns us, then, against betting heavily on any one of the possible patterns that might emerge. The problem is not that the approach is barren as concerns counterfactual generation. It is rather the reverse—it is *excessively fecund*. Its conception of agent adjustment to non-equilibrium under uncertainty generates what an impatient social scientist looking to identify a dependable causal account would view as far too many potential, plausible counterfactuals.¹⁸

The ABM-complexity strategy for managing the Fundamental Problem requires that we take flight from actual agents in actual environments, to establish idealized agents and environmental conditions for counterfactualizing that cannot obtain in any actual world. The result is a counterfactual narrative of a fantasized world. *Fantasy squared*. Applying the results of the model to the world of actual agents and environments entails a further counterfactual leap. Were the real agents and real environments just like those specified in the model, the result of the model would hold in that real world. *Fantasy cubed*. The power of the research method, then, derives not from its avoidance of counterfactual reasoning but from its imaginative embrace.

5 Causal fictions in the future and in the past

We can now understand why economists debating the minimum wage might disagree not just about prospective impacts of proposed policy but also about the realized effects of measures instituted in the past. The Fundamental Problem of Causal Inference impairs equally knowledge generation of the past and the future. When they adjudicate the matter of past effects, *they put into the arena their contesting counterfactuals—their respective fictions of what would have been but for the policy intervention* (cf Tetlock and Belkin 1996, 4). But that historical path, the path involving no increase in the minimum wage, was foreclosed the moment the policy intervention occurred. Our economists cannot ever observe the simultaneous treatment and non-treatment of the same individual or community. *The resulting ignorance is irreparable*. Given the extent of human ingenuity and imagination, and the in principle innumerable counterfactual

¹⁷ A contemporary approach that incorporates complexity into economic modeling is quantum economics, where the economy is represented as a quantum system. See Orrell (2018), which emphasizes contingency, emergence, and fluidity. The parallels with Althusser's (1969) notion of overdetermination are striking.

¹⁸ Katzenstein and Seybert's (2017) notion of the "protean power" that agents exert under conditions of uncertainty is consistent with economic complexity. For Katzenstein and Seybert, protean strategies can yield novel outcomes that cannot be anticipated in advance. See DeMartino and Grabel (2020) and the other essays in the symposium on Katzenstein and Seybert in *International Theory* (12:3).

stories that can be told and methods for generating them, we should have no trouble understanding why it is that the best minds with the best techniques in economics can and do disagree about something as apparently simple as the effects of increasing the minimum wage—its prospective effects, and even its effects in the past.

In *this* regard, then—in specifying causal relationships—there is no epistemic distinction between the past and future. As soon as we insert a counterfactual event (such as a policy intervention) into or remove an actual event from the past and ask what would have happened subsequently, we “are then telling an imagined story forward from that moment, time = T. *It is not a story that actually happened.* ... all counterfactuals are histories of the future, even when they are situated in what we normally think of as past ‘calendar’ time” (Weber 1996, 276–277).

Counterfactuals set in the future are just as temporally complex. Knight’s (2014 [1921], 202) discussion of the “two-fold inference” that informs planning for the future illuminates the issue: “We must infer what the future situation would have been without our interference, and what change will be wrought in it by our action.” In this account, present-day analysts must position themselves in the *future-future*: at a point in time when (they presume) some or all of the effects of a future policy intervention would have emerged. From that vantage point, the analysts must compare the course of history as (they presume) it would have unfolded without our interference with what (they presume) would have occurred had we counterfactually taken the action under review. This positioning requires analysts to look back in time from the future and to re-run history as it would have gone with and without the action.¹⁹

Counterintuitively, there is good reason to suspect that ascertaining causal connections among past events is in some respects more difficult than anticipating future causation. “History is a terrible teacher,” Tetlock and Belkin (1996, 9) remind us, which too often misleads:

When we look back into the past from the present, we occupy a privileged but also easily abused position. We know which one of the many futures that were once possible has actually occurred. With the benefit of this retrospective knowledge, it becomes relatively easy to find antecedents that depict the consequence as the inevitable result of some ‘inexorable’ causal process.²⁰

In theorizing the future, on the other hand, “we cannot avoid the complexity and indeterminacy of possible relationships among antecedents and consequences” (Tetlock and Belkin 1996, 9; 31, fn 4; see also Weber 1996, 275–279).²¹

¹⁹ Knight concludes on this basis that claims about the future are terribly undependable. See Hicks’ (1980, 10) related notion of “double vision” and Beckert’s (2016, 10) concept of “doubling of reality” in the creation of fictitious expectations.

