**Jasmin Hasić***

# IDENTIFYING AND TESTING CAUSAL MECHANISMS BY APPLYING J.S. MILL'S INDIRECT METHOD OF DIFFERENCE

## ABSTRACT

This paper is aimed at investigating Mill's ideas of causation, with particular reference to his Indirect Method of Difference. More specifically, I will try to inspect *whether an application of Mill's Indirect Method of Difference yields a positive result that allows a strong inference of a causal relationship being present*; and if so, *does one still need to specify a causal mechanism and test it with different cases*?

In the following pages, I will offer a comprehensive interpretation of Mill's basic methods, and argue that his Indirect Method of Difference is intended to ascertain the presence of a probable cause, within a set of all positive cases and the absence of it within the set of all negative cases. I deem that such condition is very rigorous, and that Mill's method is only hypothetical in identifying the presence of both necessary and sufficient conditions, as the cause of a certain phenomenon.

**Keywords:** John Stuart Mill, Indirect Method of Difference, Method of Agreement, causation.

---

* LL.M, doktorski kandidat na Fakultetu političkih nauka Univerziteta LUISS Guido Carli u Rimu.

## INTRODUCTION

John Stuart Mill is a transitional thinker, one of the greatest philosophers whom England has produced since Hume[1], advocating for both utility maximization and personal autonomy as highest principles of morality. In his master piece, *A System of Logic*[2], Mill puts forward his outlook on "*the principles of evidence and the methods of scientific investigation*". In keeping with his philosophical thought, social sciences ought to be founded on *"the laws of nature of individual man"*, since prevailing features of societal life are the actions and passions of human beings united together in the social state[3]. These actions and passions are seen as *"obedient to the laws of individual human nature"*, as they encompass "*laws of mind*" and "*laws of the formation of character*", thus making the logic of moral sciences a key factor which allows an individual to classify and recognize these laws[4]. All phenomena in the society are thought to be phenomena of human nature, and according to Mill's theory, they are grouped by the action of external conditions and a sufficient amount of societal actors. Consequently, the observable fact of human thoughts, actions, feelings and other emotions can be considered as a subject of unchanging laws, which are meant to conform with fixed laws as the end result of their prior relations. Finding and uncovering "*these laws*" is ultimately "*the object of the social science*". Once they are attained and understood, they will help scientists "*to explain and predict the whole history of society*", yet with timeline limitations of "*thousands of years to come*".

Mill argues that "*by identifying causal laws and conditions in which they operate*", the explanation and prediction of the content of social reality are, in fact, accepting the fact that actions may be completely caused and thus completely predictable. Free action is not a randomly occurring act or an act whose explanation is beyond our knowledge; free action is possible, but freedom also assumes a causal order and is in tune

---

[1] H. Sidgwick, *John Stuart Mill*, The Academy, 1873, 193.
[2] J.S. Mill, *A System of Logic*, Ratiocinative and Inductive, 1843.
[3] J. Barnes, *The Presocratic Philosophers*, Routledge, 1982, 105.
[4] J.S. Mill, *A System of Logic*, Ratiocinative and Inductive, 1843.

with deterministic laws. Epistemologically, Mill upholds *an empiricist outlook* that "*knowledge is a matter of experience*"[5]. His notion of free action is not related to a random occurrence or an action beyond our understanding, but rather to a compatibility of freedom and determinism, which presuppose a causal order. His views are also in support of naturalism. Consequently, Mill believes that characters of human beings in society "*are derived from, and may be resolved into, the laws of nature of an individual*", which entails a particular and unique logic of scientific explanation for all social phenomena.

Mill is deeply involved in idea of causality. His interpretation is often concerned with specific mechanisms of functioning within particular historical settings, where "*necessities are not only those of universal laws but also those of particular productive forces and their functioning*". Two of Mill's scientific methods are of significant bearing to all case-oriented studies: the Method of Agreement and the Indirect Method of Difference.

Hereafter, this paper is aimed at investigating Mill's ideas of causation, with regards to his Indirect Method of Difference. More specifically, I will try to investigate *whether an application of Mill's Indirect Method of Difference yields a positive result that allows a strong inference of a causal relationship being present*; and if so, *does one still need to specify a causal mechanism and test it with different cases*?

