World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:9, No:6, 2015

Internal and External Validity in Experimental Economics

H. Chytilová, R. Maialeh

***Abstract*—**Experimental economics is subject to criticism with regards to frequently discussed the trade-off between internal and external validity requirements, which seems to be critically flawed. This paper evaluates incompatibility of trade-off condition and condition of internal validity as a prerequisite for external validity. In addition, it outlines the imprecise concept of artificiality, which is found to be rather improving the external validity and seems to strengthen the illusory status of external versus internal validity tension. Internal validity is further analyzed with regards to Duhem-Quine problem, where unpredictability argument is significantly weakened trough application of inductivism within the illustrative hypothetical-deductive model. Our discussion partially weakens critical arguments related to the robustness of results in experimental economics, if the perfectly controlled experimental environment is secured.

***Keywords*—**Duhem-Quine Problem, external validity, inductivism, internal validity.

I. Introduction

EXPERIMENTALmethods played only a negligible role in empirical economics in the past, which has been voiced in some influential methodological writings. In the recent decade, the experimental method has quickly become a popular tool for economic research. However, as it is relatively new discipline, its discussion of important methodological issues lags behind the one in other disciplines. According to [22], serious debate considering major drawbacks and criticism of experimental methodology is addressed only by very few papers. Commonly shared belief is that controlled experimentation has little to offer and economics should be regarded as a non-laboratory science because it is almost impossible to conduct controlled economic experiments.

The situation has changed significantly during the last 20 years, with growing number of experiments conducted to address economic problems. It seems that experimental research has become a relevant branch of empirics, which is documented by frequent publications in leading international academic journals. However, most of the experimental methods are subject to criticism, with common arguments like experimental results may be spurious because it has been generated in the artificial environment. One might consider, whether claims of opponents [2], [10], [16], [19], are justified, or outdated nowadays.

Ing. Helena Chytilova, Ph.D., M.A.and Ing. Robin Maialeh are with the University of Economics, Faculty of Economics, Department of Economics, Prague, W.Churchill Sq.3, 10367, Prague 3, Czech Republic, (email: helena.chytilova@vse.cz,xmair900@vse.cz).

This note tries to address delicate methodological issues concerning the viability of experimental theory in contemporary economics with the aim to consider if some defense might be done against current criticism concerning the validity of experimental results. The main reason, behind the intention to focus on experimental methodology is, among others, that the field has not become ready yet regarding the rules of good scientific practice. There is still lively methodological debate, which is highly influential. Is there any potential in experimentation? There is a serious discussion about the character of experimental setting. Is the view of the opponents about the low external validity justified in the sense of the low potential of the artificial experiment? Should be experimental outputs considered as highly unpredictable on the basis of Duhem-Quine thesis and unstable human factors? Are there any remedies or options, which could weaken the contemporary critique? Internal and external validity in contemporary experimental economics, link the inside world of laboratory to the real world outside and stand in a relationship characterised by a trade- off as frequently mentioned in the literature. This issue mentioned above is a common argument for opponents, who built their critique mainly on the low potential of artificial laboratory experiments and therefore inability to produce valid results in this field and make relevant conclusions. The main goal will be to assess the critically methodological eligibility of trade-off between external and internal validity, based on the improper concept of artificiality. Taking views of experimental economists and opponents together, an evaluation is made about the current methodological state of experimental economics regarding these particular issues. If external and internal validity trade-off will be found illusory in their definitions and in their mutual trade-off based on artificiality, this could have serious implications for the methodological framework of experimental economics. Additionally, internal validity of experimental outputs will be further analysed with regards to Duhem-Quine problem, where question is posed whether perfectly controlled experimental design might not be rather of advantage within the inductivist approach, when dealing with unpredictability argument.

II. External and Internal Validity in Economics

Contemporary and frequently used the concept of internal and external validity, well known among experimentalists is a relatively young concept, which evolved only recently within the relatively young field of experimental economics. The basic principle lies in ability to distinguish between an inside

world of the laboratory experiment and the outside real world environment.

