Counterfactuals in Logically Formed Economic Analysis

David J. Loschky
University of Missouri - Columbia

Economists continually strive to construct persuasive (true) explanations (predictions). These explanations derive their persuasive power from various sources. Among these sources we certainly must include the validity of all argument structures used.[1]

All, or nearly all, of our explanations are organized around some explicit or implicit argument structure.[2] The reiterative appeal of « this proves », of

[1] Validity and proof will be used here only in their logical sense. For this reason an argument is considered only more or less persuasive. This is a weaker criterion than asking the argument to be true. Whatever we assert about persuasive arguments, therefore, applies *a fortiori* to arguments which are held to produce truth.

The term argument structure refers to an argument's form. The same form, obviously, can be applied to many specific arguments. If the form is valid in one usage it is valid in all usages. An example of an argument form would be the familiar one; premise 1) all men are mortal, premise 2) Socrates is a man, therefore 3) Socrates is mortal. The general form of this argument can be seen as the statements 1) for all x, if x is an M then x is D, premise 2) S is an M, and therefore 3) S is D. This form, as can be seen by inspection, is valid. Validity simply means that if the premises are true, then the conclusion is true. Validity, therefore, is a characteristic of arguments' forms and is independent of whether the premises are true or false. An argument form can be valid, therefore, while the conclusions drawn from a particular instance of the form may be false simply because some one or more premises are false.

For a discussion of argument forms and of validity see A. N. Prior, *Formal Logic,* 2nd Ed., (Oxford, at the Clarendon Press, 1962) pp. 1-3 or Irving M. Copi, *Symbolic Logic*, 3rd Ed., (Macmillan Company), pp. 4-6 and pp. 19-25.

This analysis of counterfactual will be based upon argument forms alone. Therefore we need not concern ourselves with a series of specific arguments to evaluate their promise and their problems.

[2] Robert Fogel bases much of his work upon just this position. In his *The Specification Problem in Economic History*, « The Journal of Economic History », Vol. 27, No. 3, Sept., 1967, pp. 283-308 he develops the concept of implicit theorizing upon which

formal proofs developed in footnotes or appendicies, of « the obvious logic of our answer », and of « we may infer, therefore », permit us to draw no other conclusion. Obviously we economists believe valid structures lend plausibility. Some rational beings even see valid structures as a necessary condition for plausible, persuasive conclusions. Equally obvious is the view that invalid structures lend no credibility to an argument. The importance attached to demonstrated errors in proofs or to the criticism « the argument is illogical » admits of no other interpretation. But these points, (1) that we all depend upon valid structures to lend credibility to our arguments, and (2) that if our arguments use invalid structures, then our arguments are much less plausible, are surely too obvious to require defense. Disagreements only arise when, faced with a specific study or class of studies, we must judge whether the argument's form is valid and, thereby, determine the credibility to give these studies' conclusions.[3]

Counterfactual arguments represent a case in point. They are a particularly interesting form which is frequently encountered. A counterfactual argument is one based upon one one or more premises known to be false and, according to such authors as Fogel, upon law-like statements.[4] Both characteristics make these arguments interesting. But the false premise creates difficulties in argumentation. Difficulties arise because some arguments, such as null hypothesis tests, cannot be employed.[5] This, of course, restricts our analysis's range. But other difficulties surface which frequently are much more serious. These latter problems revolve around the argument's consistency. As we shall show below, consistency frequently is not maintained, and, often enough, when it is, it is achieved only through great effort and cost. The frequency with which consistency is not maintained, or with which

---

he bases his series of counterfactual studies. Implicit theorizing is defined to mean the argument structure implicit in any line of argument, and includes, as well, a specific instance of the general form.

[3] The case is similar when mathematical analysis is used to support an argument. If a proof is found which used argument structures inconsistent with accepted usage the value of the proof is certainly called into question. For example, if a derivative is formed from a sequence in which all values are defined to the nth term and if a value of 1 is assigned to all terms ending in even values and 0 is assigned to all terms ending in odd values, the implication (the derivative) must surely be called into question. The same argument applies, *mutatis mutandis*, to any analysis purporting to be logical in structure. If, in fact, it uses invalid steps in a derivation the method cannot lend plausibility to the general argument.

[4] Fogel's quotation of Hemple at the start of his railroad study certainly suggests this view. ROBERT W. FOGEL, *Railroads and American Economic Growth: Essays in Econometric History*, (Johns Hopkins University Press, 1964), p. 1.