²⁰ Research on hindsight and other cognitive biases substantiates concern about our ability to know the past and to make good judgments under uncertainty. See Fischhoff and Beyth 1975; the essays in Kahneman et al. (1982); and in Roese and Olson (1995). Cognitive biases are particularly likely “in the most subjective of all methods of inquiry, the counterfactual thought experiment” (Tetlock and Belkin (1996, 37).

²¹ See Hicks’ (1980, 11) distinction between the “Voluntarism” that shapes the future, and “Determinism” regarding the past. See also Hay’s (2020) critique of Hobsbawm. Weber (1996, 275–276) provides a corrective: “I want to establish that when we say we have ‘explained’ an event in past time, it does not necessarily mean that we have also ruled out the possibility that things could have been radically different than they were.”

In summary, standard causal claims in economics are fictitious in the sense that they depend upon the construction of fictitious counterfactuals. Counterfactuals are stories we narrate to discover causal mechanisms and to convince ourselves and others that we understand why things did and did not, or do and do not, or will and will not happen. In this limited but fundamental respect, the future and past are epistemically indistinct: we run up equally against the limits to knowledge in our confrontation with causal relationships in the past and the future. What allows us to check off the past as “known” is a conventional exercise, momentarily professionally sanctioned, in which researchers accept the same causal model and its associated counterfactual.

6 Criteria for constructing and assessing counterfactuals

Scholars who recognize both the prominent role of counterfactuals in causal explanation and the problems posed by the proliferation of counterfactual accounts must offer what they take to be scientific criteria for restricting the field of plausible counterfactuals (see Levy 2015). But given the nature of the problem—*knowing the unknowable*—there should be no surprise in finding that though there is overlap among the criteria on offer, there is no consensus about what counts as legitimate counterfactualizing. The various criteria are also specified and applied in distinct and even inconsistent ways. Here I survey just two principal criteria.²²

²² A reviewer for this journal asks whether, since I posit all counterfactuals as fictitious, I am to be read as claiming that all counterfactuals are equally untrue. The question is important and deserves far more attention than I can give it here. Surely there are contexts in which certain counterfactuals are far more *dependable* than others—dependable in the sense that we are warranted in relying on them when intervening in the world and in expecting that the presumed causal connections are stable and will generate the anticipated outcomes. Engineering science, for instance, generates counterfactuals that we reasonably treat as dependable in teasing out causal connections. Here the basis for confidence is the ontological simplicity of some of the relevant systems, the methodological ability of the researcher to isolate a system of concern from others so as to construct counterfactuals experimentally, and the epistemic utility of corraling distinct entities into “kinds” so that the results of an experiment on an entity that represents the kind can be taken as applicable to all other entities of the same kind. But the social sciences do not generally face these ontological, methodological, and epistemic conditions. Social scientists do not typically face simple systems—complexity is the rule. In addition, social patterns and outcomes reflect the interaction of diverse systems—some of which may be unstable—and in the social sciences, it is generally impossible methodologically to isolate the effect of one particular system from others. Speaking of the natural sciences, Cartwright (1980, 162) issues a warning that applies with particular force to the social sciences: “We have many detailed and sophisticated theories about what happens within the various domains. But we have little theory about what happens in the intersection of domains.” Finally, the designation of kinds in the social sciences—say, gender—is contingent, driven by researchers’ theoretical commitments and purposes (Fullbrook 2007). The designation requires abstracting from many features of entities that may in fact influence outcomes. It is therefore epistemically hazardous to treat one entity as representative of a kind for the purposes of establishing counterfactuals that will apply to other “similar” entities. Taken together, these features of the social world create severe epistemic limitations that hamper the discovery of dependable counterfactuals (even by experimental methods). The upshot is not that all counterfactuals are equally unconvincing. It is instead that the uniquely correct counterfactual in the social sciences is generally epistemically unavailable.