Internal validity is achieved, when cause-effect relation (interaction among factors) has been properly understood by the experimenter. Reference [12, p.142] claims that the experiment E and its results are internally valid in the inside world of the laboratory, if the production of an effect Y is attributed to a factor X, and "X really is a cause of Y in E. Furthermore, it is externally valid if X causes Y not only in E, but also in a set of other circumstances of interest, F, G, H.". In other words, these results should be transferable across laboratory walls. Moreover, internal validity is understood as a precondition for external validity and trade-off is posited between internal and external validity. This [12, p.144] represents a turning point, because trade-off is implemented trough artificiality, where "the more artificial the environment, the better for internal validity, the less artificial, the better for external purposes". This methodological stance in terms of artificiality might provide sufficient methodological framework for the field of experimental economics, compared to previous poor definitions. At the same time however this may also induce powerful arguments for opponents of the laboratory experiments.

A. Different Views on Internal and External Validity throughout the History

The initial interpretation of validity, not connected yet to economics that time, was firstly mentioned by [4]. A distinction was made between internal and external validity relevant for experiments in social settings. Validity understood from point of Campbell differed a bit from the later definitions of his followers. The experiment was considered to be internally valid only if it contributed with observations not yet obtained by previous research and if there was causal relationship between two variables and results were statistically significant. External validity was understood as an approximate validity, where this causal relationship can be generalized across settings. The important point is that these types of validities were not presented by Campbell as trends going in opposite direction. Bad experiment might be inferior in both validities; the good experiment would score high on both internal and external validity[13]. This suggests that initial interpretation indicates no clear sign of trade-off between internal and external validity as proposed by [12].

It was not until 1987, when it was firstly introduced in connection with Experimental Economics in the article of [3], "The External Validity of Experimental Economics Techniques: Analysis of Demand Behavior." However, this article in no way addresses internal validity issues yet. Moreover, implementation of validity into the methodological framework of experimental economics was not accepted by many experimental economists of that time and it took a long time till it disseminated fully into economics.

The definition of internal and external validity became stricter after its expansion to economics. Turning point was Guala´s contribution [12], when both validities started to be understood as opposing and excluding forces in the mutual trade-off, where artificiality plays a role.

What is often neglected by economists is the artificiality of setting, which is a major obstacle to external validity. Reference [17, p.33] advocates context as close as possible to the real world environment and says that economists "have not been able to avoid the problem of low external validity that is the Achilles heel of all laboratory experimentation."

Reference [22] also understands the problem of external validity as the problem of artificiality. It refers to the artificial world within the laboratory, which is rather incomparable to the natural world outside the laboratory. If the laboratory environment does not sufficiently mirror the real outside world, the loss of external validity may be significant. Therefore, it also asserts internal versus external validity trade-off. However, to what extent should laboratory mirror the outside environment depends mainly on the type and the goal of the experiment.

Reference [5, p. 220] states; "It is a well-known methodological truism that almost in all cases there will be a trade-off between internal and external validity. The usual complaint here is about the artificiality of circumstances required to secure internal validity."

Reference [25] asserts that there is a considerable limitation of the outside world issues, which can be replicated in the laboratory, but the discussion is limited to types of the experiments aimed to test economic theories. Therefore, supporting the view, that discussion of internal versus external validity is subject to the type of experiment, which is under scrutiny. Reference [1] argues, that laboratory environment in its artificiality cannot replicate the real environment and how humans behave. Reference [7] makes a statement that the differences between the inside world of the laboratory and outside world are too large. Many studies also take a view that internal validity is a prerequisite for external validity, for instance [14]-[18]. This condition is crucial, because as soon as causal relationships regarding the basic hypothesis are not sufficient, external validity cannot be further build on invalid grounds. Reference [11] in "Experimental methods for economists" also underpins the significance of external and internal validity and possibility of threatening external validity.

Pioneer experimental economists were reluctant to make any reference even to the validity phenomenon introduced by Guala and his followers. Many of them were unsatisfied with the new methodological framework and shared their reservations according to [13]. For instance pioneer in experimental economics, Smith downplays the importance of the two validities, emphasizing that it is merely an empirical question and equates validity with parallelism, which is sufficient precept to make results applicable to other environments, [24]. Moreover, in an email to the author Daniel Friedman, Vernon Smith made a remark that "he has never been especially enthusiastic about internal and external validity." [12, pp.12], (e-mail Vernon Smith to Daniel Friedman, 2 July 2009). John List and Glenn Garrison, other well-known experimental economists, made similar notes (e-

mail Glenn Harrison to Daniel Friedman, 2 July 2009; e-mail John List to Daniel Friedman, 3 July 2009), [12, pp.12]. Possible explanation of these attitudes might be that acceptation of this methodological framework, was considered as major threat to the field of experimental economics, because it created strict division between inside world of the laboratory and the world of outside, [13]. The second reason might be the insufficient specification of newly adopted methodology, mainly frequently discussed the trade-off between internal versus external validity connected with artificiality.

