A Philosophy of Science Compatible with Clarke and Primo (2012) and Ashworth et al. (2021)

Nicholas Ray

June 11 2025

1 Clarke and Primo's (2012) Critique

Clarke and Primo (2012), henceforth CP, primarily claim that 1) theories cannot be true or false and that 2) empirical social science uses models of data to test hypotheses. These two claims, they argue, render the classic logic of falsification as an inappropriate justification of social science. Consider the following truth table:

A	B	$A \rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

Table 1: Truth table for *modus tollens* ($\rightarrow$).

CP show that most political scientists envision A to represent a theory while B is the hypothesis used to test the theory. The logic of *modus tollens* is powerful because it demonstrates that the only unique information we can attain about the truth value of a theory (A) is when the theory suggests a hypothesis (i.e., $A \rightarrow B = \text{T}$) but the hypothesis is rejected (i.e., $B = \text{F}$). In this case, highlighted above, the theory can be concluded to be false. If the theory does not actually imply the hypothesis or a researcher fails to reject the hypothesis, a definitive conclusion about the truth value of the theory cannot be reached. Thus, one justification for social science relies on the logic of falsification, which says that social scientists should be in the business of "falsifying" theories by systematically evaluating their hypotheses. If a theory's hypotheses are often found to be rejected, we can logically conclude that the theory is not true.

Briefly, CP argue that this is nonsensical. Theories (or models) are composed of assumptions or premises that deductively lead to conclusions. At best, if the assumptions are true and conclusions from the theory sound (i.e., $A = \text{T}$ and $A \rightarrow B = \text{T}$), then testing a hypothesis implied by the theory is redundant. It simply must be the case that $B = \text{T}$. At worst, which is more common, a theory comprised of false assumptions (i.e., $A = \text{F}$) is impossible to learn about. Ultimately, since theories are models, it makes little sense to test them in the manner suggested by the logic of falsification. Since they are constructed by researchers for a particular purpose, it should only be asked if theories are useful for that purpose. Further, because "models of data" are typically constructed to test hypotheses, it similarly makes little sense to assess the truth value of a hypothesis (B).

2 What Is There To Be Done?

On both counts, CP are correct. A theory or model should not be viewed as true or false and we should also recognize that hypothesis testing relies on models of data. However, I argue that it is premature to conclude that testing theories is nonsensical. Furthermore, clarifying what it means to test a theory or model allows us to incorporate recent advances in research methods that link models of data to theoretical models, namely the contributions of Ashworth et al. (2021) (henceforth ABBdM).

I suggest that what we mean by "testing" a theory or model is not evaluating its truth value. As CP have already elaborated, a theory or model is a collection or system of deductively valid statements that may even feature unrealistic assumptions. What we really want to know is 1) how much *evidence* (defined below) is there that our theory actually explains what is going on in the world and 2) how does this evidence compare to the evidence for other theories explaining the same thing? This is the core problem of science: what is the evidence for our theories?

As demonstrated in Table 1, *modus tollens* provides a convincing logic for which to

evaluate the evidence for theories. However, the issue with using *modus tollens* is in defining what A and B are, and I argue that this is why CP are premature to conclude that the logic of falsification is nonsensical. An entire theory or set of models is not the logical statement captured in A that is being evaluated. Rather, A represents the statement "the implication of my theory accurately describes the real world," which does have a meaningful truth value. This also makes it more clear what B represents, which is "the implication of my theory agrees with a commensurate empirical quantity." Thus, as shown in Table 2, I suggest that the logic of falsification relies on the conditional statement "if the implication of my theory accurately describes the real world (A), then it should agree with a commensurate empirical quantity (B)" ($A \rightarrow B$).

A "the implication accurately describes the real world"	B "the implication agrees with a commensurate empirical quantity"	$A \rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

Table 2: Truth table using *modus tollens* to "test" a theory.

If this appears to be an argument over semantics, it should be of no surprise since CP conclude in chapter 3 that their contribution may be a change in language, although one that has important implications for how we conduct science. Similarly, I argue that the updated truth table in Table 2 has the important implication of allowing us to continue to use *modus tollens* while reconciling CP and ABBdM.

To see this, firstly note that we are no longer interested in the truth value of a theory or model, which is indeed nonsensical as already discussed at length. A now better captures the idea that we are interested in if the behavioral implications of our theory actually correspond to real world changes in behavior, regardless of how simple or "false" our theory or model may be. Secondly, we are now explicitly acknowledging the fact that models of data are used in empirical social science through the definition of B. In order for an empirical quantity

to be commensurate with a theoretical implication, the research design or empirical model must target the same theoretical target with enough similarity (see ABBdM). Indeed, the more similar the research design and target, the more likely it is that the evaluation of B is a true positive or a true negative. Relatedly, the more similar the theory and target, the more likely it is that A is true. However, we are not typically interested in theories that maximize target similarity (e.g., a life-sized map), which is why there is uncertainty over the real world accuracy of our theoretical implications (i.e., the need for science).

Where does this leave us for assessing the evidence of a theory? Based on the above discussion, I suggest that a logically valid definition of the evidence for a theory is:

$$evidence(T) : -\sum_{i=1}^{n} \mathbb{1}\Big((A \rightarrow B)_i = \mathrm{T} \;\&\; B_i = \mathrm{F}\Big) \;\forall\; i = 1, \ldots, n. \tag{1}$$

In words, the evidence for a theory (T) quantifies how many n implications of a theory do not accurately describe the real world according to the logic of falsification in Table 2. The quantity is negated because the more that A is falsified the less evidence there is for a theory. This quantity also allows for the comparison of evidence between theories. If two theories explain the same thing, the theory that has fewer implications that fail to accurately describe the real world has more evidence in its favor. Thus, the problem of science that I laid out earlier may be defined as a maximization problem, where we are trying to find the theory that has the most evidence out of the set of theories that explain the same thing ($\bar{T}$):

$$\max \;\; evidence(T \in \bar{T}). \tag{2}$$

References

Ashworth, Scott, Christopher R. Berry, and Ethan Bueno de Mesquita (2021). *Theory and Credibility: Integrating Theoretical and Empirical Social Science*. Princeton University Press.

Clarke, Kevin A. and David M. Primo (2012). *A Model Discipline: Political Science and the Logic of Representations*. Oxford University Press.