6.1 Historical consistency

Tetlock and Belkin (1996, 23ff) offer six criteria to guide counterfactualizing.²³ The list includes an intuitively appealing requirement of “historical consistency”: a counterfactual must be consistent with well-established historical facts. The criterion is widely accepted. For instance, King et al. (1994, 78) argue that “although [counterfactuals] are obviously counter to the facts, they must be reasonable and it should be possible for the counterfactual event to have occurred under precisely stated circumstances.” The criterion is implicit in the “closest possible world” frameworks of Lewis (1973) and Stalnaker (1968). In the words of Lewis (1973, 560), “a counterfactual is nonvacuously true if it takes less of a departure from actuality to make the consequent true along with the antecedent than it does to make the antecedent true without the consequent.” The criterion appears alternatively in Levy (2015) and others who articulate the “minimal re-write rule,” in the “plausible world” approach of Hawthorn (1991), and in the dynamic framework of Elster (1978, ch. 6), which he offers explicitly as a corrective to what he sees as the flawed approach of Lewis and Stalnaker. The factual consistency criterion helps to explain why contingent events like assassinations are attractive subjects for counterfactualizing since it is easy to specify the world being identical in every respect but for the fact that the assassination failed (the gun did not go off, the assassin missed the target, etc.).

The historical consistency criterion is deceptively complex. Alternative worlds may be like the actual world in some respects (e.g., sharing its laws) and unlike in other (e.g., different background conditions), as Lewis acknowledges. Deciding which of the alternative worlds is “closest” to the actual world is then a challenging, *theory-dependent* activity (Elster (1978, 187). Rigorous application of the criterion requires specification of *correct* theoretical claims that permit accurate measures of “closeness” between alternative worlds and “minimal” re-writes. But as concerns the social world, such theory remains elusive (cf. Lebow 2010, 285). One theoretical account might plausibly posit world A’ to be closest to world A, while another might plausibly hold A” to be closer still. Weber concludes that causal complexity of the social world “butchers” the minimal re-write rule (1996, 271). Heckman (2008, 16) concurs, arguing that discovery of closest possible worlds “founders on the lack of any meaningful metric or topology to measure ‘closeness’...” (cf Russo et al. (2008).

The historical consistency requirement faces a second, distinct challenge. Some particularly useful counterfactuals employ what Fearon (1996, 41ff) calls “miracle” causes (antecedents) that are explicitly inconceivable in the historical context into which they are inserted. For instance, it might be useful to theorize the trajectory of the Soviet Union had Trotsky somehow displaced Stalin, even though we may have good reason to believe “that Trotsky could never have been in Stalin’s position” (Lukes 1980, 149). Hume usefully employs a miracle cause in a “thought experiment” that probes money neutrality. Schabas (2008, 161) provides this passage from Hume’s *Essays* to make the point:

²³ In addition to the criteria I explore in the text the list also includes cotenability, clarity, statistical consistency, and projectability (18ff).

For suppose, that, by miracle, every man in GREAT BRITAIN should have five pounds slipt into his pocket in one night; this would much more than double the whole money that is at present in the kingdom; yet there would not next day, nor for some time, be any more lenders, nor any variation in the interest [rate]. (Hume, 1985, p. 299)²⁴

Miracle causes appear not just in explicitly fanciful speculations, like Hume's, but also within the most prosaic research methods. A notable example is large- n statistical research that seeks to explore nomothetic (generalizing) causes or "the logical implications of a theoretical framework" (Levy 2015, 383). As Fearon (1996, 56) points out, "Authors estimating causal effects using regression analysis on nonexperimental data *never* ask whether it would have been actually possible for each case in the sample to have assumed different values on the independent variable." We do not demand of statistical studies plausible stories about whether or how different values on the independent variable could have come about. We accept that the historical-causal chain of events that generated $X = x_i$ could be broken miraculously to set it equal to some other value—just like five pounds being slipt into every man's pocket. We then infer the likely effect of the change in x_i on the dependent variable y_i .²⁵

Miracle causes are often legitimate and, as in the case of large- n studies, inescapable. But they nevertheless raise a perplexing methodological question: just what limits should be placed on miracle causes, if any? The scenario, *had Napoleon's arsenal included Stealth bombers he would have prevailed at Waterloo* (Fearon 1996), is sometimes cited by critics as an example of the problem of the undisciplined nature of miracle counterfactuals, and we should concede the point. On the questions, then, when are miracle causes useful, and what kinds of miracle causes are permissible and useful, there is dissensus among counterfactual methodologists about the standards such causes must meet to be taken seriously and to generate useful insights.

