DEBATES

The Discipline and They: Notes on Counterfactual Methodology and the « New » Economic History *

Stefano Fenoaltea
Amherst College

I.

The economic historian's analysis of a particular body of past reality may be said to involve three fundamental moments: first, the construction of a model, i.e., the choice of a set of assumptions that provide a manageable summary of that reality; second, the solution of the model in terms of some particularly significant variables, i.e., the deduction of some particularly interesting implications of the assumptions; and third, the interpretation of the model's solution, i.e., the drawing of conclusions attributed not to the model but to reality. The salient weakness of the « old » economic history appeared to be in the second (and most narrowly analytical) of these moments; a generation ago, the invidious distinction implied by Robbins's *obiter dictum* [1] reflected an opinion that was certainly common, and perhaps justified. If that opinion now appears *passée*, it is no doubt thanks to the analytical rigor stressed by the « new » economic history, and the consequent influx into the discipline of increasingly sophisticated economic theory. Largely by explicitly analyzing counterfactual hypotheses — in the teeth of traditional historiographical proscriptions — the « new » economic history has immensely reduced the errors and *non-sequiturs* in the logical deduction of the implications of the chosen assumptions. This development can only be applauded; but it covers only the second of those three moments of historical analysis, and one

(*) This paper owes much to the comments and suggestions of Michael Edelstein, James Garson, John Lambelet, Stephen Ross, Lawrence Schofer, and Richard Sylla.

[1] L. ROBBINS, *The Economic Basis of Class Conflict* (London: Macmillan, 1939), p. 9.

is struck by the lack of parallel progress (if not actual retrogression) in the first and third. To the extent, at least, that the « new » economic history is represented by some of its most celebrated products, its answers to historical questions have often proved far from satisfactory, either because the model ill fits the historical reality, or because the experiment the model was used for ill fits the problem at hand. In either case, the misjudgment appears to reflect an over-hasty borrowing of common contemporary economic models and methods; these weaknesses in the construction of historical models and the interpretation of their solutions may thus be as characteristic of the genre as its strength in formal manipulation. The first of these twin weaknesses is examined in section III below: the problem is empirical and historical, and brings into question the very meaning of economic history. The second, to which we now turn, is rooted instead in logic and language, and informs the vexed relationship of causes and counterfacts.

II.

Not long ago, Alexander Gerschenkron — speaking as president of the Economic History Association — declared his dissatisfaction with the « new » economic historians' understanding of the relationship of causal statements and counterfactual analysis.[2] He chose to challenge the belief that causal statements imply counterfactual propositions; the choice appears unfortunate, and the challenge unsuccessful. Gerschenkron's purported example of a non-counterfactual verification of a causal statement (« The bear died, because I shot it through the heart ») is indeed nothing of the sort. The problem of establishing whether or not two facts (1. the bear died; 2. I shot the bear through the heart) are related as effect to cause does not begin to be solved by either of his « factual » techniques. By the first of these, one establishes that the bear was not shot at all (and *a fortiori* not through the heart or by me): one of the supposed facts turns out not to be a fact after all, so that there can no longer be any question of a causal relation between the two, and the problem is effectively assumed away. By the other, one establishes that the bear died from some other cause: a causal relationship is assumed to have been established between two facts (1. the bear died; 2. e.g., the bear was struck by lightning), and the problem at hand (or, given the change in one of the two facts, a problem formally identical to the one at hand) is assumed to have already been solved. The particular — now common — opinion that Gerschenkron attacked thus remains uncontroverted: clearly, the nature of

---

[2] A. GERSCHENKRON, *The Discipline and I*, «Journal of Economic History», XXVII (December 1967), p. 456; the relevant passage was reprinted, with slight amplifications, in A. GERSCHENKRON, *Continuity in History and Other Essays* (Cambridge: Belknap, 1968), pp. 53-54. The quoted phrases are from the later source.

any relation between particular facts or situations is established by means of controlled experiments, which compare results obtained under carefully differentiated sets of conditions; equally clearly, the economic historian can neither expect to find all those exact sets of conditions in historical experience, nor create them in actual fact as a natural scientist might in his laboratory. The factual experience of history provides *one* relevant set of conditions and results; but the other sets are hypothetical, and represent not facts but counterfacts. The « new » economic historian's claim that any assertion of historical causality is at least implicitly counterfactual means, simply, that the empirical basis for such a statement is experimental, and that the historian's experiments are thought-experiments; as such, the claim is indubitably justified. For all that, the instinct at the root of Gerschenkron's dissatisfaction appears unexceptionable, and his fire no more than misdirected: causal relations are indeed hardly captured by the « new » economic historian's counterfactual paradigm, and the inappropriateness to the historical question ostensibly under investigation of the counterfactual experiment with which « counterfactual history » burst upon the scene bears no little measure of responsibility for the ill-grace with which the product was received.

The bare bones of this criticism may be summarized in the following propositions: (1) there exist two separate patterns of relationships between events or situations, say « necessity » on the one hand and « sufficiency » on the other; (2) the standard form of the counterfactual historical experiment reveals « necessity »; (3) historical importance, causation, etc., involve « sufficiency »; (4) *ergo*, the standard experiment fails to document historical importance, causation, etc. (The positive conclusion is (5) *ergo*, the standard experiment should be modified so as to reveal « sufficiency » and document historical importance, causation, etc.) The precise interpretation of these propositions will be developed forthwith; while the general analysis is perforce relatively abstract, it will be related as far as possible to very familiar material derived from Robert Fogel's path-breaking analysis of the importance of nineteenth-century American railroads.[3] In essence, Fogel's measure of that importance was obtained by comparing actual production (national or regional income) to what it would have been in the absence of the railroad, behavior patterns (including of course the reaction to investment opportunities) and non-railroad opportunities (including of course opportunities for non-railroad transport investment) being held as they were.[4] The intuitive basis of the methodological weakness of this procedure is perhaps best brought out by its application to this caricature of the empirical case Fogel dealt with: an economy producing some given aggregate income NI^*, with

[3] R. W. Fogel, *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore: Johns Hopkins, 1964).