[5] See the discussion of JACK MELITZ, *Friedman and Machlup on the Significance of Testing Economic Assumptions*, « Journal of Political Economy », Vol. 73, No. 1, Feb., 1965, pp. 37-60, and D. V. T. BEAR and D. ORR, *Logic and Expediency in Economic Theorizing*, « Journal of Political Economy », April, 1967, pp. 188-196.

the problem is lightly brushed aside, suggests this difficulty is either not appreciated or ways to circumvent it are not known.[6] Our aim is to develop counterfactual analysis within an analytic framework in such a way that the development itself will suggest procedures to overcome some difficulties. We will not attempt to banish counterfactuals from argumentation, but rather will qualify their usage.

Our concern centers on counterfactuals used to test theories in cases which involve two models. As we are concerned with Economic History, I will illustrate the discussion with examples drawn from that field. The conclusions, however, will be perfectly general. The arguments will apply, for example, to the counterfactual analysis required by anti-trust legislation; counterfactual arguments showing the desirability of one fiscal policy as opposed to another will be implied; if a counterfactual argument suggests one growth plan as opposed to another, the argument should conform to the procedures we will develop; and if an econometric test is based upon one or more counterfactuals, these same problems will be encountered. But continued illustration of this argument's range is unecessary. For it is clear that various basic assumptions of economic theory itself are reasonably questionable and, therefore, *may* be counted as counterfactual.[6]

Counterfactual tests are not only used widely between fields but extensively within fields. In Economic History, for instance, we see the validity of many important historical interpretations tested by counterfactual analysis. Were railroads important for United States growth? Did British Imperial Policy retard development in the North American colonies? Would the Northwest Territory have grown as rapidly had canals not been built? — these represent only a few of the historical interpretations tested and found wanting. Would it be fair to say that such counterfactual test results have changed the face of Economic History?

Some would say yes; others, no. The difference of opinion hangs upon the plausibility of the counterfactual test and that in turn depends, I believe, upon whether one intuitively feels the counterfactuals have been handled properly and with due sensibility. Some economic historians have embraced counterfactual tests with a display of open arms, others with arms openly displayed. For instance we find William N. Parker asserting « In this effort,

---

6 This paper is meant to contribute to the field of Economics, not the field of Philosophy. References to « the literature », therefore, relate to the field of Economics only.

7 An example of a questionable assumption is, of course, the assertion that firms attempt to maximize profits. When a model employing this assumption is used the universal statement must be converted into one concerning some particular firm or set of firms operating at specific times and places. See the discussion of this point by Bear and Orr, *Logical Expediency*. When this particular assertion is made concerning a specific firm or set of firms the assertion may be correct, and then again it may be quite false. In some cases, therefore, the economic model employs a counterfactual.

in which I must admit or confess to a contribution, we have been forced to face explicitly the problems of hypothetical history [he refers to counterfactuals], a use of imagination and reasoning without which, I feel, analysis in historical work is really not possible ».[8] On the other hand Robert Fogel roundly denounced Temin's analysis of wage differentials between the United States and the United Kingdom simply because the analysis, according to Fogel, is counterfactual.[9] Between these two particular positions we find many others asserting that counterfactual tests are plausible only if proper precautions are taken: for instance, they must not be too large; or, somehow (and in some cases) other conditions must be wrenched (by some means) into conformity with the counterfactual; or they are informative (so long as we do not take them too seriously). Our analysis will display the reasonable ground for the sceptical, intermediate position. It will also suggest ways to lend greater credibility to counterfactual studies by establishing a few procedural rules.

One more limitation upon this argument's range must be mentioned. Instrumentalism, as espoused by Friedman, has been attacked as counterfactual.[10] Be that as it may, our argument will not bear upon that particular controversy. Suffice it to say that, if you accept instrumentalism, this argument will in no way shake your belief in ' as if ' statements. If you do not accept instrumentalism the argument may sound convincing. The problem obviously is antecedent to logical analysis of the type we present here. Nevertheless, even Friedman would grant that some assumptions do not reflect accurately peoples' actions. These assumptions are therefore counterfactual for Friedman. Our argument will apply to the use of any such assumptions.

Generally speaking, our conclusions will be that counterfactuals used to examine alternative hypotheses have a place in our profession. But it will be shown that they are far less powerful than studies which do not utilize counterfactuals; that they require rather unusual freedom with the rules for

[8] William N. Parker, *From Old to New to Old in Economic History*, « The Journal of Economic History », Vol. 31, No. 1, March, 1971, p. 7. It is worth mentioning that this is the presidential address delivered to The Economic History Association in the fall of 1970.

[9] See Robert Fogel, *The Specification Problem*. Fogel bases his criticism of Temin upon the assertion « The last four assumptions are obviously at variance with reality », p. 300, in other words he criticizes Temin for arguing in a counterfactual vein. Fogel follows the argument through, using such remarks as « Temin provides no justification at all for assumption (C)... » and « Here again Temin's evidence is insufficient to establish his proposition », both p. 301, and « What then are the implications of Temin's assumptions? And how would his findings be altered if assumptions (c), (d), and (g) were changed along the lines suggested above? », p. 302. Fogel proceeds 1) to suggest assumptions less flagrantly counterfactual, and 2) to show that these changed and more realistic assumptions lead to quite different conclusions, pp. 298-308.