B. Trade-off between Internal and External Validity Justified?

General logic of experimental design, which is emphasized by many experimental economists among others [11], [12] might help to put internal and external validity in the context of methodological issues solved. The primary goal of an experimenter is to set causality that factor X is a cause of Y. Experimenter has to secure that all other confounding factors, which might enter in this causality are kept constant, (see Table I).

TABLE I
LOGIC OF AN EXPERIMENTAL DESIGN [12]

Treatment		Effect	Other Factors
Experimental Group	X	Y1	Constant
Control Group	0	Y2	Constant

The examined phenomenon is isolated trough creation of experimental and control group, where the former is exposed to the factor X, whereas the latter not. If the control by the experimenter is sufficiently secured, significant difference between outputs Y_1 and Y_2 should be attributed directly to the single factor X. This process of ensuring sufficient internal validity with help of direct control might be according to critiques too artificial and is responsible for lower external validity. The reasoning is that relationships, which have been set in the laboratory under controlled circumstances, cannot be applied to the outside world.

Question then arises, whether there is necessarily trade-off of internal versus external validity when experiment is carried out? As seen from the text above, no tension between internal and external validity is mentioned, [4]. Moreover, it is very unclear, how this trade-off is defined. In economic terms, when two variables stand in trade-off to each other, this means the more it is received of one, the less it can be received of the other. When applied directly to the situation of trade-off, either we can have experiment which has higher external and low internal validity and vice versa or there is possibility of rising internal validity at the expense of external validity and vice versa.

The second option is the topic discussed in current methodological literature as it mentions the tension between the two types of validities [15]. In a given experimental setting, the adjustment can be made in order for the experiment to have more internal validity at the expense of external validity, (1a). Or in a given experimental setting, the design can be altered in order to obtain more external validity at the expense of internal validity, (1b). Another claim which has been raised above is that internal validity has to be precondition of external validity, (2) After putting propositions (1a) and (1b) together with (2), the question arises about compatibility of trade-off between internal and external validity, with simultaneously imposed assumption of internal validity as a prerequisite for external validity. In order to satisfy condition (2), experimental design may be altered in direction of more internal validity at the expense of external validity, but not in the opposite way if we have to guarantee minimum internal validity. That is (1b) does not seem to be tenable. Taking it from opposite view, interpretation of condition (2) is that experiment does not have any external validity, if there is not enough internal validity in the beginning. However, once internal validity is supposed to be precondition for sufficient external validity, changes in the design that increase internal validity should leave external validity unaltered or in better case, external validity should be enhanced. This makes also (1a) impossible. In other words, external validity is ensured by sufficient level of internal validity, however at the same time imposing condition of internal validity at the expense of external validity leads to lower external validity, therefore making these two goals contradictory. Therefore, incompatibility of these conditions (1a), (1b) and (2) serves as an evidence of shortcomings and insufficiently defined concepts in experimental methodology.

III. ARTIFICIALITY DEFENSE

Artificiality is frequently mentioned argument by opponents, which is according to them accountable for the internal and external validity trade-off. Basically the more artificial feature the experiment possesses, the higher is internal validity of experiment at the expense of external. Artificiality might seem to be understood as substitute for internal validity resulting from artificial character of laboratory environment and strengthening thereby the argument of low external validity of each experiment. The following section will try to refute commonly held view about significant role of artificiality in validity trade-off.

A. Improper Definition of Artificiality

The problem lies in the conceptualization of artificiality, which is rather unclear. Many studies in experimental economics mention artificiality only in a very general way, mostly negative, emphasizing that it is accountable for too high internal validity at the expense of external validity. However, this is very imprecise definition, when artificiality is considered as an attribute of the degree of intervention in experiments [11]-[15].