[4] This is a purposely simplified idealization of Fogel's approach; the appropriate qualifications are introduced below.

much internal transportation which was actually provided exclusively by railroads, and which could have been provided — at an infinitesimal increase in cost — by a system of canals, none of which was in fact ever constructed. A Fogelian experiment reveals that the railroad did not matter at all: given the alternative provided by canals, income would have remained (infinitesimally close to) *NI** even if railway technology had not been available, and the effect of the railroad is found to be (infinitesimally close to) nil. By the same token, of course, canals were not important either: given the (plu)-perfect substitute provided by the railroad, income would have remained *NI** even if canal technology had not been available. What does turn out to have been important is cheap transportation, as in the absence of both railway and canal technology national income would have been lower than *NI** by a conspicuous amount ΔNI. This last conclusion is quite satisfactory; but one may object to the first on a variety of grounds. It seems gratuitously paradoxical to argue that cheap transportation *was* important (as the cause of ΔNI), but that everything that actually provided that cheap transportation *was not*; and it seems downright silly to say that the railroads which *were* built were no more important as a *historical* cause (of the ΔNI attributed to cheap transportation) than the canals which *were not* built. The railroad may not have been necessary for the cheap transportation that kept income at *NI**, rather than below it by ΔNI; but it was sufficient, and was the only sufficient means actually in use. One may be free to define terms in unusual ways; but the common use of the language ties empirical « causality » and « importance » neither to necessity nor to potential sufficiency, but to actual sufficiency, and the Fogelian experiment does not begin to measure the railroad's effect on national income — its importance — in the common conception of the terms.

Causality involves more than just necessity or sufficiency; indeed, its empirical attribution (about which more below) involves more than rigorous logic, and is correspondingly difficult to handle *in vacuo*. Necessity and sufficiency are instead rigorous concepts, and thus provide a solid starting point; but their strict logic must be carefully spelled out, for necessity and sufficiency are related at a variety of levels. (a) If y exists only if x exists, y does not exist if x does not exist, x exists if y exists, and x does not exist only if y does not exist: if x is a necessary (but not sufficient) condition for y, *non-x* is a sufficient (but not necessary) condition for *non-y*, y is a sufficient (but not necessary) condition for x, and *non-y* is a necessary (but not sufficient) condition for *non-x*. (b) Alternatively, if y exists if x exists, y does not exist only if x does not exist, x exists only if y exists, and x does not exist if y does not exist: if x is a sufficient (but not necessary) condition for y, *non-x* is a necessary (but not sufficient) condition for *non-y*, y is a necessary (but not sufficient) condition for x, and *non-y* is a sufficient (but not necessary) condition for *non-x*. Each of these two separate complexes of relationships

(a) and (b) thus involves both necessity and sufficiency, in twin symmetric pairs. On the assumption that y in fact depends on x (« causality runs from x to y »),[5] rather than vice-versa, the first pair in both (a) and (b) is *real* (i.e., deals with conditions of existence), while the second pair is *cognitive* (i.e., deals with conditions of knowledge): for instance, the existence of trucks is (really) necessary for the existence of trucking, the non-existence of trucks is (really) sufficient for the non-existence of trucking, the existence of trucking is (cognitively) sufficient for (i.e., the existence of trucking proves) the existence of trucks, and the non-existence of trucking is (cognitively) necessary for (i.e., the *non*-non-existence of trucking would *dis*prove) the non-existence of trucks. One relation in each pair, furthermore, deals with the variables' existence (x, y), the other with their non-existence (*non-x*, *non-y*); mnemonically, presence is probably superior to absence, and real relations superior to cognitive ones, so that these two complexes of four relations may be conveniently identified by the real relationship between the variables' existence. Complex (a) may thus conveniently be labelled « necessity », and complex (b) « sufficiency », as suggested in proposition (1) above. Of course, two variables may be related simultaneously by both (a) *and* (b) (y exists if and only if x exists, etc.); but while « necessity » and « sufficiency » may thus be combined, what matters for the moment is that they need not be.

Indeed, the standard Fogelian experiment reveals « necessity », and not « sufficiency ». The attractiveness of that experiment — at least to economists — is that it appears to do exactly what a development economist would do to evaluate the effect, say on national income, of a particular investment project, say railroad construction: he would measure the difference in national income with and without the railways, other things being held as they are.[6] The calculations of the economic historian and of the development economist appear formally identical; but they are substantively very different. In both cases, a change in a given variable is assumed, and the resulting change in another variable is estimated; the latter change is of course that for which the former is sufficient (under the conditions assumed by the experiment). The development economist, however, works in prospect, while the economic historian works in retrospect; the one adds the railway, and measures the income rise its addition is sufficient for, while the other removes the railway, and measures the income drop its removal is sufficient for. The relation of (conditional) sufficiency is thus between the *presence* of the railway and of incremental income in the one case, and between the *absence* of the railway and of incremental income in the other. If x is sufficient for y, we are in the presence of « sufficiency »; if *non-x* is sufficient for *non-y*, x is necessary for y,

[5] On the direction of causal relations, see H. A. Simon, *Models of Man* (New York: Wiley, 1957), ch. 3.