[10] Instrumentalism is associated with the use of « as if » statements as premises or basic assumptions. An example would be the premise « Firms act as if they were attempting to maximize profits ».

formal argumentation; that they do not permit one to test theories despite claims to the contrary; that they require a great deal more analysis to provide an interesting study than is the case with non-counterfactual works; and lastly, *that logical analysis can lend no plausibility even to the limited range of inferences outlined above unless extraordinary precautions are taken, precautions usually honored in the breech.*

These conclusions will be developed in the following sequence. Section I will prove that if a counterfactual enters in a non-trivial way into any premise of an argument designed to test a model, then either the counterfactual must appear in the argument's conclusion or if it does not appear in the conclusion then it must have been assumed. The logical content of a hypothetical implication with a counterfactual antecedent will be shown. Section II will discuss the implications of assuming a counterfactual. Various problems associated with its assumption will be pointed out. A concluding reference to hypothetical statements will be couched in a discussion of inductive proofs. Section III will suggest several techniques for coping with counterfactuals. Section IV will present a summary of our conclusions.

Section I

A

Difficulties with counterfactuals in a deductive argument arise only when the counterfactual is assumed or when it is either the consequent or the antecedent in an implicative conclusion. In this section we will prove if a counterfactual is introduced into the typical testing argument, then either the counterfactual is assumed or it must appear in the argument's conclusion. Section II will developed some implications of this proof.

Central to most tests in economics are two models. For example, one is the economic model and the other the statistical model; or one may be a historical model and the other an economic model; or one may be one growth plan and the other another growth plan; or we may contrast two policy models.

From these models [including the attendent data input] [11] inferences are derived. The logical principle that, if the models are accurate, then the conclusions must be accurate, is then used to conclude the test. For instance, if we wish to test an economic theory, utilizing a statistical test, we can

[11] Our usage will distinguish between theories and models. Theories will contain no data statements, only generalizations and logical relations. Models will refer to all cases in which theories are fit to specific circumstances. Any statistical test would represent one instance of a theory's being modeled according to this usage. Also, by this usage, those things normally called models in economics would be called theories unless they were being fit to a particular time or event.

visualize the testing process in the following manner. Call the economic theory E and the attendant data D. The inference to be tested we call I. We assert « If both E and D, then I_1 », or phrased differently, E and D imply I_1. The statistical theory, S, is then introduced along with the data and from it an inference is derived, an inference which has the same form as I. We assert that S and D imply I_2. For illustrative purposes let us presume I_1 is the assertion that quantity sold rises. Now if the inference from the statistical procedure is that quantity sold falls, we assert that some one or more particular statements included within either or both E and D are inaccurate. The purpose of the test, in other words, is to see if one can deduce that either E or D is false. Presuming care has been taken in collecting the data, we can deduce that E is false in some one or more respects.

A concrete illustration will show the applicability of the above outline to a particular case. This case, however, will be representative of many studies. We choose Robert Fogel's important railroad analysis.[12] Fogel wishes to test the historical interpretation which asserts that railroad construction and usage in the United States significantly contributed to subsequent industrial development. This interpretation is composed of a theory, data statements, lines of argumentation and conclusions. The historical model includes a number of statements; several lines were left out. See other margin for these lines which connect, among others, such things as railroads, corporate structures, population levels and distributions, development of financial markets, costs of transportation, and tonnage hauled with economic growth. The model is implemented by introducing data on costs, tonnages and the like. From the model the deduction is drawn that growth was materially affected. For emphasis one piece of data is isolated from the rest. That piece is, of course, the assertion concerning railroads. Fogel then proposes an economic model. This economic model is composed of statements which connect, among others, such things as railroads, costs of transportation and tonnage hauled with economic growth. The model is completed by adding data assumptions. From the model, conclusions are drawn concerning America's growth rate. Again one data statement is isolated from the remainder for purposes of emphasis. Again that data statement is one concerning railroads. The inference drawn Fogel claims to be contradictory with the inference drawn from the historical interpretation. This, Fogel feels, permits him to conclude the historical interpretation is false in one or more respects. This description of Fogel's study perfectly fits the general testing pattern set forth above. Moreover, Fogel's study is a paradigm for a multitude of similar studies in Economic History. What we may say about Fogel's work, therefore, applies, *mutatis mutandis*, to those other counterfactual studies.

