EXPERIMENTAL MACROECONOMICS, A CHALLENGE FOR THE FUTURE?

Helena Chytílova, Ing. Ph.D., MA
University of Economics, Prague, Faculty of Economics, Czech Republic

Abstract
Standard econometric approaches have been many years a well-established workhorse in economic research. However, widespread application of macroeconomic models with explicit micro-foundations since rational expectations revolution in 1970 gave birth to experimental macroeconomics recently, where controlled environment might be used to get insights regarding the effects of alternative institutions and policies with direct implications at the aggregate level. This article aims to enrich present discussion about appropriateness of this method, since a very few studies deal with this new phenomenon among others Ricciuti (2008), Duffy (1998). We claim that experimental macroeconomics shouldn’t be omitted as one of the possible methods in economics, since individual and aggregate outcomes might be assessed. As a result, it should represent complement to standard econometric techniques. Furthermore, illustrative case study is used to demonstrate that macroeconomic experiments are not a distinct ones from microeconomic ones and represent powerful instrument for evaluation of micro-founded macroeconomic models.

Keywords: Macroeconomic experiment, controlled environment, aggregate outcomes, external validity

Introduction
Standard econometric approaches have been many years a well-established workhorse in economic research. Recently, experimental economics has gained considerable standing as an alternative method. However, a very special subfield called experimental macroeconomics is still very unheard of, despite being utilized by many well-known economists as mentioned in studies of Ochs (1995), Duffy (1998) and Ricciuti (2008).

Experimental macroeconomics as a relatively new discipline aims to address aggregate economic phenomena with help of controlled laboratory methods in which specific assumptions and predictions of macroeconomic models are tested. Without rational expectation revolution (Lucas 1972), development of experimental macroeconomics would have been impossible. Since then, macroeconomic propositions started to be based on individual maximizing behavior, given the constraint. These micro-foundations of macroeconomic models opened the space for experimental macroeconomics, which observes interaction of individuals in artificial environment followed by individual and aggregate outcomes. Based on that, powerfulness of micro-foundations of macroeconomic models might be assessed.

Although there are many possible insights to be gained, still experimental results are met with skepticism with regards to macroeconomic implications from rather insufficient experimental samples. As argued by Sims (1996, p. 107) “Economists can do very little experimentation to produce crucial data. This is particularly true of macroeconomics.”

This article aims to contribute to relatively new research, outline relatively new subfield of experimental macroeconomics and disprove common misleads about experimental macroeconomics. After delineation of position of experimental macroeconomics within the laboratory economics, identification of the main burning downsides of macroeconomic
experiments mentioned by critiques is the case. Based on that, arguments in favor of this new method are discussed in order to alleviate frequently mentioned deficiencies of experimental design. Furthermore, illustrative case study is used to demonstrate that macroeconomic experiments are not distinct from microeconomic ones and represent powerful instrument for evaluation of micro-founded macroeconomic models.

**The Nature of Experimental Macroeconomics**

In order to understand the true nature of experimental macroeconomics, its position within the laboratory economics has to be delineated. According to Ricciuti (2008) two classes of experiments are present in macroeconomics. The first approach considers an experiment which concentrates on single market. Not only it is easy to apply ceteris paribus condition in case of a single market, but this approach is more consistent with the current character of macroeconomic modeling, based on micro-foundations. As a result, most experiments fall rather within this category. The main purpose of these types of experiments is to test microeconomic predictions and assumptions of macroeconomic models with direct implications at the aggregate level. This type of experiment may be summarized in vein of Duffy (2011, p.6) who claims: “In practice, experimental macroeconomics is not distinct from microeconomic laboratory experiments, there is just a different focus or interpretation. A macroeconomic experiment is one that tests the predictions of macroeconomic model or its assumptions or is framed in the language of macroeconomics.“ Even Ricciuti (2008) emphasizes that it is almost impossible to find a substantial difference of macroeconomic experiment over microeconomic one, since these type of experiments are based on individual behavior of agents and the only macroeconomic feature is analysis of aggregates such as inflation, unemployment, etc. The second approach was firstly applied by Charles Plott in several of his experiments and is the one, which has a real macroeconomic content due to its focus on inter-relations between several markets and the spill-over between them. This Walrasian type of the laboratory experimentation is related to examination of the system effects, equilibration and spillovers between markets. To sum it up, although laboratory economy is not as complicated as the real economy, its simplified version may provide clue whether the model can or cannot be applied to more complex real world. If a simplified version of the economy in the laboratory does not reject the model of macroeconomic behavior, it means that the model may be actually plausible.