[6] See S. Lebergott, *United States Transport Advance and Externalities*, « Journal of Economic History », XXVI (December 1966), p. 437.

and we are in the presence of « necessity ». The difference between measuring marginal contributions (*ceteris paribus*) in prospect and in retrospect is thus the difference between « sufficiency » and « necessity »; in prospect and in retrospect, the « same » counterfactual experiment will produce different results because it captures two fundamentally different relationships. On a Fogelian (retrospective) calculation, as noted, the contribution of railroad technology in general is limited to its margin of superiority over the best alternative (canals), and that margin may be trifling indeed; on the same basis, no specific, individual railroad system could be considered to contribute anything at all, as the best alternative is at worst a marginally different individual railroad system (say one occupying a barely different right-of-way, or built to a barely different gauge) that is all but indistinguishable from, and thus obviously a perfect substitute for, the system under consideration. The development economist's calculation of the (prospective) contribution of any one project is, instead, independent of its possible substitutes: the prospective increase in national income attributed to the creation of any particular railway system is invariant to the possibility of building any different railway system or any system of canals, and so on (with the result that the calculated contributions of different individual systems are positive and generally different, rather than uniformly zero, and therefore meaningfully comparable). « Necessity » turns on the non-availability of substitutes; « sufficiency » does not.[7]

It is of course « sufficiency », and not « necessity », that is at the root of the common notions of historical importance, causation, and the like. Effects are what causes are sufficient for; concrete historical effects are attributed to concrete historical causes, even though perfect substitutes for the latter obviously abound (prosperity is attributed to that particular railway network, even though a slightly different one would have achieved exactly the same results). To be sure, Fogel's reconstruction of alternative transport investment opportunities is a historiographical contribution of the first magnitude; and he was perfectly aware that his experiment revealed « necessity », i.e., that it tested what he called « the axiom of indispensability ». He did not, however, appear to realize that the development economist's *ceteris paribus* experiment which he (ideally) imitated in form but not in substance reveals not « necessity » but « sufficiency » — and that it is « sufficiency » with which historical « importance » is generally identified. Indeed, of the very statements which Fogel quoted (1964, pp. 1-9) to define the issues he was ultimately concerned with — to wit, the « importance » of

[7] The difference in the results of the standard prospective and retrospective experiments is thus based on their very logic, and does not require such gimmicks as the irreversible economies of scale (through learning-by-doing and the like) suggested by Paul David, *Transport Innovation and Economic Growth: Professor Fogel On and Off the Rails*, « Economic History Review », XXII (December 1969), pp. 514-519.

the railroad (to the aggregate level of production) — many clearly refer not to « necessity » at all, but to « sufficiency ». This is true of the earliest ones, as the 1847 paragraph by J. W. Scott (p. 3), and the many statements from the 1867 *North American Review* (pp. 3-4); it is true of later ones, as those by C. W. Davis in 1891 (p. 7); and it is true of the latest ones, as the words of C. W. Wright in the 1940s, and A. C. Bolino in 1961 (p. 9). Of the other statements on the importance of the railroad, some are ambiguous (e.g., that by S. M. Peto, p. 5, or H. Krooss, p. 9), and some irrelevant (e.g., those concerned with the railroads' power to set largely arbitrary rate structures, quoted from various sources, pp. 6-7, or that concerned with locational distribution rather than aggregate production levels, quoted from G. C. Fite and J. E. Reese, p. 9); in fact, the *only* source which gave Fogel extended passages which clearly argue for the railroad's « necessity » was a polemical article written in 1891 by Sidney Dillon (quoted first on pp. 5-6 and again, for good measure, on pp. 8-9) — and one may well wonder whether it reflected the honest opinion of its author, let alone « the dominant opinion at the close of the Gilded Age » (p. 5). All of these statements, however, are proclaimed to reveal belief in the « axiom of indispensability », i.e., in the railway's « necessity » (pp. 1-9, *passim*); clearly, Fogel could establish the relevance of his experiment to the question of the railway's « importance » — as defined by his own sources — only by ignoring the gap between « necessity » and « sufficiency ». By the same token, of course, the experiment's results are misinterpreted, to the extent that the evidence that the railway was not « indispensable » is (wrongly) taken to imply that it was not historically « important » in the common sense of the terms. The root of the problem, to repeat, appears to be the ill-considered application of contemporary applied economists' methods to historical problems. The economists' standard counterfactual experiment, which keeps « other things » as they are, is indeed appropriate to revealing « importance » — i.e., (conditional) « sufficiency » — in its normal prospective use; but when applied to historical questions, this very method reveals « necessity » rather than « sufficiency ». In a word, the experiment no longer accomplishes precisely that which makes it attractive — and therefore familiar — in the non-historical uses for which it was developed.

The solution, of course, is so to modify the form of the experiment as to retain its essence, i.e., the ability to make manifest relations « sufficiency », and therefore historical « importance ». The Fogelian experiment reveals « necessity »; if « necessity » is not « sufficiency », it may nonetheless be combined with it, in which case the experiment will reveal « sufficiency » as well as « necessity ». Thus, if railroad technology had no *potential* substitutes — either because canals could not be built, or because they were already built (and thus not potential but actual) — the prospective and retrospective experiments that keep alternative opportunities and behavior patterns as they

were yield the same results (the same ΔNI, albeit of course of opposite sign). At times, then, the Fogelian experiment will (happen to) reveal « sufficiency »; but at other times, it must be modified to be made to do so, essentially by assuming away substitutes until the y that x is necessary for has grown to coincide with the y that x is sufficient for. This is not to say that the historian's thought experiments need not be *ceteris paribus*: clearly, other things must be held constant, if the experiment is to be a *controlled* experiment; but for the experiments' results to be comparable, other things need be held constant only across experiments. There is no logical need to (also) hold these other things equal to what they actually were, i.e., there is no need to restrict the interpretation of *ceteris paribus* to *ceteris* factually *paribus*, to the exclusion of *ceteris* counterfactually *paribus* — particularly as the latter solution may make for the more meaningful experiment.