Can Fogel derive the implication that the historical interpretation is

[12] Robert Fogel, *Railroads and American Economic Growth*.

wrong without having to assume the counterfactual? (For our purposes we will assume the counterfactual, the statement contradictory to fact, to be the assertion that railroads were not built in the United States before 1890). We symbolize his argument to simplify discussion.[13] Let his economic model be E, the historical interpretation H, the assertion that railroads were not built before 1890 $\sim R$, the deduction that growth was slowed $\sim G$, and the deduction that growth was not slowed G. We include all other data statements within the terms E and H. The conclusion Fogel desired to draw becomes $\sim H$, i.e., H is false.

Proving that $\sim H$ cannot be deduced from this argument without assuming the counterfactual is straightforward. We do this using the notions of heredity and independence. Our logic is hereditary with respect to assignments of truth and falsity.[14] Consequently if the premises are made true (by assumption), then any valid inference drawn from them must also be true. Any statement which can be false while the premises are true must be independent of these premises.

B

This portion of section I contains the proof alluded to above. In C, below, the proofs' implications will be discussed briefly.

Table I

1.	$[(H \cdot \sim R) \supset \sim G]$	Premise 1
2.	$[(E \cdot \sim R) \supset G]$	Premise 2
3.	$\sim H$	

---

13 Specifically, we shall use the system of natural deduction developed by Irving Copi in his *Symbolic Logic*. It should be pointed out that any logic based upon, in part, the traditional definition of material implication, and « free from self-contradiction » has « only one consistent and complete enlargement, namely the ordinary two-valued system of logic ». William Kneale and Martha Kneale, *The Development of Logic*, (Oxford at Clarendon Press, 1962), p. 574. In other words, we may consider this single system. For those who are interested in such matters, Copi specifically shows his system to be an alternative to the Hilbert-Ackerman system to the Principia Mathematica calculus of elementary propositions and to many others. See Chapter 8, Copi, *Symbolic Logic*.

We will use the symbols for material implication ($\supset$), for conjuction ($\cdot$), and for negation ($\sim$). These are defined as follows:

$A \supset B$	$A \cdot B$	$\sim A$
T T T	T T T	F T
T F F	T F F	T F
F T T	F F T	
F T F	F F F	

Implication reads « if A then B ». Conjunction means « both A and B ». Negation reads « not A ».

14 The Copi system is hereditary with respect to value assignments to the propositions in the arguments. See Chapter 7, Copi, *Symbolic Logic*. See pp. 187-190 for a preliminary discussion of completeness and independence.

Let the values of H, E, R and G be true. Then premises one and two are true and $\sim$H is false.

Table 1 proves $\sim$H independent from premises one and two. $\sim$H therefore can never be deduced from these premises without the assumption of other premises.

Can any assumptions be made, other than $\sim$R,[15] which permit $\sim$H's deduction? Firstly we do not want to assume $\sim$E, H or $\sim$H. It is not reasonable to assume $\sim$E. If we do the model supporting premise two is declared false and so premise two can be disregarded. If H is assumed correct, then it cannot be false. And the assumption of $\sim$H begs the question. This leaves E, R and $\sim$R, G and $\sim$G, premise one and premise two as reasonable additional assumptions. We shall use weak induction on the possible combinations of E, R, G, $\sim$G, one and two (as additional assumptions) to prove that no combination can be found which is not independent of $\sim$H.

We begin by noting we can write any possible set of assumptions using only the notions of conjunction and negation.[16] Also, premises one and two and the conjunction of them are true when E, H, and R are true and whether G is true or false. We use no other assignments in the proof so that one and two and the conjunction of them will always be true. We denote premises one and two by the symbols 1 and 2 respectively. Their conjunction is signified by A.

Case 1. We assume A and either R, G, $\sim$G, E, 1 or 2 but no more than one additional assumption in conjunction with A.

(*a*) We add either R, G, E, 1 or 2. Let H, R, G and E all be true. Then the conjunction of A (which is true) with a conjunct which is true will be true. $\sim$H is false.

(*b*) We add $\sim$G to A. Let H, R and E be true and G be false. Then the conjunction of A (which is true) with $\sim$G (G being false, G$\sim$ is true) will be true. $\sim$H is false.

Case 2. We assume A and n additional premises. We now add one more assumption to yield n + 1 additional assumptions. We may note first that it is not useful to assume both G and $\sim$G, so we presume if G is assumed then $\sim$G is not, and if $\sim$G is assumed then G is not.

(*a*) We assume A and n additional assumptions and we now add the assumption of either E, R, G, 1 or 2 but no more than one of these. Let

15 $\sim$ R is by assumption the counterfactual.

16 If we can prove $\sim$ H from premises 1 and 2 and combinations of E H, R or G, it can be done in a system using negation and conjunction alone. See Chapter 7 for a proof of this statement, Copi, *Symbolic Logic.* There Copi proves his system complete.

the value for E, R, G and H be true. Then if A· (n additional assumptions) is true, its conjunction with the additional E, R, G, one or two (each of which is true) will be true also. ∼H is false.