In the following pages, I will offer a comprehensive interpretation of Mill's basic methods, and argue that his Indirect Method of Difference is intended to identify the presence of a probable cause, within a set of all positive cases, and the absence of it, within the set of all negative cases. I deem that such condition is very rigorous, and that Mill's method is only hypothetical in identifying the presence of both necessary and sufficient conditions as the cause of a certain phenomenon.

---

[5] R. Carnap, *Intellectual Autobiography*, in P. Schilpp, ed., *The Philosophy of Rudolf Carnap,* Berkeley University of California Press, 1967, 47.

## 1.    Mill's Methods and their Conceptualization

Initially, John S. Mill sought to find solutions to Hume's problem of induction[6]. Namely, the analysis of event causation involving regularity (also known as *nomological approach*) is related to the notion that all associated events are set within a general rule of consistency. In other words, if *x* is the cause of *y*, there has to be other type of events, A and B, such that *x* is of kind A, x is of kind B, and events of kind A are regularly followed by events of kind B[7]. Many authors regard this occurrence to be a "*constant conjunction*" of the two event types involved in the equation; others believe that such "*nomological necessity*" implies the existence of a "law" or even a "causal law".[8]

In resolving the nature of regularity analysis and "*subsumptive*" laws, Mill departed from a premise that an explanation can only occur "*if causal laws and the conditions in which they operate are previously identified*". Mill's causal images and model of causality, in that sense, are of a specific type, working within particular historical conditions. In his view, an individual fact can be explained by identifying its cause; namely, by affirming the *law or laws* of causation of its origin or *"a law or uniformity in nature is said to be explained when another law or laws are pointed out, of which that law itself is but a case, and from which it could be deduced*"[9]. Both Hume and Mill argue that for any generalization originating in experience can never be sufficient to eliminate all possibility of doubt. However, as suggested by Hollis[10], such events are not - or not exclusively - emerging of general and universal laws, but rather from particular *productive forces* and *their effects*.

---

[6] A. Rosenberg, *Philosophy of Science*: A Contemporary Introduction, Routledge, 2011.
[7] D. Hume, *A Treatise of Human Nature.* L. Selby-Bigge Ed. Oxford: Clarendon, 1973.
[8] Cambridge Dictionary of Philosophy, 2nd edition, Cambridge University Press 1999, s.t. causation.
[9] J.S. Mill, *A System of Logic*, Ratiocinative and Inductive, 1843.
[10] M. Hollis, *The Philosophy of Social Science: an Introduction*, Cambridge University Press, 1994.

The initial, original and the most straightforward of Mill's scientific methods is the *Method of Agreement*, generally regarded as an inferior technique, often leading to faulty empirical generalizations.

> "*If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon*"[11].

In a nutshell, if two or more instances of a certain investigated phenomenon concur with one of the several possible causal circumstances, there is a rational assumption that they might be the cause of the observed phenomenon. The *Method of Agreement* directs to the laws of phenomena which also do not necessarily have to be laws of causation, or where the subject of causation has not yet been resolved or determined. The *Method of Agreement*, especially in comparative social science and case-oriented investigations, proceeds by simple elimination.

For example, if a number of people who are suffering from a certain disease have all been without an access to certain vitamins, but have been keeping different diets, been in different conditions, have had different hereditary backgrounds etc., implying that the lack of these vitamins is the only characteristic common to all of them, then one can conclude that the lack of vitamins can be understood as the cause of this particular disease[12]. The set of instances *A B C, A D E or A F G*, suggests that *A* is the factor which could be considered as a foundation for determining a certain effect observed; or at least as an inferior resource. In other words, the cause can be identified as a common feature within a number of different cases when a certain effect takes place; or in other instances, when the cause of a perceived outcome can be understood as the one where the observed effect is different from of another comparable case when the same effect did not happen.

---

[11] J.S. Mill, *A System of Logic*, Ratiocinative and Inductive, 1843.
[12] J. Mackie, *The Cement of the Universe: A Study of Causation*, *Eliminative Methods of Induction*, 1980, available online:

> "*If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon*".[13]

Similar to the *Method of Agreement*, Mill's *Method of Difference* is designed to establish models of relationships between certain observed causes and effects of certain occurrences. It is to be observed as the investigation of certain subjects where artificial experimental design is seen as unrealistic. Thus, the Method of Difference generally accounts for more efficient social processes, which can establish causes or even laws.