Imagine, for instance, the trivariate world summarized in Table 1: it admits four states, i.e., four different combinations of six particular values of the three variables. In the context of these four states alone, Δa (like Δb) is necessary (but not sufficient) for Δc, and *non*-Δa (like *non*-Δb) is

TABLE 1

	State of the world	State of variable 1	State of variable 2	State of variable 3	Relationships implied by the alternatives **	Relationships suggested by the alternatives **
Trivariate alternatives:	1.	a	b	c	$\Delta a \rightarrow \Delta c$, $\Delta a \not\Rightarrow \Delta c$ $\Delta b \rightarrow \Delta c$, $\Delta b \not\Rightarrow \Delta c$ $(\Delta a+\Delta b) \rightarrow \Delta c$, $(\Delta a+\Delta b) \Rightarrow \Delta c$	$(\Delta a+\Delta b) > \Delta c$
	2.	$a+\Delta a$	b	c		
	3.	a	$b+\Delta b$	c		
	4.	$a+\Delta a$	$b+\Delta b$	$c+\Delta c$		
Bivariate alternatives:						
experiment α:	3.	a	$b+\Delta b$ *	c	$\Delta a \rightarrow \Delta c$, given $(b+\Delta b)$ $\Delta a \Rightarrow \Delta c$, given $(b+\Delta b)$	$\Delta a > \Delta c$
	4.	$a+\Delta a$	$b+\Delta b$	$c+\Delta c$		
experiment β:	1.	a	b *	c	$\Delta a \not\Rightarrow \Delta c$, given (b)	$\Delta a \not> \Delta c$
	2.	$a+\Delta a$	b	c		
experiment γ:	2.	$a+\Delta a$ *	b	c	$\Delta b \rightarrow \Delta c$, given $(a+\Delta a)$ $\Delta b \Rightarrow \Delta c$, given $(a+\Delta a)$	$\Delta b > \Delta c$
	4.	$a+\Delta a$	$b+\Delta b$	$c+\Delta c$		
experiment δ:	1.	a *	b	c	$\Delta b \not\Rightarrow \Delta c$, given (a)	$\Delta b \not> \Delta c$
	3.	a	$b+\Delta b$	c		

* Variable kept constant under the *ceteris paribus* clause.

** $x \rightarrow y$ signifies that: x is (really) necessary for y; the lack of x is (really) sufficient for the lack of y; y is (cognitively) sufficient for x; the lack of y is (cognitively) necessary for the lack of x.

$x \Rightarrow y$ signifies that: x is (really) sufficient for y; the lack of x is (really) necessary for the lack of y; y is (cognitively) necessary for x; the lack of y is (cognitively) sufficient for the lack of x.

$x > y$ signifies that: x caused y; the lack of x caused the lack of y.

$x \not\Rightarrow y$ signifies that $x \Rightarrow y$ does not hold.

$x \not> y$ signifies that $x > y$ does not hold.

sufficient (but not necessary) for *non*-Δc. (If Δa and Δb represent the addition of a real variable — e.g., the addition of trucks and of roads, Δc being the appearance of trucking — we are in the presence of « necessity »; if Δa and Δb represent the removal of a real variable — e.g., the removal of railroad-building opportunities and of canal-building opportunities, Δc being the removal of a conspicuous amount of production — we are in the presence of « sufficiency ».) This conclusion does not depend on whether the factual set (i.e., the actual historical situation) was 1., 2., 3., or 4.; it is reached by an exhaustive (i.e., trivariate) investigation of *all* possible states of the world, factual and counterfactual, through controlled experiments α, β, γ, and δ. Of these experiments, α and γ are positive, in the sense that they bring relations to the fore; β and δ are instead negative, and serve essentially to balance α and γ by disproving the existence on a *tri*variate level of one of the relationships established by α and γ on a *bi*variate level. The « importance » of Δa for Δc (and *non*-Δa for *non*-Δc) is thus revealed by experiment α: the *cetera* that are kept *paria* are factual if the factual set was 3. or 4., and counterfactual if the factual set was 1. or 2. (if the factual set was 1. or 2., the experiment that keeps *cetera* factually *paria* is β, which fails to reveal the sufficiency of *non*-Δa for *non*-Δc, etc.). Indeed, if the factual set was 3. or 4., « necessity » and « sufficiency » coincide, and the *ceteris* factually *paribus* counterfactual experiment does reveal « sufficiency »; if the factual set was 1. or 2., « necessity » and « sufficiency » do not coincide, and the *ceteris* factually *paribus* counterfactual experiment fails to reveal « sufficiency ». If the factual set was 1. or 2., then, the experiment which reveals the « importance » of Δa is performed by *making* « necessity » and « sufficiency » coincide — by assuming away *non*-Δb, which is here a perfect substitute for *non*-Δa, and substituting *ceteris* counterfactually *paribus* for *ceteris* factually *paribus*. In analogous fashion, of course, one would reveal the « importance » of Δb for Δc (and *non*-Δb for *non*-Δc) by performing experiment γ.