6.2 Theoretical consistency

For present purposes, a second criterion, already suggested above, is most important and yet inescapably fraught. It is the requirement of "consistency with well-established theoretical laws" (Tetlock and Belkin 1996, 25; see Heckman 2008, 3-4; Schabas 2008; Reiss and Cartwright 2003). This criterion is particularly problematic in the social sciences where such laws may be missing, contested, or minimally, paradigm specific: "[D]ifferent schools of thought in world politics [and, I would add, economics] have very different notions of what constitutes a theory or law or what is 'well established'" (Weber 1996, 285). Tetlock and Belkin (1996) cast doubt on the criterion even as they offer it: "Consistency with *well-established* theory is a reasonable standard for gauging the plausibility of counterfactuals but we should expect disagreement about what counts as well-established theory in world politics" (27; see also Levy 2015).

²⁴ A thought experiment "commence[s] with a jarring counterfactual that transports the mind to a different and distant world, as opposed to a proximate alternative world" (Schabas 2008, 62).

²⁵ In doing so we also assume that breaking the historical-causal chain to set x_i to an alternative value does not alter any other causal relations. See Reiss and Cartwright's (2003) examination of Pearl's semantics in this connection.

Why is well-established theoretical law so rare in economics (and other social sciences)? Why have theoretical controversies between neoclassicals, Austrians, Marxists, feminists, Post-Keynesians, institutionalists, and advocates of other theoretical frameworks by now not been resolved empirically? The most common answer is that theoretical dispute reflects the absence of scientific objectivity that prevents consensus around the right framework. Partisans complain that their opponents are prevented from seeing the evidence right before their eyes by a tangle of unacknowledged ideologies, interests, or incompetence. But the allegation is *ad hominem* and epistemically naive. A more satisfactory account recognizes that theoretical adversaries hold to distinct, theoretically driven, fictitious counterfactual narratives—and *the world cannot ever tell us definitively which fiction (if any) is correct*. Holland's warning is pertinent: world events unfold just once. Economists do not get to run the tape n times, alternately treating and not-treating identical agents, to ferret out causal relationships. They can presume that they can simulate multiple re-runs with agent-based models, RCTs, or large- n studies, but each of the i 's that sum to the n are one-time events, occurring under innumerable unique, epistemically unavailable conditions that bear on outcomes.

There is an inescapable epistemic conundrum regarding counterfactualizing that takes the form of a circularity problem. *Generating the right counterfactuals requires the right causal theory, while establishing the right causal theory requires the right counterfactual*. Weber (1996, 272) makes the point while problematizing the minimal re-write rule: "once we knew what 'minimal' meant well enough to validate counterfactuals, we would no longer need to use counterfactuals because we would already know what we want to know and understand what we want to understand."

The circularity problem can be sidestepped but not resolved. It is sidestepped by a *convention* (Heckman 2008) that a particular theoretical framework is correct (and/or that a particular method is legitimate), so that only its associated counterfactuals are appropriate when theorizing causal connections. But this requires that we presume what it is we thought we were testing when we do empirical work—the adequacy of causal theory. Only if we are prepared to presume *ex ante* the unambiguous superiority of one theory over alternatives can we presume that its associated counterfactuals are uniquely correct. If we reject that presumption, as we should, then we are led to see that theoretical pluralism is an inevitable feature of social science—not a pathology (cf Fullbrook 2007). It is inevitable because *the fight is over what is not seen and cannot ever be seen*. The fight is over the features of and outcomes in a fictitious world—and there are no secure, scientifically unimpeachable means for sorting out which if any of the many hypotheses about that world is correct. How, after all, is one to do knock-out, conclusive empirical research of worlds that do not exist?

7 The virtues of counterfactualizing

The epistemic iconoclasts are right to emphasize the limitations to what economists can know about the future. I have attempted to show that economists face irreparable ignorance equally in knowing causal processes situated in the past and the future. The problem stems from the irreducibly fictitious nature of counterfactuals. With Fischhoff (1982, 335), "we should hold the past in a little more respect when we attempt to plumb its secrets."