Precise origins of macroeconomic experiments are rather not clear according to Duffy (2008). Explicit statement of the existence of macroeconomic experiment within economics are mostly dated to Robert Lucas’s (1986) invitation to macroeconomists to conduct controlled laboratory experiments to cope with coordination problems, which remained unsolved by macroeconomic theory. However, Vernon Smith’s (1962) double auction experiment is prevalently considered as the first macroeconomic experiment conducted ever. Sometimes it is pointed to Phillips’ (1950) experiment, which used a colored liquid-filled tubular flow model of the macroeconomy, which however did not involve human subjects. Additionally, John Carlson’s (1967) experiment, which aimed to examine price expectations in stable and unstable versions of the Cobweb model is also one of the first possible attempts.

**Possible Insights and Common Downsides of Macroeconomic Experiments**

The basic difference between experimental approach to testing theories and standard macroeconomic modelling might be illustrated with help of Figure 1 below. According to Noussair (2012) both methods have the same chain starting from the structure of the economy, followed by agents’ behavior with outcomes at the end. The difference lies in the way how to reach the final outcomes. Whereas theoretical models simply assume the structure of the economy, macroeconomic experiments set up the structure of the economy in the laboratory environment. Secondly, in case of theory the behavior of agents follows from assumption
given by the structure of the economy. In case of macroeconomic experiment it ensues from the interaction of individuals and their decision-making in artificial environment.

Lastly, in case of theoretical models simulated outcomes and equilibria are outputs which might be compared with the data acquired on the basis of the laboratory experiment.

Figure 1: Experimental Approach to Testing Macroeconomic Theories

The basic advantage of experimental economics as opposed to real data is that it offers full control over parameters in the lab, since the experimenter is flexible in modelling and can set parameters in a desirable manner and direction. As a result, precise information about the factors, which changed during the experiment are available, (Ricciuti, 2008). This is also connected to micro-level causal relationships, since experimenter can directly isolate phenomenon which he aims to test via parameters setting. Moreover, unaccounted factors are highly eliminated. Causal relationship in terms of so called internal validity is secured by specification of experimental group, which is exposed to tested effect and control group, which is not exposed to this effect, but otherwise groups face the same conditions ceteris paribus. Critiques point out to Duhem-Quine problem related to under-determination (and impossibility to secure all ceteris paribus condition), due to which it is highly tough to secure internal validity and thereby examine given phenomenon. However, this is general problem in experimental economics, (See for instance Guala 2005). Furthermore, if we make comparison of experimental data with field data, they cannot be described by the same degree of internal validity in terms of causal relationship, (Duffy 2008). Additional argument, which goes in favor of laboratory experimentation is replication of the experiment (with adjustment) if doubts about internal validity emerge, (Smith, 1982). Concerns about internal validity are common for both microeconomic and macroeconomic experiments together. However, what are the most striking methodological concerns related directly to experimental macroeconomics?

The most frequently mentioned issue, which is considered to be a delicate problem, is external validity. External validity refers to “inferring from the special circumstances created in the laboratory to the phenomena that takes places in the real world”, (Guala 2005, p.141). Firstly, external validity is also common for microeconomic experiments. As a result we should delineate external validity in case of macroeconomic experiments, which is rather related to the problem of sample size. Secondly, artificiality seems to play also some role, but that is also the case for microeconomic experiments. As opponents point out, substantial internal validity at the expense of external validity may lead to high artificiality of experiment, whose results is impossible to apply in the real world, (Starmer 1999), Cartwright 2007). For instance Lowenstein (1999) points out that economists “have not been able to avoid the problem of low external validity that is the Achilles heel of all laboratory experimentation,” (Lowenstein 1999, p.3).