Any historical event is, obviously, the result of its sufficient conditions. Some of these conditions are redundant, and are therefore conditionally (i.e., given the others) sufficient but not necessary; others are not redundant, and are therefore conditionally necessary as well as sufficient. *Ceteris* factually *paribus* experiments allow one to attribute an event to any one of its conditionally necessary (and sufficient) conditions; *ceteris* counterfactually *paribus* experiments extend the range of options to include conditionally sufficient (but not necessary) conditions. The experiment one actually chooses will depend on the historical condition (of the sufficient set) one chooses to highlight as particularly « important »; « the cause » of an event identifies not just any conditionally sufficient condition, but one that is selected as being somehow particularly meaningful. Meaningfulness, of course, is in the eye of the beholder: it will depend on the observer's interests, frame of

reference, and *forma mentis*. Some see significance in the apparent disproportion of cause and effect (« for want of a nail... »); others, in the element of human control (thus those who stress the impact of public policy on development); and so on.[8] Generally, however, opportunities that *were* taken advantage of will be viewed as more meaningful than those that were *not* taken advantage of; and if « importance » is attributed to an opportunity, it will typically be attributed to the taking of that opportunity, in its historically individual form. Thus, to return to the examples discussed above, one may wish to attribute *non-*Δc not just to « either *non-*Δa or *non-*Δb », but to one of these in particular: if one chooses to select cheap transportation (rather than, say, natural resources) as « the cause » of development, and finds various technological opportunities (canals, railroads) only *one* of which (railroads) was taken advantage of, one will attribute development to the railroads even though canals could have been built instead; and one will attribute it to the railroad network actually built, even though different railroad networks could have been built instead. By allowing *ceteris* counterfactually *paribus* experiments, then, one allows the attribution of concrete historical events to other concrete historical events (as is commonly and not improperly done) rather than just to such formal characteristics of those concrete events as are sufficiently abstract to be devoid of substitutes.

It thus seems possible, on this basis, to eliminate the confusion and resolve the controversy on at least the formal aspects of « counterfactual » history. So far, one of the most common reactions to Fogel's work on the part of those who found his numbers unsatisfactorily low has been to declare his analysis irrelevant *qua* counterfactual. Even Albert Fishlow,[9] incredibly, lent his voice to this chorus, as if his own estimate of the railroad's direct benefits involved « measuring what never occurred », in a « hypothetical world... without the innovation of the railroad » any less than Fogel's. Fishlow even took Fogel's statement that those potential canals *would* have been built in the absence of the railroad to mean that they *should* have been built even in the presence of the railroad — apparently in the belief that if the development of the North Central States depended on cheap transportation but not on the railroad (see FOGEL, 1964, p. 110), then it had to depend on the (unbuilt) canals! As noted above, the Fogelian method would no more attribute importance to the imperfect substitute for the railroad (the canals) than to the railroad itself: to return to Table 1, if the factual set is 1., the *ceteris* factually *paribus* experiments are β and δ, and there is no reason whatever to believe that experiment β (which does not tie Δc to Δa),

[8] For a more extended discussion of the relevant criteria, see S. FENOALTEA, *Counterfactual History, Public Policy, and the Stages of Growth: Reflections on the Italian Industrial Experience from the Risorgimento to the First World War* (forthcoming).

[9] A. FISHLOW, *American Railroads and the Transformation of the Ante-Bellum Economy* (Cambridge: Harvard University Press, 1965), pp. 57-58.

say, must be followed by experiment γ (which would tie Δc to Δb). More generally, of course, the point is that one need not reject counterfactual analysis *per se* if one is inclined to reject the Fogelian (*ceteris* factually *paribus*) experiment; for the former is in no way exhausted by the latter, and the one is compelling as the other is not. Empirical analysis, to repeat, requires experiments, and the economic historian's experiments are thought-experiments, which produce and utilize not « facts» but « counterfacts »: analytical history is thus at least implicitly counterfactual, and therefore often best explicitly so. Causal historical importance, in turn, is identified with « sufficiency », and attributed to (particularly significant) individual, concrete historical events or situations; but the Fogelian counterfactual paradigm reveals only « necessity », and thus allows the attribution of significance only to conditions that are devoid of substitutes — and therefore, almost inevitably, only to some abstract general characteristics of actual historical phenomena. That counterfactual paradigm thus in general fails to come to grips with the standard notions of historical importance; but its deficiency may readily be remedied, by assuming away possible substitutes until « necessity » coincides with « sufficiency » — i.e., by assuming away possible individual cases which share the relevant abstract characteristics until the latter may be found only in the actual individual case of history. The appropriate controlled experiment is necessarily *ceteris paribus*, and certainly counterfactual; indeed, it differs from the standard paradigm only in that the *cetera* that are kept *paria* are allowed in principle to be not only what they actually were, but also, to the required extent, what they actually were *not*. The justified complaint with the standard counterfactual paradigm is thus met not by abandoning its counterfactual aspects, but by extending them, if one will, to include experiments *ceteris* counterfactually, as well as factually, *paribus*. If Fogel had imitated the substance of economists' counterfactual experiments, instead of letting himself be seduced by their mere form — if he had measured the contribution of the railroads without allowing for possible canals, making clear that that counterfactual experiment was, appropriately, *ceteris* counterfactual *paribus*, rather than *ceteris* factually *paribus* — the false controversy on the legitimacy of counterfactual analysis *per se* might have been avoided altogether... but such counterfactual historiography is perhaps best not pursued too far.

III.