The speculative freedom provided by counterfactual analysis, and its legitimizing of theoretical pluralism, is a bane to economists seeking to lockdown causality. But the freedom presents invaluable service to society if we know how to exploit it. Used properly, counterfactualizing promotes scenario thinking that helps us prepare for and adapt to an unknowable world, where things might go very differently from what established theory leads us to expect, so that we can begin to imagine how we might adjust. The approach is at the heart of what is called “decision-making under deep uncertainty” (DMDU) which is practiced by scholars in many domains and especially by those confronting climate-related policy questions. Under DMDU experts explicitly reject the idea that policy-making requires and is well-served by determinate or probabilistic predictions of policy effects. Instead, the approach explores thousands to hundreds of thousands of plausible scenarios and then searches for “robust” policy, policy that appears to do well enough under a very wide range of possible futures (see Marchau et al. 2019). Rather than optimize on outcomes, which requires dependable causal knowledge of the past, present, and future that we typically do not have, the approach looks to maximize robustness. The epistemic grounding in uncertainty requires that DMDU decision-making be adaptive in response to new knowledge about the relevant systems that can only emerge over time. Decision-makers must monitor the systems and adjust policy as circumstances evolve. It bears emphasis that DMDU practitioners recognize counterfactual scenarios as invaluable instruments—*not* for proving causality but for managing better in a world that presents both opportunities and dangers that we cannot ultimately know and that we can influence but never control.²⁶

DMDU entails humility on the part of economic experts and inclusion in policy design and implementation of critical stakeholders who will bear the consequences of policy interventions. In economics those who have historically understood the epistemic problem have also challenged the view of the prescient economist with the capacity to exert sufficient control in order to engineer optimal outcomes. Hayek, McCloskey, and Easterly are representative of the classical liberal tradition in this regard. They are joined by most of the epistemic skeptics on the right and the left cited above (see fn 1). In Hayekian terms, we might say that the project looks to bring into policy design, implementation, and adaptation the opportunities for experimentation and learning that the classical liberals associate with market competition.²⁷

Good social scientific counterfactuals can prepare us for the unknown—for surprise, discontinuities, and tipping points that generate new and unpredicted trajectories. Weber (1996, 268) conveys the point: “counterfactuals can be used to open minds, to raise tough questions about what we think we know, and to suggest unfamiliar or uncomfortable arguments that we had best consider.” He continues: “The diversity of ideas is a survival asset for a human society living in an uncertain environment with an uncertain future (273).²⁸

²⁶ DMDU philosophies and strategies are being developed by applied economists at the RAND Corporation, the World Bank, and beyond. See the website of the Society for Decision-Making Under Deep Uncertainty for an introduction to the field and links to relevant work: <http://www.deepuncertainty.org/>.

²⁷ See Witt (2003) for an evolutionary approach to policy formation that engages the implications of epistemic insufficiency and learning by economic actors and policy experts.

²⁸ I am concluding this paper as COVID-19 sweeps across a world that is extraordinarily unprepared in even the most basic ways to deal with it. I conjecture, *counterfactually*, that had social scientists better prepared policy-makers to understand the value of counterfactual exercises in imagining alternative futures in policy design, the world would have been much better prepared for the pandemic. Public health and climate change scholars are far ahead of economists in this regard.

Recognition of the speculative value of counterfactualizing provides the grounding for a non-apologetic defense of theoretical pluralism. Pluralism has been resisted in economics at least in part, I submit, because alternative theoretical perspectives cast doubt on the truth content of the counterfactual scenarios generated by economic orthodoxy and, as a consequence, challenge orthodox causal claims. But that is precisely the virtue of contending theoretical perspectives in economics. They serve to generate alternative plausible causal linkages that are missed when a profession assembles within one particular church and professes the truth of its sacred texts. Convergence around one theoretical approach generates unwarranted overconfidence in theoretical propositions and empirical inferences, suppresses recognition of alternative worlds, and restricts the proliferation of alternative scenarios that just might prepare us for unwelcome futures.²⁹

The significance of the first epigraph to this paper should now be apparent. The opening lines to *Nothing Twice*, by Nobel Laureate Wislawa Szymborska, provide a warning that illuminates the epistemic problem confronting economists.³⁰ In economics, “nothing can ever happen twice,” let alone $n > 2$ times. Our ability to navigate responsibly in the context of irreparable ignorance, “without the chance to practice,” depends on our abilities to counterfactualize creatively, pushing back against the professional and psychological drive for theoretical closure and causal finality.