The two instances to be compared have to be similar in all circumstances, except for the one which has to be investigated. Strictly speaking, there needs to be a relationship between *A B C* and *B C*, or between *a b c* and *b c*. In this case, the preexisting similarity of circumstances does not have to be extended to those conditions that are already known or can be completely irrelevant to the result.

Mill's *Indirect Method of Difference* (also known as the *Joint Method of Agreement and Difference*) "arises by a double employment of the *Method of Agreement*"[14]; where each *proof* being used is to be independent of the each other, but of mutual support. In Mill's view, this method can only be understood as an extension and/or an advancement of the *Method of Agreement*, but is still not as convincing as that of *Difference*, due to the fact that the *Method of Difference* is a substitute for an artificial experiment, while the Method of Agreement is only used where experiment is impossible.

> "*If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in*

---

http://www.philoscience.unibe.ch/documents/kursarchiv/WS99/mackie-appendix.pdf (accessed in Jan. 2015).

[13] J.S. Mill, *A System of Logic*, Ratiocinative and Inductive, 1843.

[14] *Ibid*

> *which it does not occur have nothing in common save the absence
> of that circumstance, the circumstance in which alone the two sets
> of instances differ is the effect, or the cause, or an indispensable
> part of the cause, of the phenomenon*"[15].

Noticeably, by "*double application*" of the *Method of Agreement*, one might expect that both positive and negative occurrences can be established, which the *Method of Difference* itself requires. However, the *Indirect Method of Difference* is designed to complete the requirements of the *Direct Method of Difference*, since "*the requisitions of the Method of Difference are not satisfied unless we can be quite sure either that the instances affirmative of x agree with no antecedents whatever but A, or that the instances negative of x agree in nothing but the negation of A*"[16]. Hence, the Indirect Method of Difference fulfills the requirement by which all positive instances should have "*nothing in common save the absence of that circumstance*", and that all negative instances should "*have nothing in common save the absence of that circumstance.*"[17]

## 2. Theoretical Principles of the *Indirect Method of Difference* and its Application

Both the *Method of Agreement* and the *Indirect Method of Difference* are insufficient and entail the notion of probability when considering cases where the number of possible causes may be greater than 1.[18]

So, if one is to examine a set of observations, where a certain factor *x* is the subject of interest, and if one assumes that there is no preceding cause of *x* other than *A*, by using the *Method of Agreement* one reaches a conclusion that certain factor *A* is the cause of the result exerted by *x*. In an ideal environment, this conclusion could be validated by setting an experiment, where value based manipulation of *A* could yield an evidence

---

[15] *Ibid*

[16] *Ibid*

[17] B. van Heuveln, *A Preferred Treatment of Mill's Methods: Some Misinterpretations by Modern Textbooks*, Informal Logic, 2000.Vol. 20, No. 1.

[18] J. Sekhon, *Quality Meets Quantity: Case Studies, Conditional Probability and Counterfactuals*, Perspectives on Politics, 2004, available online: *http://www.nyu.edu/g-sas/dept/politics/faculty/beck/sekhon.pop.pdf* (accessed in January 2015).

of effects it has on the *x*, or to illustrate whether the same effects take place when the factor *A* is missing. Since conducting such an experiment is not possible, the double application of the Mill's *Method of Agreement* remains a credible option. In this case, one has to assume that there must be another set of observations. Hypothetically, one can safely conclude that, within the theoretical premises of the *Indirect Method of Difference*, *A* could be the cause of the *x*.

For that reason, in all cases where the presence of an *incident* leads to an occurrence of another one and the absence of an *incident* leads to the absence of entirely another incident, one can conclude that there is a causal connection between the two observed phenomena.