Artificiality is relative concept, according to the typology of experiments ranging from less artificial to more artificial. In addition, artificiality not properly defined is subjective view of the experimenter and experimental subjects, immeasurable and therefore undetectable in degree of artificiality. Moreover, artificiality is not a strictly unified concept, but has several characterizations [15]. The first characterization is Hawthorne

effect, referring to the situation of subjects knowingly being under scrutiny and therefore behaving not in natural way. This effect of artificiality can be easily corrected or weakened if experimental design is sufficient, where differences between experimental and control group are secured trough sufficient identification of possible factors involved in examined effect. Moreover, this effect is not restricted to the experimental method and is rather common for other methods as well. Second characterization of artificiality is failure to capture some theoretical entities, as most formal models leave out details. Owing to practical difficulties, construct validity is thus threatened. This type of validity can jeopardize external validity, but is not accountable for the tension between internal and external validity and thus trade-off between them, [6]. The last issue related to artificiality is a criticism towards particular experimental procedures, for example anonymity of players, which is not the case in the real circumstances, therefore threatening external validity with respect to connection to the real world, [15]. However, here reasoning might be provided by consideration of what is meant by the real world in the laboratory, where classification is being made dependent on type of the experiment, which is subject to analysis in the next section.

From above, artificiality might account for some threats to external validity, but not for the tension between internal and external validity due to its disintegrated and immeasurable character. In addition, owing to the subjective character of artificiality depending on the type of experiment, criticism about a general artificiality of the laboratory environment is meaningless.

B. What is the Real World? Is Artificiality Real Obstacle?

According to above mentioned critiques, experiments are rather insufficient representation of the real world. Artificial character of laboratory experiments creates less real situations than studies in natural settings.

However, the laboratory world is part of our real world. This seems to be also supported by economists, who consider laboratory markets as real markets, where general principles of economics hold like in any other market [20]. Reference [20, pp. 1486] claims that "While laboratory process is simple in comparison to naturally occurring processes, they are real processes in the sense that real people participate for real and substantial profits and follow real rules in doing so. It is precisely because they are real that they are interesting." Therefore using the word artificial in this sense was considered by Plott as rather inappropriate. Moreover, laboratory environment offers possibility to test general theories and models, where they are expected to hold, as it is assumed they will work in the special conditions. The aim of most experiments is not to exactly mirror patterns of behaviour, but identify causal relationships, which is not possible to isolate outside the laboratory environment.

Indisputably, the advantage of artificiality lies in the fact that it eliminates irrelevant variables, therefore making generalization more probable. Moreover, the advantage of perfectly designed laboratory experiments is that participants respond only to theoretically relevant factors. It is necessary to point out that only choice of factors was made artificially, but responses of participants are real. References [9]-[18] point out that although experiments are unrealistic in their abstraction from the reality, their simplicity like in a model case, is often a virtue, because it increases understanding of the interaction of relevant variables.

According to the analysis above, artificiality seems to be an advantage, because it enables to examine only relevant factors, which are connected to the theory, with elimination of unnecessary elements. Artificiality therefore does not reduce external validity needed, because generalization is more probable in simplified environment, which isolates irrelevant variables.

C. Type of Experiment and Artificiality

What is a proportion in which internal validity should be present in experiment compared to external validity? If we add postulate mentioned in [25] that it depends on the type of experiment, it suggests that setting of internal versus external validity, when they are supposed to be in mutual trade-off, is very unclear. In other words, this postulate supports Smith´s view that internal and external validity issue is rather empirical thing and it is up to the critics to falsify parallelism of any specific experimental output.

Reference [16] also highlights that external validity is more important for experiments aimed to search for empirical regularities compared to theory testing experiments. Reference [24] indirectly states that more attention regarding parallelism should be paid to experiments that do not aim at testing theories. This view is also supported by [22], where the external validity required depends on the goal of the experiment. Compared to previous studies it provides thorough analysis of experiments, according to the intensity of external validity needed. Theory testing experiments, in which category most of the experiments fall according to this study (after rough categorization of 69 papers, where 33 papers fall in category testing theories), do not require external validity at such level, like other types of experiments. In this case internal validity is preferable to external, mainly because of ambitions not going beyond the walls of laboratory in terms of generalization. Reference [9] also argues that for this sort of experiments, which aim to test a theory or find a failure, evidence is important exactly for theoretical framework, but not for a closer understanding of the real world. Theory stress tests and experiments searching for empirical regularities are more important in terms of external validity. And finally, category of experiments aimed to advise policy makers are highly demanding in terms of external validity. This suggests that validity of laboratory outputs is matter of separate evaluation of each experiment.