To illuminate a historical problem, the counterfactual experiment must fit not only the abstract problem, but the concrete historical reality. Apart from the difficulties that arise when the model is used to perform the « wrong » experiment, then, one must face those which arise when the experiment is

performed on the « wrong » model. In some sense, of course, any model may be considered historically « wrong » because it is an abstraction, and « new » economic historians have in fact been faulted for their willingness to operate with extremely succinct summaries of reality; but this criticism appears obtuse. Even if one seeks as rich a picture as possible, the rough sketch with which one begins cannot fairly be blamed for being shorn of details — while if one seeks a rapid overview, details may well be just counterproductive clutter. The « wrongness » of a model is here to be understood rather differently, as a misinterpretation of reality *given* the level of abstraction. To sustain the metaphor, the rough sketch is « wrong » not because it is a rough sketch, but because it grossly distorts the relative proportions of the object. It is precisely on this more meaningful definition of « wrongness » that one is struck by the tendency of « new » economic historians to use models which appear to be « wrong » — and « wrong », typically, not because they depart from the usual economics-textbook assumptions about behavior, but because they fail to do so, despite the at times painfuly manifest inappropriateness of those assumptions to the historical reality in question.

Perhaps the most glaring example of assumptions which are as historically inappropriate as they are familiar relates to the conditions of factor supply. The standard assumption of textbook economics is that factor availabilities are given: most often, only one (closed) economy is contemplated at all, and even the theory of international trade has traditionally been concerned with the exchange of products in the context of factor immobility. No one would deny the appropriateness of the standard assumption to a host of problems, or the richness of the standard analysis, for instance of allocative efficiency; by the same token, one can readily understand Ricardo's use of that assumption, if his model was to establish the benefits of Anglo-Portuguese commodity trade on the basis of comparative advantage (rather than the depopulation of England for Portugal on the basis of absolute advantage).[10] Theoretical usefulness, however, is not empirical veracity: the nineteenth century was the age of mass migration of labor and free movements of capital, and the standard fixed-factor-endowment model misrepresents — at a minimum — those national economies that were the sources or recipients of major factor flows. For all that, « new » economic historians typically treat the textbook model as a valid summary of past reality; in discussing nineteenth-century America, for instance, international factor mobility is often ignored by the model and calculations, and appears only in the verbiage of qualifications. Once again, the paradigm may be found in the classics of the discipline: thus Fogel (1964) measured « the loss in national income occasioned by the absence of the railroad » by estimating the difference in resources consumed

[10] D. Ricardo, *Principles of Political Economy and Taxation* (London: Everyman's Library, 1965), ch. 7; note esp. p. 83.

in the transportation sector with and without railroads (« the social saving »), « on the assumption that national income would have fallen only because more productive resources were required to provide a given amount of transportation service and that all productive resources not used for transportation would have remained fully employed » (pp. 20-21). The model thus lets the railroad influence national income only by altering average factor productivity; factor-supply response to demand shifts is assumed away, and the impact of the railroad on demand is accordingly irrelevant. International migration is mentioned (p. 226) only as one of many readily dismissed qualifications to the conclusions — a negligible « derived effect », in contrast to the « primary effect » represented by the « social saving » (pp. 208, 224). Fishlow's (1965) approach is fundamentally similar, if slightly more complex, as he shares the assumption of fixed factor endowments, but does not assume full employment. The railroad's « direct benefits » thus remain identified with the « social saving », which is however viewed as an increase in *potential* output only; that estimate is duly followed by a (rather impressionistic) discussion of the railroad's relation to aggregate demand. The assumption of fixed factor endowments underlies virtually the whole discussion of the railroad's impact on aggregate production through « direct benefits » and macroeconomic backward linkages (pp. 18-118; see esp. pp. 54, 99; also pp. 14-16); the sensitivity of the factor endowment itself to railroad demand (through international real factor movements) receives only brief mention in the very last paragraph of this long section (pp. 117-118) — even though a later chapter on the railroad's extension into the West argues at some length that « railroad construction... exerted a considerable influence on [European] immigration » (pp. 200-202). Well it may have: in a world where internationally mobile labor and capital could choose among a variety of national markets, the elasticity of factor supply to any one of these was presumably quite high; but this implies that equilibrium national products and factor endowments were determined by relative opportunities for remunerative factor employment, and that (unless the contemplated equilibria are wholly within the band defined by the factor-import and factor-export points) such constructs as « full employment », « potential output », and *a fortiori* « social saving » are irrelevant to the problem at hand.

This tendency to misrepresent indubitably relevant behavior (such as factor-supply response to demand shifts) is only one result of « new » economic historians' apparent willingness to borrow their models from contemporary economics textbooks, rather than construct them on the basis of their historical knowledge; another is the tendency to put what appear to be inappropriately narrow limits to the proper scope of the discipline. It is instructive, in this connection, to ponder Fogel's (1964) definition of the « true » social saving (pp. 20-21n). Fogel there contemplates four levels of income (Y), characterized by two independent binary variables: the presence