<div align="center">

For a Positive Set
A B C are followed by a - b - c
A B D are followed by a - b - d
A C D are followed by a - c - d
For a Negative Set
B C D are followed by b - c - d
C D E are followed by c - d - e
B E F are followed by b - e - f

</div>

According to Ragin[19], Mill's *Indirect Method of Difference* has three distinct phases of application: double application of the *Method of Agreement* ("*the cross-tabulation of cause and effect*") and a third phase which involves the "*rejection of competing single-factor explanations through paired comparisons*". The major difference between the *Method of Agreement* and *Indirect Method of Difference* is the utilization of negative cases with the aim of emphasizing the conclusions that were previously drawn from positive cases. Thorough examination of negative cases allows the researcher to identify a set of observed factors, which encompass the phenomenon of interest. For this reason, the principle of elimination in both these methods is applied to eliminate the accidental factors and to find out the genuine cause, i.e. what cannot be eliminated without affecting the phenomenon has to be, in principle, causally related

---

[19] C. Ragin, *The Comparative Method*, Berkley University of California Press, 1987.

with it. This definition of the set, however, should not be a part of known hypothesized causes or instances of the effect.

In this sense, Ragin states that, while this closer approximation of experimental design is scientifically more applicable, it suffers from some of the same limitations as the *Method of Agreement* in situations of "*multiple causations*".

For instance, if a researcher wants to observe cases of *land hunger*, s/he would find a correlation between land hunger and social/ political revolts. However, by using the second phase of the *Indirect Method of Difference,* one might observe "*that land hunger may not be the cause of revolts because the rapid commercialization can also cause revolts by itself*", in the absence of land hunger; thus, there are situations where the absence of land hunger can be connected with social revolts. Parallel study of rapid commercialization would also lead to parallel dilemmas: "*are there any revolts caused by land hunger and in the absence of rapid commercialization*"? In other words, a cross-tabulation of the results, with either of the causal variables, would independently lead to rejecting of both variables as the single causes of the effect. Such display of results would match up with "*perfect zero-order statistical correlation*" and with established "*patterns of invariance*". The previous notion has a couple of implications. The most important is that Mill's deterministic methods must have no measurement error in order to work plausibly. Also, it is a noteworthy to mention that the type of paired comparison used in the indirect *Method of Difference* can be seriously weakened by *conjunctural causation*[20]. Namely, if one presumes that social revolts occur only when both land hunger and rapid commercialization concur, and that all cases of land hunger can also be considered as the instances of rapid commercialization, but not the other way around, then the researcher would be completely entitled to believe that the land hunger alone causes revolts. All the data available would support this conclusion.

All cases of land hunger would also be examples of revolt, and all cases where the land hunger is absent would show no revolt. In such case,

---

[20] *Ibid*

the *Indirect Method of Agreement* (the 3rd phase) would then allow one to reject the "rapid commercialization" as a cause of revolts mainly because some examples of the absence of revolts show "*rapid commercialization without land hunger*". To put it plainly, a matching comparison with a positive instance (where land hunger and rapid commercialization are paired to produce a revolt) and with a negative instance, (where rapid commercialization with no land hunger was unsuccessful to produce a revolt), could lead to the "*rejection of rapid commercialization as a cause of revolts*"; in fact, it could be the concurrence of "*land hunger and rapid commercialization*" that causes revolts. In the language of the statistical method, "*the cross-tabulation of the outcome with either of the causal variables would lead to independent rejection of them both*".[21] This model could not be observed due to the fact that researcher assumed land hunger alone to be the sufficient cause of a revolt. Consequently, such *conjunctural causation* critically weakens the type of paired comparison that is a part of the *Indirect Method of Difference*.

Moreover, as Ragin points out, the dependence of the *Indirect Method of Difference* on negative cases "*to reject competing arguments*" is also faulty. Neither land hunger, nor rapid commercialization, can be considered as redundant in examples, which show that absence of revolts and a presence of one of these two factors may both independently cause these revolts. The fact that both causes can be simultaneously "accepted or rejected", demonstrates the inconclusive nature of the *Indirect Method of Difference*, when dealing with cases of multiple causation.

The argument continues further with the ontological position of deterministic causations and whether these deterministic relationships are entirely possible. So, what are the epistemological issues related to causality with the *Indirect Method of Difference*? A concurrence between qualitative and quantitative scholars on "*complexity-induced probabilism*" might be observed.[22] Even if one admits that deterministic

---

[21] *Ibid*

[22] A. Bennett, *Causal Inference in Case Studies: From Mill's Methods to Causal Mechanisms,* American Political Science Association Conference in Atlanta, 1998.

social relations are existent, it is still uncertain how one would find out about their presence, if one were able only to observe a several causes interacting within a complex setting.