IV. Debate about Proper Experimental Design

It is widely discussed by opponents and defenders of experimental economics, what is the appropriate experimental design. When experiment is being set up, one might propose, that it should resemble closely the complexities of a real-world

environment, in order to avoid artificiality critique, frequently mentioned by sceptics. For instance, [17] advocates experiments designed in the environment, which is similar to the behaviour of agents in the real world. Reference [24, p. 33] claims that inability of researchers to avoid low external validity is considered by him as "Achilles heel of all laboratory experimentation." However, based on above-mentioned discussion, external validity problem depends on the type of the experiment that is held, rather than on trade-off implied erroneously by improperly defined artificiality.

Based on this, it is important to note, that experimentalists do not rule out, that attention should be paid to external validity issues, like [23]. However, this concerns experiments, whose focus is not on theory testing. Most experimentalists support the classic view of non-significance of external validity, [22]. This further supports the view of the dependence of external versus internal validity setting on a particular type of the experiment. We may wonder if there is any justification for this approach. Besides practical issues, concerning the budget or impossibility to capture some details, there is a major threat of so complex environment, where it would be practically unfeasible to distinguish causes and effects. Control of the experimenter over the factors, which enter the experiment, would be almost disabled and therefore, internal validity significantly weakened. Moreover, economic theories do not typically refer to such complexity and exhibit simplicity in order to explore the phenomena without disturbing influences. Reference [20, p.906] states that it was believed by many researchers, that "the only effective way how to design an experiment would be to mirror in every detail, to simulate, so to speak, some ongoing natural process." However, similarly to economic theories, the experimental design should be able to analyse relationships while abstracting from the other variables, which would complicate the investigation and is not part of the phenomena. The true merit of the laboratory experimentation is the possibility to close the door on other factor and isolate the phenomenon via the introduction of experimental and control group.

However, how closely should be followed the formal model by experimental design? The major difficulty is that models omit some important details. As an example could serve a rational expectations model, where traders orders are based on observed market-equilibrating prices, [11]. The practical issue is whether market-clearing prices should be announced before, or after traders place orders. Thus, the experimenter is forced to make choices, which are arbitrary in terms of the theory, but significant in terms of behaviour. Reference [11] notes that even if the experiment would be in close connection with the formal model and perfectly mirror all of its assumptions, it does not have any value added, only it can confirm, that model is not flawed. It does not capture model's explanatory power. More profitable it seems to design experiment, where assumptions of the model are a bit relaxed and mechanism of the model not precisely followed step by step. The model generalizability and applicability can be detected, as theory embedded in this model does not hold only in this specific case but is applicable in a wide range of cases.

A. Internal Validity and Dealing with Unpredictability Argument

If we adopt an approach to setting the laboratory experiment in a way proposed above with regards to the external validity, still doubts remain about the validity of experimental results. An open question remains if really experiment provides an appropriate test of a given theoretical hypothesis by means of internal validity. The derivation of an empirical prediction, as widely accepted in the philosophy of science, usually involves a set of assumptions and not a single hypothesis as noted by [12]. In other words, one is testing the hypothesis, which is supplemented by a variety of auxiliary hypotheses. This presents a challenge because it is impossible to detect conclusion from the empirical test with certainty. We may fail to accept the right hypothesis because we are not sure about the completeness of the auxiliary hypotheses. This may lead to the situation, where we exactly adopt the hypothesis which is false. This is commonly presented as the Duhem-Quine problem, [7]. With help of hypotheticodeductive model of testing (HD model), it might be shown, that Duhem-Quine problem is not a problem, but rather challenge in line with inductivist approach, where experimental economics might provide support in order to improve predictability of outputs.

In the left column of Table II, H stands for scientific hypothesis and E for empirical evidence.

TABLE II
HYPOTHETICAL-DEDUCTIVE MODEL OF TESTING [12]

Basic HD Model	HD Model With Auxiliary And Background Assumptions
Scheme A (Refutation)	Scheme C (Refutation)
H —› E	(H & I & K) —› E
~ E	~ E
~ H	~ H
Scheme B (Confirmation)	Scheme D (Confirmation)
H —› E	(H & I & K) —› E
~E	~ E
Probably (Or More Probably) H	Probably E

Notations used: the arrow—› stands for the relation of implication („if.... then...."), ~ stands for the negation.

Scheme A and B are rather different. In scheme A, *modus tollens* argument is present, where conclusion logically follows from assumptions, where hypothesis H is set that something holds and prediction e is formed, what we shall observe, therefore setting valid deductive inference. On the contrary, in scheme B, deductive approach alone cannot be used to confirm hypothesis directly on the basis of assumptions mentioned above. It should be understood in terms, that the observation of empirical evidence makes the hypothesis H highly likely to be true. It can be said that empirical evidence confirms, supports or indicates H, but not that H is true, which would represent a delusion.