(*b*) We assume A· (n additional assumptions)· ∼G. Let E, R, H be true and let G be false (then ∼G is true). If A· (n additional assumptions) is true, then A· (n additional assumptions)· ∼G is true. ∼H is false.

This proves it is impossible to expand the argument by adding any combination of E's, R's, G's, ∼G's, 1's or 2's and produce a system from which one can deduce ∼H. In fact even the addition of ∼E or H does not enable one to infer ∼H. This can be seen by assigning a value of false to E and of true to H and following the above methods of proof. Generalizing further it can be shown that no statement implying something about growth can be deduced from premises one and two without the assumption of ∼R.

C

Two conclusions can be drawn from this discussion. Firstly, if we use the common logic (one with material implication) and if we employ counterfactuals in the testing situations described above, then we can eliminate them from our analysis only if we assume the counterfactuals. In the following section we see what this means for our analysis' rigor. Secondly, if we do not assume the counterfactual, then we can only infer the conditional statement (or something equivalent to it): « If the facts had been otherwise, then... ». While this may seem a small price to pay for the use of counterfactuals such is *not* the case; for from a logical point of view, as footnote 13 shows, a false antecedent (in this case ∼R) implies any conclusion whatsoever. Fogel, of course, contends there is more to the matter than the definition material implication. We also consider this argument in the following section.

Section II

Only two types of conclusions can be drawn from the argument set forth in Section I. Either the conclusion is hypothetical: « If the facts had been otherwise, then... » or the counterfactual has been assumed. If the conclusion is hypothetical, then no appeal to the facts in question can support the statement; it is true by definition alone provided only the antecedent is counterfactual. This type of conclusion appears quite weak. Will our deduction be stronger if we assume the counterfactual?

Analytical difficulties arise when the counterfactual is assumed. These difficulties stem from several properties of logical systems: consistency, heredity and completeness. Absolutely anything can be proven within an inconsistent system.[17] Such a system, therefore, is completely useless. Heredity means that if all premises are true, then all valid inferences are true. Accordingly we can freely add to *any counterfactual argument* any assumptions which we know to be true since the additional assumptions will lead us to no false inferences. In particular, taking the above example of railroad histories, we may assume that railroads were built in the United States before 1890. But this assumption makes our system inconsistent since we have assumed both $\sim$R and R. Consequently our analysis becomes worthless.

Undoubtedly the counterfactualist would resist making his analysis inconsistent. But the difficulty of preventing inconsistency seems to be underestimated, particularly if the analysis is complete.

Which assumptions can produce inconsistency? Certainly more than R. Fogel's study concerns itself with the railroads' influence upon the United States' growth. Growth rates are estimated in terms of factor savings associated with a shift in the transportation production function. Many assumptions about society in 1890 are needed before the estimates can be derived. For instance, Fogel needs to know the population, its distribution, the structure of financial markets, corporate structures and the like. If the study is to be consistent these assumptions must not be contradictory with the historical analysis' implications. The historical analysis implies consequences other than cost reductions. It suggests, among other things, that without railroads population growth would have been slower, that financial markets would not have developed as rapidly, that corporate structures would have been more primitive in 1890, and that western settlement would have been delayed.[18] Consequently if the counterfactual study assumed the 1890 values for these variables, the study becomes inconsistent because it is possible to derive contradictory values from the historical model.[19] The analytic

---

17 Using an inconsistent system it is easy to prove, for instance, that the moon is made of green cheese. Take as the inconsistent system the two assertions that railroads were built in the United States before 1890 and that railroads were not built in the United States before 1890. If it is true that railroads were not built (we make this true by assumption) then it is also true that either railroads were not built or the moon is made of green cheese. But if it is false that railroads were not built (again by assumption) then it must be true the moon is made of green cheese.

The criticism concerning consistency developed in this section represents an adaptation of Nelson Goodman's criterion of cotenability, NELSON GOODMAN, *Fact, Fiction, and Forecast,* (Bobbs-Merrill Company, 1965), Chapter 1, passim.

18 For one discussion of railroads' influence upon western settlement see ALBERT FISHLOW, *American Railroads and The Transformation of The Ante-Bellum Economy,* (Harvard University Press, 1965), Chapter 4, pp. 163-204.

19 That the study becomes inconsistent is easy to see. Assuming the counterfactual, the historical model yields a population level below that actually observed in 1890. When the actual 1890 population is assumed we have two contradictory assertions in the

difficulties can be seen in an analogy. If we tested the correctness of a fiscal policy which involved a substantial deficit and if the analysis assumed no deficit and simultaneously assumed the investment and consumption expenditures which actually occurred, the test then probably would show the policy to have been unnecessary. It is also likely the test would be unacceptable. The analogy with any other counterfactual test is exact.