For instance, if it would be possible to precisely approximate the probability distribution of *A* causing the effect *x*, does one imply that these instances can also account as an explanation of the particular occurrence of *Y*? These complex interactions and several existing causes among deterministic relations could appear highly probabilistic if there would be no way to accurately test for the extent of effects these complicated causal mechanisms put forth individually. Thus, it might be more useful to focus on practical matters and of how one learns about the causes. In addition, the quest for finding answers to philosophical questions regarding the ontological status of some probabilistic ideas of causality might be temporarily suspended.

Evidently, the *Direct Method of Difference* is a more advanced method than the *Indirect* one, mainly because it involves a stronger manipulation. In Mill's view, one can manipulate the preceding factors in order to eliminate an assumed cause. The cause *A*, eliminated at will, might actually lead to *x* if the constant value of all other factors is maintained. If there was a preset and fixed relationship between preceding cause *A* and the effect *x* and if one is able to measure the value of either *A* or *x* (taking into account the possible occurrence of error), the resultant of the observed relationship would be nothing but probabilistic, mainly because it is only possible to observe a situation where one is mistakenly presuming that s/he is observing *A* (due to the stochastic measurement error) while, in fact, not observing the *x*. In this instance, the process of elimination would only strengthen the conclusion that *A* is not a cause of *x*.

Deduction, certainly, plays a prominent role in this approach; Mill places greater importance on deduction in the three-step process of "*induction, ratiocination, and verification*" while the issue of deductive reasoning has been largely neglected. When encountering multiple causes and interactions, one can think of two prevailing reactions, which are not

mutually exclusive. To quote Wesley Salmon[23], "*the primary moral I drew was that causal concepts cannot be fully explicated in terms of statistical relationships; in addition, I concluded, we need to appeal to causal processes and causal interactions*". Firstly, it is necessary to take into account formal theories, which describe the accurate empirical calculations. Those theories are usually tested within laboratory environments and controlled external factors, practical absence of measurement error, with strong manipulations of variables, which enables them to reach a precise result. Secondly, based on understanding of essential, or at least satisfactory, causes rely on conditional probabilities and counterfactuals (formal theories and statistical tests). Having stated that, one can conclude that if the theories are weak, the causes are complex, one cannot circumvent conditional probabilities.

Scholars are often uncertain about the differences between the *Indirect and Direct Methods of Difference*, many times claiming that they have been using Mill's *Method of Difference*, when they were actually employing the *Indirect Method of Difference*, i.e. in these cases, it would be complicated to define a set which would include all negative instances of social revolution. For example, Skocpol[24] while investigating the case of nineteenth-century Germany, as a negative instance of social revolt, and comparing it with other positive instances occurring in France, Russia, and in China, was using only, at best, the weaker *Method of Agreement* twice, while stating that she is using both *Methods of Agreement* and *Difference*. This is evidently not entirely correct, since it would be impossible to manipulate those factors of investigative interest. Causal complexities one perceives in social sciences are serious, and this limitation would play a very important role. In case-oriented inquiries, it is almost impossible to define inclusive sets, mainly because interest in specific cases or categories of cases often motivates the research itself. However, Germany experienced a massive upheaval in 1917-1918, which has almost grown into a social revolution. According to Ragin[25] and

---

[23] W. Salmon, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, 1984.
[24] T. Skocpol, *States and Social Revolutions,* Cambridge University Press, 1979.
[25] C. Ragin, *The Comparative Method*, Berkley UC Press, 1987.

Tilly[26], Skocpol's findings were not entirely correct, since Germany was on the verge of (at best) a negative instance. In their opinion, further states the selection of negative cases is random, mainly due to "the absence of strong theoretical or substantive guidelines". They believe that researchers who are often involved in extraordinary outcomes tend to rely on the *Method of Agreement* as a "universalizing" strategy.

## CONCLUSION

Mill's methods of inductive inference are suitable and valid only if the finding the relationship between preceding factors and effects is "*unique and deterministic*". Clearly, both investigated methods can be employed to cover any number of assumptions.