Modus tollens with refutation of the hypothesis as illustrated in scheme A represents pure hypothetical deduction according to Popper, where there is no need for confirmation

or induction. This was subject to criticism from various positions, because the aim of scientists is not only to propose hypotheses and theories to reject them, as it does not indicate that the theory is valid, but hypotheses should be used to predict future events and identify some phenomenon, [12]. Philosophers of science indicate that moving in scheme B, which requires some logic of induction, apart from the deductive logic, is inevitable. The predictive success requirement followed by deductive approach [21] is not sufficient in this sense. Neither, it is satisfied by instrumentalism of Friedman [10], where scientific theories are just tools for anticipating future events, without any attempt to explain the mechanism and their components. For e to confirm H, it is necessary and sufficient that e is logically implied by H, according to deductive approach. The relation between e and H requires a more complicated framework. However, model B is not sufficient in this sense. We need to build in important elements, which should strengthen the causation. Definition of initial conditions (noted by symbol I), is necessary together with a hypothesis for the derivation of evidence. Initial conditions usually include basically known facts, (for instance law of demand), from which then might be generated other empirical statements about examined phenomena.

Auxiliary and background assumptions (noted by symbol K), are needed in order to derive a prediction. These are crucial in order to specify the close behaviour of agents with regards to examined phenomenon. More specifically this means the definition of the particular function according to which agents form decisions and are supposed to maximize the value of available possibilities. Additionally, also the thorough specification of parameters such as interval in which price is selected or occurrence of some shock is included in these background assumptions. These K assumptions are crucial, mainly for experiments in economics itself, because they help to ensure correct function of the instruments, zero occurrences of disturbing events, specify correct parameters and error term. Initial and auxiliary and background assumptions are included in models C and D in Table II. These do not represent a pure deductive model. In model C, for hypothesis H to be false, at least one element from H (hypothesis), I (initial conditions), and K (background and auxiliary assumptions), must be false. Thus, it eliminates pure deductive view, that the evidence of negation of e implies the rejection of the hypothesis. Refutation of the hypothesis has to consider the whole set of assumptions mentioned above [12]. Similar arguments would apply to Scheme B for confirmation of the hypothesis.

Duhem-Quine thesis was considered to be a problem in the process of hypothesis verification, because of the potential absence of some important auxiliary and background assumptions. However, since inductivism finds support within the scientific community by not employing pure deduction, rather the inclusion of additional background assumptions is preferred. This helps to improve predictability of phenomena and confirm rather than reject the hypothesis. As a result, it is preferable to find ways how to cope with Duhem-Quine problem.

However, inductive logic is a tricky issue. How can we ensure completeness of auxiliary hypotheses, often mentioned by critics? Reference [8, p.16] notes that here exactly experimental economics might be helpful since perfectly controlled laboratory experiment allows for "significant reduction in the number of auxiliary hypotheses involved in the verifying primary hypothesis." The Duhem-Quine issue can be solved better in this environment, because we can adopt mechanism, which enables us to control theory and manipulate or omit exactly variables and auxiliary assumptions in the way that enables to isolate phenomenon we wanted and strengthened causation. This also provides argument against critiques of internal validity of the laboratory experiments, since advantage of the controlled environment is as stated by [12] that collection of the evidence occurs in ideal circumstances, where we are sure that the background conditions are right.

B. Methods of Induction Dealing with Duhem-Quine Problem and Predictability

In order to deal with Duhem-Quine problem and improve predictability, experimentalists set up the perfectly controlled experiment as mentioned above. Direct control enables to hold some variable constant and enforce this in form of the rule during the experiment, or alternatively work with treatment variable, where this variable is controlled at two or more different levels with different outcomes for each part of the experiment, [11]. The group in certain experimental conditions behaves differently from the other similar group, which has different experimental conditions imposed by the change in one or more parameters (treatment). Similarity of groups with respect to some key characteristics is achieved by matching procedure. It happens, that experimenter is not able to control for every possible flaw and some of them are even not known to him, given the present conditions. The list of possible errors is potentially infinite, but at least we can indirectly affect the other factors, which are yet unobserved by the method of randomization. This ensures the independence of uncontrolled variables over treatment variables, by assigning chosen levels of the treatment variables in random order, which spreads unknown factors evenly across the treatments, [11]. It is not the intention to describe here the logic of perfect experimental design thoroughly, but to depict its significance for inductive methods, which strengthen predictability of phenomenon and provide additional counter arguments for critiques of experimental methods.