Acknowledgments Katie Aldrich, Abby Brown, Lauren Craig, Joe Downes, Sophia Gonzalez-Mayagoitia, Quentin Good, Amanda Hayden, Tasia Poinsatte, and especially Holden Fitzgerald, provided wonderful research assistance on this project. Thanks to Ilene Grabel, Marco Nathan, and Elizabeth Ramey for comments on earlier drafts; to many members of the Societies for the History of Economics (SHOE) (including among others, Spencer Banzhaf, John Davis, Claude Diebolt, Mirus Fitzner, Menno Rol, Eric Schliesser, Mark Silverman, Brian Simboli, and Adrian Yee) for helpful leads into the literature; and to Dogus Aktan, JJ Fancone, Jack Donnelly, Rob Garnett, Stefano Guzzini, Peter Katzenstein, Louis Pascarella, Zoe Sherman, and Kai Thaler for thoughtful suggestions. I am particularly indebted to the extraordinarily generous referees for this journal.

References

- Abbott B (1974) Some problems in giving an adequate model-theoretic account of cause. In: Fillmore C, Lakoff G, Lakoff R (eds) *Berkeley Studies in Syntax and Semantics*, vol I. University of California, Berkeley, pp I-1–I-14
- Althusser L (1969) *For Marx*. Penguin, London
- Ariely D (2008) *Predictably Irrational*. Harper Collins, New York
- Arthur WB (2013) Complexity economics: a different framework for economic thought. SFI Working Paper #13-04-012, Santa Fe Institute. Available at: <https://www.santafe.edu/research/results/working-papers/complexity-economics-a-different-framework-for-eco>. Accessed 24 Nov 2020
- Arthur WB (2015) *Complexity and the economy*. New York, Oxford
- Beckert J (2016) *Imagined futures*. Harvard University, Cambridge
- Cartwright N (1980) The truth doesn't explain much. *Am Philos Q* 17(2):159–163
- Cartwright N (2004) Causation: one word, many things. *Philos Sci* 71(5):805–819. <https://doi.org/10.1086/426771>
- Cartwright N (2007) *Hunting causes and using them*. Cambridge University, Cambridge

²⁹ The point was made forcefully by leading economists to explain the failure of the economics profession to anticipate even the possibility of the financial crisis of 2008. See DeMartino (2011b).

³⁰ Thanks to Zoe Sherman for bringing this poem to my attention.