It becomes clear that guarding against inconsistency in counterfactual analysis is at best a difficult problem. At worst it is even more serious. In particular, the historical analysis may be complete. If it is, then every statement which can be written in the system's terms is either a valid inference or its negation is a valid inference. It should be pointed out that completeness is a desirable property since it permits one to prove that one and only one solution to a particular problem can be deduced from the system. But completeness means that the range of inconsistent assumptions can become greatly multiplied. Clearly the problem is more serious than the simple question of whether or not to assume R.

Nevertheless, it may be felt that as practical men we must overcome our hesitant, tentative doubts concerning the premises and concentrate upon the conclusions instead. We must strike at the heart of the matter and for the moment forget the peripheral issues. We may take this meaty conclusion to assert, in the case developed above, that the historical interpretation was wrong. Footnote 20 shows that given the assumption of $\sim R$ we can deduce $\sim H$. But this proof deserves careful attention on at least two counts. Both are related to the counterfactual's presence.

H represents the historical model on the one hand and an array of data statements on the other. Footnote 20, given the premises, proves that *one*

---

model. These two contradictory assertions make the analysis inconsistent. Consequently it can not lend plausibility to the analysis of which it is a part.

Fogel appears to see the consistency problem in part. To lend his analysis greater plausibility (to make it more consistent) he projects a canal system for the United States which could have been built in the absence of railroads. He also restricts the cultivated acreage. Our analysis of consistency is meant, in part, to clarify the precise basis for his felt need to incorporate canals. It will also provide an exact criterion for those changes in data assumptions which must be made to preserve consistency.

[20] We use a variant upon a *reductio ad absurdum* argument. What we show in Section I below is that statements one to five are inconsistent. In other words not all of them can be true simultaneously. Attention is directed towards statement five, which asserts that the historical interpretation is correct. We may ask whether statements one to five could all be true simultaneously if we changed five to read « The historical interpretation is false ». Section II of this proof shows that if we change statement five to read « The historical interpretation is false », then the set is no longer inconsistent. We prove consistency of this set of statements by finding some proposition that is not provable within the system. We chose as our proposition $\sim G$ and we show that statements one to five can be true with a certain assignment of truth values to the terms, while with that same assignment $\sim G$ is false. This proves consistency, for

or more statements in H are contradictory (this is the meaning of proving H false). But this is not news. For we have already seen that the 1890 values for population and a variety of other factors are inconsistent with the implications of the historical interpretation in which $\sim R$ is assumed. Since we knew to begin with that some statements in the study were inconsistent, and since this is all the proof permits us to assert, the study therefore permits us to deduce absolutely nothing about the truth or falsity of the interpretative framework, H.[21] This analytical paucity must always exist so long as it has not been shown that *all* data employed in the computations are consistent with the counterfactual's assumption. In no counterfactual study published thus far, which aimed to test a historical interpretation, has this requirement been met. None, therefore, permits one to deduce that the historical model itself is false.

But let us assume all data statements are consistent with $\sim R$'s assumption.. Can we now deduce that the historical model is false? The first section

recall that if a set of premises is inconsistent, then absolutely anything can be proven from them. The proof is as follows:

Section I

1. $[(H \cdot \sim R) \supset \sim G)]$
2. $[(E \cdot \sim R) \supset G]$
3. $\sim R$
4. E
5. H

If these five statements are consistent then it will not be possible to assume all premises true and at the same time give some assignment of values to the propositions within them such that one or more propositions will have to be simultaneously assigned a value of both true and false. We give such an assignment here. If all premises are true then in particular three, four and five must be true. If they are true then E and H must be true and R must be false. If E is true and R is false then $(E \cdot \sim R)$ is true. If premise two is true and $(E \cdot \sim R)$ is true then G must be true. If premise one is true and if G is true then $(H \cdot \sim R)$ must be false. But if R is false and $(H \cdot \sim R)$ is false, then H must be false. However, we have already seen that H must be true if premise five is true. It should be noted that using the same method, but going from premises three to five, next to premise one and lastly to premise two we can prove that only by assigning contradictory values to E can we have all statements true simultaneously. Alternatively, assigning contradictory values to R would permit all premises to be true simultaneously.

Section II

Change statement five to read $\sim H$. Then posit $\sim G$ as another statement. We wish to determine if $\sim G$ is independent of our new statements one through five. Let E and G be true and let R and H be false. These assignments make statements one through five true. But $\sim G$ is false. This proves that statements one through five, as changed, are now consistent. It also proves we may deduce $\sim H$ from statements one through four.