The last point concerning predictability of examined phenomena concerns view of the sceptics, who cast doubts on the reliability of experimental results based on the behaviour of individuals. Human action is a product of self-conscious choice and thus prone to variability, [25]. This is closely connected not only to internal but also to external validity requirement and ability to extrapolate experimental findings beyond the lab. As human behaviour is sensitive to a variety of factors, results might significantly differ in the laboratory and real-world setting. Reference [25] notes that it is hard to

make an objection against experimentation on the basis of unpredictable behaviour, without undermining the predictive role of economic theory. Prediction rests on the right or wrong presumption that there is at least some regularity in human behaviour, in which economists are interested. Induced value theory offers solution how to deal with agent´s innate characteristics, [11]. In this sense following conditions are sufficient: Monotonicity ensures, that subjects prefer more reward medium to less and non-satiation. Salience secures that the reward medium is related to subject's actions according to the rule the subject understands. The last rule, dominance should secure that the rewards from the experiment should be sufficiently high to neutralize other influences.

Apart from these rules which help to support auxiliary assumptions, experimental setting has one advantage in order to weaken the variability of human behaviour. Suppose we want to confirm some theory prediction experimentally, and behaviour of subjects is not in line with the prediction. We should not reject the theory, but apparently the experiment did not incorporate all of the conditions, which are necessary for confirmation of theoretical statements. Therefore, we can vary experimental setting and run the new experiment when some new important aspect emerges, and it seems that it could correct the limitation of the previous experiment. This could correct the distorted behaviour of subjects, induced by inappropriately designed experiment. So it is also suitable for the situation, when the experimental design was appropriately set up with applying direct, indirect control and still some disturbance is present due to the fact that some auxiliary assumptions are omitted.

V. Conclusion

This paper tried to tackle some important methodological issues concerning the criticism of current experimental economics. In the recent decade, the experimental method has become a popular tool for economic research quickly. However, as it is a new discipline, its discussion of important methodological issues lags behind the one in other disciplines. Serious debate considering major drawbacks and criticism of experimental methodology is addressed only by very few papers. This article addresses the problem of internal and external validity, where the inside world of the laboratory and real world environment stand in a relationship characterised by a trade-off. The strict division between inside and outside environment is a common argument for opponents, who built their critique mainly on the artificiality of laboratory experiments and therefore inability to produce valid results in this field. However, this frequently cited tension between internal and external validity seems to be critically flawed. Incompatibility of trade-off condition and condition of internal validity as a prerequisite for external validity is presented, which indicates rather illusory character of internal versus external validity tension. Moreover, further evaluation of artificiality concept, often mentioned to be accountable for excessive internal validity at the expense of external validity, suggests its meaningless role in internal versus external validity trade-off. Also, the experimental environment rather advantage, because it enables to examine only relevant factors in the identification of causal relationships, therefore making generalization more likely. Finally, the vagueness of trade-off concept is strengthened by the unclear specification of the level of internal versus external validity required, dependent mainly on the type of experiment. This refutes the commonly held view that experiments are artificial in general. According to this it rather appears that still the border between external and internal validity is very weak, depending on the type of the experiment.

In the process of discussion of sufficient experimental design, which reflects the issue of the validity, the question was posed if experimental economics really provides an appropriate test of a given theoretical hypothesis by means of internal validity. The second half of the study was devoted to the evaluation of Duhem-Quine problem, which might be considered as a threat to the predictability of examined phenomenon in terms of secured causality. With the help of the hypothetical-deductive model, inductivist approach was introduced, where the whole set of background assumptions was considered as opposed to pure deductive view. It was shown that Duhem-Quine problem within inductivist approach is rather challenging since the completeness of auxiliary and background assumptions is not secured. However, Duhem-Quine problem associated with unpredictability argument seems to be solved partly if it goes hand in hand with experimental economics, which is exactly aimed at the elimination of Duhem-Quine problem and, therefore, strengthens inductivist approach. Thus, it seems that experimental economics may be sufficient fundamental in increasing predictability of examined phenomenon and therefore strengthening its internal validity rather than decreasing. This is a secured trough perfectly controlled experiment, which deals with this critique with the help of direct and indirect methods. In addition, the frequently mentioned problem connected to the reliability of experimental output with respect to human variability constraint might be partly solved through induced value theory. Moreover, if the examined phenomenon is not certain regarding its results with respect to correct specification of experimental design, the major advantage of experimental economics is a possibility to vary environment and replicate the experiment. This might help to eliminate potential deficiencies and deal with disturbance due to the potential omission of some auxiliary assumptions.