To clarify the former, macroeconomic experiments are described by rather small sample size. Involving small groups of experimental subjects interacting for a short period of time represents serious problem according to opponents, in order to utilize experimental outputs and derive implications for the real world. The analysis of aggregate economic
phenomena or even testing of predictions or assumptions of models is met with some 
skepticism due to non-representativeness of the laboratory experiment from macroeconomic 
point of view. However, with the current trend of macroeconomic models with 
microeconomic foundations, the issue of number of subjects may not represent a problem. 
Classes of macroeconomic experiments centered on single market are considered to be rather 
simplification of the real economy, which is aimed to test exactly microeconomic foundations 
of macroeconomic models. In other words, definition of macroeconomic experiment itself 
provides justification for acceptable external validity, consistent with the current trend of 
macroeconomic modeling. Moreover, their scientific strength is in building on economic 
theory, (Ricciuti, 2008, Duffy, 2008). Additionally, studies of Smith’s double auction (1962), 
Forsythe (1982), Plott and Sunder (1982), Sunder (1995) with partial equilibrium approaches 
and Lian and Plott (1998) with a general equilibrium approach, generally conclude that small 
population of 5-10 subjects with enough trading experience is sufficient in order to achieve 
efficiency consistent with competitive equilibrium in various market environments. Statement 
about not necessarily large sample size is documented by various experimental studies of 
macroeconomic nature like Fehr a Tyran (2008), with number of subjects n=76, Fehr, 
Kirchsteiger, Riedl (1998), n=52, Adam (2007) n=30, Duffy a Fisher (2005), n=10, Arifovic, 
Sargent, (2003), n=12, Van Huyck, et.al (1994), n=40 and others. Additionally, we should not 
rest on too restrictive definition of macroeconomic experiment, since there are many 
experiments like coordination issues, which have purely microeconomic content with 
macroeconomic flavor as noted by Duffy (2008).

To clarify the latter, internal validity at the expense of external should not be 
considered as serious problem, since the degree of external validity is dependent on type of 
the experiment. Schramm (2005) emphasizes that highly theoretical experiment doesn’t 
necessarily require high degree of external validity, but rather of internal validity in order to 
secure strong causality. In this case, experiments are used to test performance of theories in 
light of working institutions or to test the initial assumptions of the theory. The other side is 
represented by experiments designed to test-bed policies. In these experiments external 
validity is of a major importance. This is usually common, when some new institutional 
design is tested and requires unique practical skills, which are out of student’s domain.

**Illustrative Macroeconomic Experiment**

Robert Lucas’s (1986) invitation to conduct macroeconomic experiments was 
followed up on by many influential economists, recently Lim, Prescott and Sunder (1994), 
thories tested in the laboratory, among others by Duffy and Fisher, (2005), Van Huyck et al. 
Duffy, Ochs, (1999), (2012), Fehr and Tyran (2007), Fehr and Tyran (2008), Heinemann, 
Nagel, Ockenfels (2004), Hey (1994), Noussair, Plott, Rizemann (2007), Lei, Noussair, 
(2007) and others. There is infinite variety of macroeconomic experiments, ranging from 
Walrasian competitive equilibrium, asset price bubbles and crashes, bank runs, sunspots, 
Keynesian coordination failures to speculative currency attacks and contagions.