[21] This, of course, is the reason that Null Hypothesis tests of counterfactuals yield no new information.

of footnote 20 shows the initial premise set is inconsistent. The second section shows a somewhat different premise set is consistent. This permits us to assert H is false. But the proof's second section could have been constructed in a very different manner. First substitute $\sim$E for $\sim$H in statement five and then substitute H for E in statement four. Now with the appropriate values assigned to the propositions we have proven $\sim$E. This proof is based upon *exactly* the same counterfactual study as our proof of $\sim$H. Evidently, therefore, our ability to deduce either $\sim$H or $\sim$E is simple a matter of whether we decide to assume E or H.

Usually such a decision is made upon the basis of a Null Hypothesis test. But as we have seen, and as was pointed out by Bear and Orr,[22] counterfactuals cannot be usefully tested this way. The alternative is an appeal for inductive support. But this appeal may not solve the difficulty, for it is at best rather ambiguous. And in the present case the circumstances are not the best, since the models come from two separate fields. These points notwithstanding, the temptation to assume E may prove nearly insuperable. But before this step is taken one should reflect upon the reception Friedman and Meiselman [23] would have received had they set out to prove the inadequacies of investment-multiplier analysis by assuming at the outset that if the two modes of analysis were inconsistent, then the investment-multiplier analysis must be in error while the monetary-velocity analysis must be correct. A moment's reflection will convince one that such a line of argument would be totally unacceptable.

---

[22] For an extended discussion of this point see D. V. T. Bear and Orr, *Logic and Expediency*. This, of course, is the reason statisticians are so disturbed by Null Hypothesis tests which assume, for example, a normal distribution when the event is not distributed normally. In such cases all a Null Hypothesis test will show, no matter how beautifully it turns out, will be that one or more of the initial assumptions was not true, and again in such cases we already knew that much.

[23] M. Friedman and D. Meiselman, *The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958*, « Stabilization Policies », ed. D. G. Brown et al, and the study of Alberto Ando and Franco Modigliani, *The Relative Stability of Monetary Velocity and the Investment Multiplier*, « American Economic Review », Vol. 55, No. 4, Sept., 1965, pp. 693-728.

Among other things this example points out the extreme difficulty in obtaining agreement on the inductive support for one theory as opposed to another, even within the same field. Where theories come from two different fields, as is the case in econometric tests or in counterfactual tests of historical interpretations by economists, the problem is even more severe. In these latter instances we would expect historians to be much more familiar with inductive support for their theories while economists would be much more familiar with the support for their theories. Similarly, we would expect the statisticians to be more familiar with the support for their theories. It is worth mentioning that the deductive validity of a theory provides absolutely no inductive support for that theory. The logical validity of a statistical theory or an economic theory, consequently, has no bearing upon whether or not it is useful in a particular application where the question of inductive support is the decisive consideration.

Recapitulating the argument, we have seen that a historical interpretation can be shown to be false (inconsistent with the results of a counterfactual study) only if we assume the counterfactual. But if the counterfactual is assumed, then firstly grave problems of consistency arise and secondly it is no longer useful to develop a Null Hypothesis test. These are not trivial shortcomings. To disregard them in using counterfactuals to deduce the falsity of an interpretation must lack persuasiveness.

Can we circumvent these problems by using the counterfactual in a hypothetical sense? Section I and footnote 20 show that, if you assume the counterfactual, you can prove that two interpretations (with the attendent data) are inconsistent. They also showed that this cannot be done if you refuse to assume the counterfactual. To retreat to the hypothetical use of the counterfactual means, therefore, that the study can no longer even prove inconsistency.

The difficulties with counterfactual analyses do not necessarily make counterfactual studies inappropriate nor do they make them unenlightening. But they do indicate that the methods of reasoning open to the counterfactualist are more narrow than those open to the non-counterfactualist.

Section III

Each of the difficulties discussed above was associated directly with one or more features of a logical argument. Means for getting around some of the difficulties can be specified within the logical framework.

It has been said that data used in a counterfactual study must be consistent with the counterfactual's assumption.[24] But each such discussion has begged two important questions, which are: which data statements must be adjusted, and, how must the adjustments be made? According to our analysis, data adjustment is needed to prevent an inconsistent argument. The argument's structure, therefore, tells us exactly which data must be adjusted and how they must be adjusted. From a logical point of view this is the restriction's extent.

While this establishes exactly which data must be adjusted, the adjustment process often appears ambiguous because the historical interpretation to be tested has been inadequately specified. I am not referring to the problems discussed by Fogel[25] which concern translating a portion of the historical thesis into economic terms, but rather to an appreciation of the entire historical

[24] Gould for one, discusses this point, J.D. Gould, *Hypothetical History*, «Economic History Review», SS, Vol. 22, No. 2, Aug., 1969, pp. 195-207. And, as pointed out in footnote 18 above, Fogel appears to have some concern with this problem.