- Colander D (2005) What economists teach and what economists do. *J Econ Educ* 36(3):249–260
- Crotty JR (2019) *Keynes Against Capitalism*. Routledge, New York
- Davis JB (2018) Agent-based modeling's open methodology approach: simulation, reflexivity, and abduction. *Economia* 8(4):509–529
- Deaton A (2009) Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development. NBER working paper No. 14690 (revised December 2011)
- Deaton A, Cartwright N (2018) Understanding and misunderstanding randomized controlled trials. *Soc Sci Med* 210:2–21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- DeMartino GF (1993a) Beneath 'first principles': controversies within the new macroeconomics. *J Econ Issues* 27(4):1127–1153
- DeMartino GF (1993b) The necessity/contingency dualism in Marxian crisis theory: the case of long wave theory. *Rev Radic Politic Econ* 25(3):68–74
- DeMartino GF (2011a) *The Economist's oath*. New York, Oxford
- DeMartino GF (2011b) The economic crisis and the crisis in economics. In: Starr M (ed) *Consequences of economic downturn*. Palgrave, New York, pp 25–44
- DeMartino GF (2013) Epistemic aspects of economic practice and the need for professional economic ethics. *Rev Soc Econ* 71(2):166–186
- DeMartino GF, Grabel I (2020) Irreparable ignorance, protean power, and economics. *Int Theory* 12(3):435–448
- Duflo E, Glennerster R, Kremer M (2006) Using randomization in development economics research: a toolkit, NBER Technical Working Paper 333. <http://www.nber.org/papers/w0333>. Accessed 24 Nov 2020
- Easterly W (2006) *The white man's burden*. Penguin, New York
- Elsner W (2017) Complexity economics as heterodoxy: theory and policy. *J Econ Issues* 51(4):939–978. <https://doi.org/10.1080/00213624.2017.1391570>
- Elster J (1978) *Logic and society*. John Wiley & Sons, New York
- Fearon JD (1996) Causes and counterfactuals in social science: exploring an analogy between cellular automata and historical processes. In: Tetlock PE, Belkin A (eds) *Counterfactual thought experiments in world politics*. Princeton University, Princeton, pp 39–68
- Fischhoff B (1982) For those condemned to study the past: heuristics and biases in hindsight. In: Kahneman D, Slovic P, Tversky A (eds) *Judgment under uncertainty*. Cambridge University, Cambridge, pp 335–351
- Fischhoff B, Beyth R (1975) I knew it would happen: remembered probabilities of once-future things. *Organ Behav Hum Perform* 13(1):1–16. [https://doi.org/10.1016/0030-5073\(75\)90002-1](https://doi.org/10.1016/0030-5073(75)90002-1)
- Fogel RW (1964) *Railroads and American economic growth*. Johns Hopkins, Baltimore
- Fullbrook E (2007) Narrative pluralism. post-autistic economics review No. 42: 22–43. <http://www.paecon.net/PAERReview/issue42/Fullbrook42.htm>. Accessed 24 Nov 2020
- Goertz G, Levy JS (2007) Causal explanation, necessary conditions, and case studies. In: Goertz G, Levy JS (eds) *Explaining war and peace*. Routledge, New York, pp 9–45
- Goldin C (1973) The economics of emancipation. *J Econ Hist* 33(1):66–85
- Goodman N (1947) The problem of counterfactual conditionals. *J Philos* 44(5):113–128
- Grabel I (2017) *When things don't fall apart*. MIT, Cambridge
- 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(suppl.):iii–vi; 1–115
- Hawthorn G (1991) *Plausible worlds*. Cambridge University, Cambridge
- Hay C (2020) Brexistential angst and the paradoxes of populism: on the contingency, predictability and intelligibility of seismic shifts. *Polit Stud* 68(1):187–206
- Hayek F (1974) *The Pretense of Knowledge*. Nobel Prize Speech, Dec 11. Available at: <https://www.nobelprize.org/prizes/economic-sciences/1974/hayek/lecture/>. Accessed 24 November 2020
- Heckman JJ (2001) *Econometrics, counterfactuals and causal models*. Keynote Address, International Statistical Institute, Seoul
- Heckman JJ (2005) The scientific model of causality. *Sociol Methodol* 35:1–97. <https://doi.org/10.1111/j.0081-1750.2006.00164.x>
- Heckman JJ (2008) *Econometric causality*, IZA Discussion Paper No. 3425. <https://ideas.repec.org/p/iza/izadps/dp3425.html>
- Hicks J (1980) *Causality in economics*. Australian National University, Canberra
- Hirschman AO (2013) [1970] The search for paradigms as a hindrance to understanding. In: Adelman J (ed) *The essential Hirschman*. Princeton University, Princeton, pp 137–154
- Holland PW (1986) Statistics and causal inference. *J Am Stat Assoc* 81(396):945–960