This study will restrict attention to microeconomic experiment with macroeconomic 
flavor in vein of Fehr and Tyran (2001) in order to further support afore-mentioned 
methodological discussion and provide possible justification for conducting macroeconomic 
experiments. Experiment is based on n-player pricing game with unique equilibrium and 
strategic complementarity in vein of Haltiwanger and Waldmann (1989), where the best 
reaction of individual in order to maximize profits is to set price, which is positively related to 
the average price set by other n-1 players. The game has 40 rounds plus one trial period, with 
a group size of n=4. Experiment is divided into a pre-shock and a post-shock phase, all of 
which has T=20 periods. Fully anticipated negative monetary shock is implemented during
the game, which is common knowledge to participants, (reduction of money supply from \(M_0=42\) to \(M_1=14\). Treatment groups received payoff functions, which provide them with information about their pricing strategy. Pay-offs of participants are expressed either in nominal or real terms. In order for subject to decide correctly about the price of his product (price lies between 1 and 30) in the nominal environment, he needs to re-count nominal pay-off into the real pay-off. The nominal pay-off is given by \(P_i \cdot \pi_i\). In order to compute real payoff, subjects have to divide their nominal payoffs \(P_i \cdot \pi_i\) by \(P_i\).

The real pay-off of subject \(i\) is given in Fehr, Tyran (2001) by:

\[
\pi_i = \pi_i \left( P_i, P_{-i}, M \right) \quad i=1,..., n
\]

where \(P_i\) stands for nominal price, \(P_{-i}\) is the average price of the other \(n-1\) group members, and \(M\) is nominal shock variable. Subjects are informed about payoffs of other subjects in the group, since \(x\) and \(y\) types players are present in nominal treatment. The need to recount nominal pay-off into real pay-off is a cognitively challenging task, which is the main barrier to optimal behaviour. This experiment aimed to test the ability of subjects in the economy to adjust to the equilibrium after full anticipated negative monetary shock depending on whether they face nominal or real environment. Results of Fehr and Tyran show that subjects, which have to face money illusion in form of nominal pay-offs adjust in much slower way in line with strategic complementarity. In other words, subjects (despite their rationality) have tendency to set prices close to the pre-shock price, consistently with development of the average price of \(n-1\) players if only few players in reality suffer from money illusion. This proved that money illusion is persistent phenomenon, causing substantial nominal rigidities. Consequently, implications are derived for the economy at the aggregate level, strengthening thereby New Keynesian predictions. In line with previous methodological discussion, question arises, whether it is possible to derive implications from such laboratory experiment at the aggregate level. Based on afore-mentioned methodological discussion, the following comments are worth mentioning with regards to laboratory experiment used for our illustrative purposes.

Firstly, one dimension concerns the well discussed number of subjects under scrutiny due to the belief that it is difficult to approximate macroeconomic world through a laboratory experiment with a small number of subjects. However, this macroeconomic experiment rests on micro-foundations in line with assumptions of New Keynesian economics built in experimental design. Figure 2 demonstrates more closely derivation of the aggregate level effects from individual behavior.

**Figure 2: From individual effects to the aggregate effects**

Wrong expectations

- Nominal economic environment
- Goals of individual \(j\) economic

Correct expectations

- Real economic environment
- Goals of individual \(i\)

Incentives

- Actions of individual \(j\) about action of \(i\)
- Actions of individual \(i\) about action of \(j\)

Aggregate outcomes

Microeconomic foundations of experiment are initially based on artificial economy with strategic complementarity, where pricing behavior of individuals is tested after the shock. Figure 2 shows that decision-making of individuals is tested in two different environments, based either on nominal or real pay-offs. Subjects j are expected to form wrong expectations about behavior of the other players i in an environment of nominal pay-offs. This is given by individual money illusion. Additionally this movement in wrong direction is multiplied since nominally confused individuals are followed even by rational subjects in an environment described by strategic complementarity. According to Fehr and Tyran (2001), these multiplied effects of individual money illusion are called indirect effects of money illusion. On the other hand, the control group in real environment is expected to form correct expectations about actions of j. Experimental results in this study proved the presence of strong indirect effects of money illusion, which are associated with nominal rigidities. This proof strengthens even more microeconomic foundations of New Keynesian theory and is additional reason for slow adjustment at the aggregate level, (others are imperfect information, implicit contracts, etc.). This is shown further in diagram, where actions of individuals are followed by aggregate outcomes. As a result the sample size need not be the relevant issue, as already suggested above and conclusion about nominal rigidities might be generalized at the aggregate level. Moreover, as emphasized by Duffy (2008), evidence from many auction experiments since Smith (1962) suggests that equilibration to competitive equilibrium occurs reliably with just a few individuals on supply or demand side market, so a large number of subjects need not be a necessary condition.