[25] Robert Fogel, *Specification Problem.*

interpretation. An example of inadequate specification of the type we are discussing here is mentioned in Section II above where Fogel's use of population data is criticized. Clearly, in principle the adjustment process in such cases is relatively straight-forward; the difficulties arise because inadequate attention has been paid to clear specification. These comments also hold in those cases in which the historical interpretation is itself less than clearly developed by historians.

Another contentious issue is whether large and small counterfactuals are equally satisfactory. Gerschenkron, for instance, asserts that small counterfactuals provide more reliable conclusions than large counterfactuals.[26] Without introducing many-valued logics, our analysis suggests both are equally satisfactory in principle, but that the small or narrow counterfactual may be much more satisfactory in practice.

In accord with common usage, define a narrow counterfactual as one whose range of implications is restricted, and a broad counterfactual as one whose range of implications is much less restricted. It should therefore be easier to trace a narrow counterfactual's implications than those of a broad counterfactual If the implications can be traced more easily, the problems of adjusting the data to prevent inconsistency should be easier. In this sense narrow counterfactuals are safer than broad ones. But it does not follow that narrow counterfactuals would be less interesting, because we have said nothing about the magnitude of the individual changes. It is entirely possible that very important changes in consequences would flow from a narrow counterfactual. In other words, the magnitude of these changes could be of intense interest.

A third point concerns the results of a counterfactuals study. We have seen that inconsistency between two interpretations can be proven. We have also seen that heroic assumptions are required before the falsity of the alternative hypothesis can be proven. This leaves the question of what we may do with our results.

Traditionally, evidence of inconsistency between interpretations has led to attempts at reconciliation. For example this is generally done when different GNP estimates are found for some remote period. Reconciliation is generally sought through data investigations. If data errors are found, it is often possible to reconcile the two estimates. An alternative procedure is to study the assumption's reality. Obviously, when using a counterfactual some limitations upon this line of inquiry must be imposed.

Nevertheless, it is a poverty-stricken hypothesis which has but one premise. If we carry the investigation to a more theoretical (as opposed to empirical) but seldom used level, we may investigate the formal properties

---

26 Alexander Gerschenkron, *The Discipline and I*, « The Journal of Economic History », Vol. 27, No. 4, Dec., 1967, pp. 443-459.

of the basic premise sets, trying to find portions of the models which are equivalent. The different outcomes around which the discourse revolves may then be attributed to the sets of premises which are not equivalent. If such sets are small, we may have learned something quite worthwhile about the nature of our theories, and, hopefully, about the world. Useful conclusions, then, obviously may be derived from counterfactual studies which find differences in implications between two or more models.[27]

Still it should be noted that there is a great deal more work involved in the counterfactual study. Since no Null Hypotheses can be tested, the study's interest must reside in the divergence between inferences that arise. As has been suggested these differences gain their interest from a discussion of the reasons for the divergence. This means that when such differences have been found, there remains the task of explaining them. The noted difference then becomes the intermediate step in the analysis.

Section IV

As stated in the introduction, no attempt has been made to discuss the question of whether counterfactual investigations not aimed at the refutation of some other hypotheses are interesting and informative. The topic has remained the value of the counterfactual in theoretical investigations. Our argument supports the contention that counterfactuals may be all right for practical purposes, but that they do not work well in theory. Specifically, where theoretical problems are concerned, the counterfactual approach is clearly more cumbersome than the non-counterfactual; it requires much more delicacy of insight to evade problems of inconsistency; it does not admit of Null Hypothesis testing; and it requires, to provide interesting results, extended analyses of reasons for differences in conclusions.

Moreover, while our illustration has been drawn from Economic History, the argument is perfectly general in application. In particular the discussion applies to econometric studies as well. It must be emphasized that the formal logical validity of these models — be they historical or statistical — is distinct from the question of whether in a particular application they are or are not counterfactual. The question at issue here is simply whether the assumptions resemble the reality to be described. And there is not more reason to single out the existence or non-existence of railroads as questionable than to question assumptions in other analyses. Whether those firms did seek to maximize

[27] It goes without question that these same differences in implications could be found without extensive data sources and traditional methods of historical investigation. This being the case there is some room to question the efficiency of many so-called cliometric studies.

profits, whether that production function was Cobb-Douglas, whether the variance was homoscedastic are as relevant to the consistency-counterfactual problem as the existence of railroads or the nature of colonial government after 1763. In all cases where inconsistency arises, whether in Economic History or in other fields, the logical nature of the argument fails to lend the degree of plausibility given by a consistent argument. This, at heart, must be the nature of Leontief's call for more work with primary sources and for less reliance upon elaborate estimation procedures within the consistency problems multiply, sometimes, beyond all reasonable bounds.[28]

[28] W. Leontief, *Theoretical Assumptions and Nonobserved Facts*, « American Economic Review », Vol. 61, No. 1, March, 1971, pp. 1-8.