or absence of the railroad (RR) on the one hand, and the presence or absence of historically present institutional constraints (IC) on the other — say Y(RR, IC), Y(RR, non-IC), Y(non-RR, IC), and Y(non-RR, non-IC), the first of these being of course the actual level of income. On the basis of these, Fogel offers «four possible ways of defining the 'true' social saving», which in the present notation are 1. Y(RR, IC) — Y(non-RR, non-IC); 2. Y(RR, non-IC) — Y(non-RR, non-IC); 3. Y(RR, non-IC) — Y(non-RR, IC); and 4. Y(RR, IC) — Y(non-RR, IC). Definitions 2. and 3. are dismissed as « irrelevant »; as the comparison excludes Y(RR, IC), they cannot measure what the railroad « actually did contribute to economic growth ». Definition 4. is also dismissed, on the grounds that it is «indeterminate since there is no way of knowing » Y(non-RR, IC), so that « definition 1. is the ' true social saving ' to which the discussion... refers ». In point of fact, it is obvious that the controlled experiments, measuring the railroad's contribution to potential and actual income respectively, correspond to definitions 2. and 4.; definitions 1. and 3. are both inadmissible because they include simultaneous changes in *both* independent variables, and thus fail to make manifest the *ceteris paribus* impact of the railroad itself. As Fogel pointed out, definition 3. overstates any reasonable measure of the « true » social saving; but by exactly the same token, definition 1. leads to an obvious understatement. Indeed, it measures a movement from a point *inside* the with-railroad production frontier to a point *on* the without-railroad production frontier; as the removal of institutional constraints might well increase production more than the removal of the railroad would reduce it, Fogel's proposed definition could perfectly well attribute a (totally irrelevant) negative « social saving » to the technological innovation under examination. Definition 1. is thus obviously deficient, while definition 4. is equally obviously correct; yet Fogel chose the former over the latter. Definition 1. and its corresponding counterfact require only the texbook model of production possibilities set by technology and resource constraints; definition 4. requires a model expanded to incorporate institutional constraints as well. As Fogel asserted that such constraints were historically present, and mattered so much that whatever neglects them is «irrelevant» to actual production, he implied that the textbook model is here too narrow, and empirically « wrong »; yet he felt entitled to retain it as the basis for his ideal experiment, and to dismiss the « right » model and the appropriate experiment « since there is no way of knowing » what the appropriate counterfact might be. In thus dismissing definition 4., Fogel echoed the very language of those critics who reject counterfactual history as they declare unknowable and infeasible what they do not know or cannot do themselves; in thus accepting definition 1., Fogel selected as ideal an experiment that was not even *ceteris paribus*, and thus, strictly speaking, scientifically worthless. Of course, Fogel did not perform his ideal experiment at all: he did not try to measure the « true

social saving », but only to limit it by a suitably low upper bound. His choice among those proposed definitions was thus irrelevant to the calculation he actually performed, and he may correspondingly have considered that choice less carefully than he otherwise would have. While his error on this point was conceptually fundamental (and methodologically revealing), then, it was largely without practical consequence; and — *teste* Columbus blundering into the New World — even a fundamental confusion is not unbecoming to pioneers.

If the Procrustean bed of contemporary economics-textbook models and assumptions could thus accomodate Fogel's actual measure of the « social saving » without damaging it overmuch, the same cannot always be said for the products of his epigones; and one could not ask for a more revealing spectacle of the « new » economic historian hamstrung by his view of the discipline than that provided by Donald McCloskey's (1970a) investigation of « Britain's loss from foreign industrialization ».[11] The title defines a supremely interesting and important historical problem: the common presumption, at least, is that German and American industrialization are at the root of the major developments of recent Western history — Germany twice channelling its economic might into military bids for supremacy, Britain declining from « top nation » almost to an American protectorate as it twice fought back Germany's challenge with American help. A quantitative historian who produced even a very crude estimate of the cost to Britain of that arms production, bloodshed, and abridgment of sovereignty — or who convincingly established that Anglo-German arms races and warfare were in fact independent of German industrialization — would go far indeed toward fulfilling the promise of McCloskey's title. But textbooks of (bourgeois) economics view nation-states as collectives that engage only in voluntary and mutually beneficial exchange; wars, threats of war, and responses to threats of war are simply ignored. So too McCloskey's analysis: the model incorporates no interaction other than international trade, and Britain's loss, as calculated, is simply the income lost from the *ceteris paribus* decline in export demand. Characteristically, Britain is taken to be a fully employed economy with fixed factor availabilities; this assumption is justified by a reference to a parallel effort (1970b) which treats Britain's low recorded unemployment rate as evidence that « constraints of supply... were the only binding constraints [on output] »,[12] instead of as evidence of the ease with which resources left Britain for employment elsewhere (so that the persistence of migration and « full employment » in a narrow statistical sense would argue for a very high, rather than very low, elasticity of factor supply). Total

---

[11] D.N. McCloskey, *Britain's Loss from Foreign Industrialization: A Provisional Estimate*, « Explorations in Economic History », 8, Winter 1970-71, pp. 141-152.

[12] D.N. McCloskey, *Did Victorian Britain Fail?*, « Economic History Review », XXIII (December 1970), p. 448.

output, then, is independent of export demand; the only loss is thus from the deterioration of the terms of trade, as « the rent Britain could extract from the rest of the world would have fallen » (1970a, p. 146).