However, the social reality dictates that circumstances are impermissible neither for one cause to shape an effect nor to identify of interactions between different causes. Plainly, if one is concerned with determining the effect *x*, it is necessary to assume that the single possible cause for *x*. If that particular cause *A* is present, the effect of x inevitably has to happen. The two conditions, uniqueness and determinism, in reality define the set of preceding factors one is taking into account, implying that all the elements in the set of causes *A, B, C, D, E* need to at least have theoretical possibility of occurring independently from one another. The condition states that these preceding factors can take place without the presence or absence of any other preceding factors, or else, it would be impossible to differentiate between the possible effects of those factors. Hence, generalizations of Mill's methods suffer from these limitations. The causal mechanism of the Mill's *Indirect Method of Difference* is a quite dependable method, due to the fact that it establishes the idea obtained through the use of the *Method of Agreement*. While the Method of Agreement implies a causal relation within the positive set, the *Indirect Method of Difference* accomplishes the same with a negative set. Also, it can be employed both in cases of observations and experiments (as a modification of the *Method of Difference*), with a wide range of applications, since it encompasses negative instances that pass

---

[26] C. Tilly, *Big Structures, Large Processes, Huge Comparisons*, Russell Sage Foundation, 1984.

up the unrelated common preceding factors. However, there is a huge possibility of a plurality of causes that could not be avoided using this method, due to the fact that particular agreeable factors might be the causes of the same effect and yet be far apart. Also, the method does not successfully separate the mixture of different effects within a conjunction of causes, permanent causes (where the inability to obtain negative instances exists) and cases of different factor's co-existence, which makes it less practical and useful in several instances dealing with complex phenomena, mainly because it assumes that causes are essential for their effects.

The *Indirect Method of Difference* cannot be generally used in some types of examination of negative cases. Namely, testing negative cases takes for granted a preexisting theory, which might allow researchers to identify a set of observations which encompasses all possible instances of the phenomenon at stake. The other presupposed assumption is that among the set of factors taken under consideration, there should always be "*one factor that is the unique cause of the effect*". However, there is no guarantee that such assumption is always accurate, due to the fact that in some instances the cause might be some intricate combination of various factors.

Lastly, I have also concluded that Mill's *Indirect Method of Difference* is intended to categorize the presence of a probable cause within a set of all positive cases and the absence of it within a set of all negative cases. It is aimed at proving the causal relation in an indirect way, and under very rigorous conditions. Moreover, the method is designed to mark the presence of only necessary and sufficient conditions as a cause of a certain phenomenon. However, I doubt that the method can effectively work for more complex mechanisms, since their purpose is not to identify both necessary and sufficient causes, but only those that are sufficient. In this regard, and as mentioned before, the method fails to produce valid and reliable results when applied to complex phenomena, and when preceding factors are in concurrence of causes and of effects or when incidents of two investigated phenomena can be found or missing in several differing instances.

# PREPOZNAVANJE I ISPITIVANJE UZROČNIH MEHANIZAMA PRIMJENOM J.S. MILLOVOG INDIREKTNOG METODA RAZLIKE

## SAŽETAK

Ovaj rad se fokusira na istraživanje ideje uzročnosti u društvenim naukama, sa posebnim osvrtom na indirektni metod razlike J.S. Milla. Preciznije, predmet istraživanja je pronalaženje odgovora na pitanja da li primjena Millovog indirektnog metoda razlike daje pozitivanu rezultatntu koja omogućava izvođenje pouzdanih zaključaka o uzročnoj vezi između varijabli, te ako je odgovor potvrdan, da li je potrebno detaljnije precizirati o kojem je uzročnom mehanizmu riječ i testirati ga u različitim slučajevima? Ovaj rad ima za cilj ponuditi sveobuhvatnu interpretaciju Millovih osnovnih metoda i utvrditi da li je njegov indirektni metod razlike dovoljan za utvrđivanje prisustva mogućeg uzroka, u okviru skupa svih pozitivnih slučajeva, te za utvrđivanje odsustva uzroka, u skupu svih negativnih slučajeva.

**Ključne riječi:** John Stuart Mill, indirektni metode razlike, uzročnost.