- Hume D (1985) [1741-1777] *Essays, moral, political and literary*. In: Miller EF (ed) *liberty classics*, Indianapolis
- Hume D (2007) [1748] *An enquiry concerning human understanding*. New York: edited with an introduction and notes by Peter Millican. Oxford University, New York
- Jevons WS (1882) *Elementary lessons in logic*. MacMillan, London
- Kahneman D, Slovic P, Tversky A (eds) (1982) *Judgment under uncertainty*. Cambridge University, Cambridge
- Katzenstein PJ, Seybert LA (eds) (2017) *Protean power: exploring the uncertain and unexpected in world politics*. Cambridge University, Cambridge
- Keynes JM (1936) [1973] *The general theory of employment, interest, and money*. London: Macmillan
- Keynes JM (1937) *The general theory of employment*. *Q J Econ* 51(2):209–223
- King G, Keohane RO, Verba S (1994) *Designing social inquiry*. Princeton University, Princeton
- Knight FH (2014) [1921] *Risk, uncertainty and profit*. Cambridge University, Cambridge
- Lawson T (2009) *On the nature and roles of formalism in economics*. In: Fullbrook E (ed) *Ontology and economics: Tony Lawson and his critics*. Routledge, London and New York, pp 189–231
- Leamer EE (2009) *Macroeconomic patterns and stories: a guide for MBAs*. Springer-Verlag, Berlin
- Lebow RN (2010) *Forbidden Fruit*. Princeton University, Princeton
- Levy JS (2015) *Counterfactuals, causal inference, and historical analysis*. *Secur Stud* 24(3):378–402
- Lewis D (1973) *Causation*. *J Philos* 70(17):556–567
- Lukes S (1980) *Elster on counterfactuals*. *Inquiry* 23:145–155
- Mackie J (1965) *Causes and conditions*. *Am Philos Q* 2(4):245–264
- Mahoney J, Barrenechea R (2017) *The logic of counterfactual analysis in case-study explanation*. *Br J Sociol* 70(1):306–338
- Marchau VAWJ et al (eds) (2019) *Decision making under deep uncertainty*. Springer, Cham
- Marshall A (1961) [1890] *Principles of economics*, 9th edn. Macmillan, New York
- McCloskey DN (1990) *If you're so smart: the narrative of economic expertise*. Chicago University, Chicago
- McCloskey DN (1998) *The rhetoric of economics*. University of Wisconsin, Madison
- Mill JS (1843 [2002]) *A system of logic*. University Press of the Pacific, Honolulu
- Morgan MS (2013) *What if? Models, Facts and fiction in economics - Keynes lecture*. Retrieved from <https://www.youtube.com/watch?v=c-tsoGmOlfq>. Accessed 24 Nov 2020
- Morgan SL, Winship C (2007) *Counterfactuals and causal inference*. Cambridge University, Cambridge
- Nathan M (2017) *Counterfactual reasoning in molecular medicine*. In: Boniolo G, Nathan M (eds) *Philosophy of molecular medicine: foundational issues in research and practice*. Routledge, New York, pp 192–214
- Nelson J (2004) *Clocks, creation and clarity: insights on ethics and economics from a feminist perspective*. *Ethical Theory Moral Pract* 7:381–398
- Orrell D (2018) *Quantum economics*. Icon Books, London
- Pearl J, Mackenzie D (2018) *The book of why*. Basic Books, New York
- Reiss J, Cartwright N (2003) *Uncertainty in econometrics: evaluating policy counterfactuals*. Reprinted in: In: Reiss J (ed) *Error in economics: towards a more evidence based methodology*. Abingdon. Routledge, UK
- Resnick SA, Wolff RD (1987) *Knowledge and class*. University of Chicago, Chicago
- Robbins LC (1927) *Mr. Hawtrey on the scope of economics*. *Economica* 20:172–178
- Rodrik D (2017) *Straight talk on trade*. Princeton University, Princeton
- Roese NJ, Olson JM (eds) (1995) *What might have been*. Lawrence Erlbaum, Mahwah
- Ruccio DF (2005) *(Un)real criticism*. *Post-autistic economics review* (35):40–49
- Ruccio DF, Amariglio J (2003) *Postmodern moments in modern economics*. Princeton University, Princeton
- Russo F, Wunsch G, Mouchart M (2008) *Potential outcomes, counterfactuals, and structural modelling: causal approaches in the social sciences*. STAT Discussion Paper 0826. <http://hdl.handle.net/2078.1/126559>
- Samuels WJ (1993) *In (limited but affirmative) defense of nihilism*. *Rev Polit Econ* 5(2):236–244
- Schabas M (2008) *Hume's monetary thought experiments*. *Stud Hist Phil Sci* 39:161–169
- Schliesser E (2017) *The sympathetic process and judgments of propriety*. Oxford University, New York
- Shackle GLS (1992) [1972] *Epistemics and economics*. Routledge, New York
- Simon HA (1952) *On the definition of the causal relation*. *J Philos* 49(16):517–528
- Skyrms B (1980) *Causal necessity*. Yale University Press, New Haven
- Stalnaker RC (1968) *A theory of conditionals*. In: Rescher N (ed) *Studies in logical theory*. Blackwell, Oxford, pp 98–112
- Sugden R (2000) *Credible worlds: the status of theoretical models in economics*. *J Econ Methodol* 7(1):1–31
- Szymborska W (1998) *Nothing twice*. In: *Poems new and collected: 1957–1997*. Harcourt Brace, New York
- Taleb NN (2010) *The black swan*, 2nd edn. Random House, New York

- Tetlock PE, Belkin A (1996) Counterfactual thought experiments in world politics. In: Tetlock PE, Belkin A (eds) Counterfactual thought experiments in world politics. Princeton University, Princeton, pp 3–38
- Thompson WR (2007) Powderkegs, sparks, and world war I. In: Goertz G, Levy JS (eds) Explaining war and peace. Routledge, New York, pp 113–145
- Weber S (1996) Counterfactuals, past and future. In: Tetlock PE, Belkin A (eds) Counterfactual thought experiments in world politics. Princeton University, Princeton, pp 268–288
- Witt U (2003) Economic policy making in evolutionary perspective. *J Evol Econ* 13:77–94