With this model, the estimate of Britain's loss from foreign industrialization, calculated for 1913, is almost certainly small; but the result lacks empirical credibility, as it hinges not only on that « wrongness » of the model which reduces Britain's loss from a decline in export demand to a terms-of-trade effect, but also on that « wrongness » which identifies Britain's loss from foreign industrialization with an export-demand effect in the first place. Even if the model were otherwise empirically sound, its failure to contemplate military relations establishes the presumption that the resulting estimate covers no more than a comparatively trivial component of Britain's actual loss from foreign industrialization; and while an estimate of the income lost from export decline would be of interest in its own right, one should be wary of letting it masquerade as anything more than what it is. McCloskey's analysis is in fact concerned only with that loss through trade relations; and even though he quite explicitly points out that *that* loss from foreign industrialization was probably dominated by other losses, through linkages his model ignores (pp. 147, 151), he nonetheless consistently refers to it as « the » (rather than, say, « this element of the ») loss from foreign industrialization (e.g., pp. 147, 150). Such phraseology has a power of its own, and the author himself is misled by it: thus, for instance, the unwarranted conclusion that Britain's real income was relatively unaffected by international developments, as opposed to domestic ones (p. 152). The most egregious slip is no doubt the claim that the model's implication that « the impact of the industrialization of Germany alone was very small » represents « a notable result in view of the contemporary alarm over Germany's rise » (p. 150), as if the alarm that fuelled the Anglo-German naval race was really over losses through the peaceful exchange of commodities in 1913. One may wish to measure the damage done by a hurricane even before there is any destruction by wind and water, ignoring the cost of any precautionary protective measures or the real loss from the prospect of uninsured damage; but who could call that *the* damage done by the hurricane, or consider the smallness of the estimate « a notable result in view of the contemporary alarm » over the impeding disaster?

McCloskey justifies the identification of Britain's loss from foreign industrialization with the loss from a decline in export demand only to the extent of indicating that the losses he ignores were « political, not economic » (p. 151). Such welfare losses as are caused e.g. by consumption cuts to allow for war production are clearly « economic » in their nature, and differ from those caused by a drop in export demand only in the mechanism that generates them; it is thus no doubt to that generating mechanism that the adjectives « political » and « economic » are to be referred. The suggestion

of such a distinction in such an article and indeed in such a journal is that the scope of models in economic history is legitimately restricted to « economic » relations, as opposed for instance to « political » ones. But how can one justify such a view? Whose business is it to examine the *total* historical impact of a change in one economic variable (say non-British industrial production) on another economic variable (say British real income), if not the business of economic history? Why should « economic history » be history impoverished by the lack of any consideration of « non-economic » linkages, rather than history enriched by the proper consideration of « economic » ones? Even if some relevant behavior were « not economic », then, an economic *historian* would have no excuse for ignoring it. Indeed, would even an *economist* ignore it? Though the focus of (bourgeois) economics has historically been on market exchange, economists hardly limit their models exclusively to market relations. For instance, such non-pecuniary externalities of construction projects as floods or flood control are taken in stride, and not dismissed as « hydraulic, not economic »; and, in particular, the tools of economics have been successfully applied to the study of arms races,[13] the effectiveness of warfare, and the like. One need not accept the Marxian view that politics are the continuation of economic relations by other means to be loath to conclude that obviously relevant « political » linkages are, any more than « hydraulic » ones, beyond the proper scope of empirical economic analysis.

In sum, while models are inevitably not portraits but caricatures of the reality they refer to, they can and will be empirically persuasive if they are lifelike. A model must therefore seek to capture the most relevant reaction patterns that mark the historical reality in question; an *a priori* empirical model is almost a contradiction in terms. « New » economic historians, however, are so partial to the familiar textbook model as often either simply to assume it applies or to conclude that it *does* apply on the strength of highly ambiguous evidence. Empirical validity thus becomes a matter of luck, as the largely predetermined model may or may not fit the reality ostensibly being investigated; if it does not, the exercise, however otherwise interesting or ingenious, is empirically valueless. There is correspondingly little merit to the suggestion that the models of economic history should be systematically predetermined at least to the extent of contemplating only « economic » linkages, no matter how fundamental « non-economic » ones may be to the problem at hand. Such a principle would place particularly stultifying limits on economic history, as problems that are economic and historical would no longer be *ipso facto* wholly within its purview; indeed, as economists typically

[13] In fact, the latest such study by J. C. LAMBELET is devoted precisely to *The Anglo-German Dreadnought Race, 1905-1914* (presented to the Tenth European Conference of the Peace Science Society (International), Vienna, August 28-30, 1973).

view the boundaries of their own discipline as elastic enough to embrace all the reaction patterns relevant to the problem under investigation, economic history so narrowly construed should most likely be considered deficient not only as history but also as empirical economics.

IV.

Where the «old» economic history was too often content with impressionistic and allusive argument, the « new » proclaims the principle that historical conclusions must be rigorously derived from explicit premises. The discipline's new-found analytical richness is an improvement indeed, and its concomitant rise to respectability as applied economics is nothing if not encouraging: at long last, economic history is in principle as « economic » as it should be. This progress, however, is not unmixed. Economists now dominate the discipline; with perhaps a hint of the arrogance of technical experts, they do not show much patience toward those who lack their training, and historians are relegated to the fringes of the field. The «new» economic history is thus if anything even less interdisciplinary than the «old»; not surprisingly, its characteristic weaknesses reflect the narrowness of its base.

Empirical problems are most readily handled by being identified with an abstract problem whose solution is already known: without the counterweight of historians' critical counsel, economists may find the familiar textbook models and procedures irresistibly attractive even though they may be fundamentally unsuited to the historical reality or problem under investigation. The formal manipulation of the model may be beyond reproach; but if the work is thus flawed at the interface of the abstract and the concrete, it is poor applied economics and correspondingly poor history. While the promise of modern analytical methods surely justifies the fervour of the « new » economic history's enthusiasts, it is sobering to note that counterfactual methodology has at times confused its devotees no less than its critics, and that some of the discipline's most conspicuously revisionist conclusions are in fact untenable. The « new » economic history of economists has set itself up as the antithesis of the « old » economic history of historians; but the best economic history, as free of the empirical vulnerability of the « new » as of the analytical laxity of the « old », is truly interdisciplinary.