Secondly, this experiment enables the collection of a type of data less directly observed in the field and even of better quality in terms of causation as suggested by Duffy (2008). For instance, the examination of expectation formation after the shock is valuable output, which cannot be obtained otherwise. Additionally, it is extremely difficult to collect individual information sets regarding the actual price in the pre-shock and the post-shock phase in the field, needed for comparison with ideal equilibrium prices to identify nominal inertia, (Tyran 1999). Also identification whether the monetary shock is anticipated or not is a serious constraint. The experimental method (as opposed to field data) possesses an immense advantage in its control over the environment and information conditions. This is closely associated with causal relations (and internal validity), which are directly under experimental control, where the frame can be easily set by specification of treatment conditions as emphasized by Fehr and Tyran (2001), (2005). In this case environment of nominal and real pay-offs was created in experimental design. Examined phenomenon is secured by specification of treatment condition with possible elimination of all other factors, which should be constant, (See Table 1 for more details). Experimental group faces environment of nominal pay-offs, which represent the need of subjects to cope with some barrier, i.e. the need to recount nominal pay-off into the real pay-off. On the contrary, control group has to work with an environment of real pay-offs when making decision about the price of production. This does not represent any cognitive barrier and subjects should be pretty comfortable in this environment. When keeping other factors constant, we are able to examine phenomenon of confusion by nominal values by comparing the difference between these two examined groups. Since otherwise equal conditions are secured for both groups with sufficient elimination of all other effects, we may consider causal relation to be strong enough. This method “allows a dramatic reduction in the number of auxiliary hypotheses involved in examining a primary hypothesis”; thereby reducing Duhem-Quine problem as emphasized by Davis and Holt (1993, p.16).

Table 1: Treatment Conditions

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Other factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Nominal pay-off</td>
<td>Constant</td>
</tr>
<tr>
<td>Control Group</td>
<td>Real pay-off</td>
<td>Constant</td>
</tr>
</tbody>
</table>

Additionally, objections might be raised regarding the external validity of experimental results. However, in the case of the more theoretical character of the experiment examined in this case, there are fewer objections against lower external validity, which is supported also by study of Schramm (2005). In case of these experiments internal validity is predominant.

**Conclusion**

Experimental macroeconomics is becoming increasingly popular method used even by well-known economists for evaluation of various economic phenomena. Like every alternative method even this field is under scrutiny of opponents, who point out that validity of experimental outputs might not be satisfactory. This article tried to tackle the most burning methodological issues of experimental macroeconomics. Consequently, illustrative case study follows, which provides further justification for conduct of macroeconomic experiments. Study shows that although economists need to be careful in making generalizations based on the results of an experiment that involves a small number of subjects, still less objection should be made against using experiments in order to test predictions of macroeconomic models based on explicit micro foundations. These experiments might provide guidance for how subjects perceive examined phenomenon with consequent generalization. For instance, in the economy with multiple equilibria it might indicate what equilibrium subjects consider as more relevant. Additionally, experimental data should be understood as a complement to standard econometric analysis of field data if there is no possibility how to gather some specific data or if field data do not possess the character, which is desirable for examination of specific phenomenon. Indeed, experimental data offer possibility, how to secure sufficient internal validity in terms of micro-level causal relationships and exhibit better characteristic in this sense than standard field data. Last but not least, we have to bear in mind in vein of Duffy (2008) that all experimental work should be judged by its findings and not deficiencies, since all empirical methods have their strengths and weaknesses.

**References:**


Ricciuti, R. Bringing Macroeconomics into the Lab. University of Siena Department of Economics WP No.4, 2005.