References

[1] K.Binmore, Playing Fair. MIT Press,1994.
[2] K. Binmore, A. Shaked, "Experimental Economics: Where Next?"*Journal of Economic Behavior and Organization*, 2010, 73 (1), 87-100.
[3] D.S.Brookshire, L. Coursey,D.W. Schulze, "The External Validity of Experimental Economics Techniques: Analysis of Demand Behaviour,"*Economic Inquiry*, 1987, 25 (2), 239-250.
[4] D.T. Campbell,"Factors Relevant to the Validity of Experiments in Social Settings, "*Psychological Bulletin*, 1957, 54(4), 297-312.
[5] N. Cartwright, "Hunting Causes and Using Them," Cambridge:University of Cambridge Press, 2007.

[6] T.D. Cook, D.T. Campbell, "Quasi-experimentation: Design and Analysis Issues for Field Settings," Boston, Houghton Mifflin Co., 1979.
[7] J. Cross, Some Comments on the Papers by KagelabdBattalio and by Smith. In Kmenta, J. And Ramsey, J., eds., "Evaluation of Econometric Models," New York University Press, 1980.
[8] D. Davis, C. Holt,"Experimental Economics," Princeton University Press, ISBN: 9780691043173, 1993.
[9] E. Fehr, A. Falk, "Why Labour Market Experiments?"*Labour Economics*, 2003,10 (4), 399-406.
[10] M. Friedman,"The Methodology of Positive Economics." in Essays in Positive Economics, edited by Milton Friedman. Chicago: University of Chicago Press, 1953.
[11] D. Friedman, D. S. Sunder,"Experimental Methods. A Primer for Economists,"Cambridge: Cambridge University Press, ISBN 0-521-45682-7, 1994.
[12] F. Guala,"The Methodology of Experimental Economics," Cambridge: Cambridge University Press, ISBN 0-521-85340-0, 2005.
[13] F. Heukelom,"Origin and Interpretation of Internal and External Validity in Economics," Nijmegen Centre for Economics, NiCE Working Paper 09-111, 2009.
[14] R.B. Hogarth,"The Challenge of Representativeness Design in Psychology and Economics,"*Journal of Economic Methodology*, 2005, 12 (2),253-263.
[15] J. Kessler, L. Vesterlund,"External Validity of Laboratory Experiments," In preparation for The Methods of Modern Experimental Economics, editors Guillaume Frechette and Andrew Schotter, Oxford University Press, 2010.
[16] R. Lipsey,"An Introduction to Positive Economics," 5th ed., London, Weidenfeld and Nicolson, ISBN 0297775979, 1979.
[17] J.W. Lucas,"Theory-Testing, Generalization, and the Problem of External Validity," Sociological Theory, 2003, 21(3), 236-253.
[18] C. Plott, "Will Economics Become an Experimental Science," Southern Economic Journal, 1991, 57 (4), 901-19.
[19] C.R.Plott, "Industrial Organization Theory and Experimental Economics," Journal of Economic Literature, 1982, 20 (4), 1485-1527.
[20] K.R. Popper, "Conjectures and Refutations," London: Routledge, 1963.
[21] P.A.Samuelson, W.D. Nordhaus,"Principles of Economics," 12th ed. 1985.
[22] V.L.Smith,"Experimental Economics: Reply."*American Economic Review*, 1985, 75 (1), 265-272.
[23] V.L. Smith,"Microeconomic Systems as an Experimental Science,"*American Economic Review*, 1982, 72 (5), pp. 923-955.
[24] Ch.Starmer, "Experiments in Economics: Should we trust the dismal scientists?" *Journal of Economic Methodology*, 1999, 6 (1), 1-30.
[25] W. Wallis, M. Friedman,"The Empirical Derivation of Indifference Functions," In: Studie in Mathematical Economics and Econometrics in Memory of Henry Schults, O. Lange, F. McIntyre, T. Yntema (eds), Chicago Press, 1942.