# STATISTIC AND EXPERIMENTAL ECONOMICS

**Petr Obergruber**

**Abstract**

This paper will be short overview of the use of experiments in economics and changing paradigm about homo economicus by showing some aspect of behavioral economics as growing field of science, relationship between behavioral economics and experimental economics and show some application of those for example on labor market.

Moreover, as science (include psychology and behavioral economics) starts to be more spread in common knowledge and science results starts to be known in business and organizations, implications of economics experiments and field studies are use in important decision making as well as every day organization live. Experiments shows how behavioral economic works. Experiments and field studies are base sources for this discipline. Contrasts with standards economics approach which is mostly mathematical or statistical analysis (although experimental economics uses statistical methods too).

To highlight potential it will be introduce some of experiments (some of them organized in The Laboratory of Experimental Economics at The University of Economics in Prague) to show critical points in experiment´s architecture and evaluation. Special focus will be paid to recognition of the actual effect of the experiment from random variables.

**Key words:** Behavioral economics, experimental economics, experiment, measuring, performance, statistical analysis, lab experiments, field studies

**JEL Classification Codes:** C44, C91, C93, D03

## Introduction

Economics expects in coming years, perhaps decades, a significant change. Economic models mainly based on rational agents and strong mathematical apparatus cannot account with new researches in psychology, sociology or neural science. Those researches show that human behavior is much less optimal and rational as economics expected. Second, economics theories often works with premise so called ceteris paribus - with other things the same, which basically means that with this premise we study some reaction on single and only action.

Other variables are strictly constant. That is very strong premise and it seems that fulfilling it is not usual in observable reality.

Behavioral and primary experimental economics have two approaches how to compensate or bypass problem with ceteris paribus assumption. Those approaches are include in two possible ways of experiments – lab experiments and field studies. Lab experiments are usually designed in such a way that they study very narrow aspect of behavior. On the one hand the ceteris paribus assumption is broken but with only few variables in testing, we can still make good conclusions.

On the other hand there are field study. Field study does not try study isolated variables. Even if research change one or small amount of variables there are not aspiration to keep ceteris paribus assumption. If field study last for long time period (where are many circumstances changes), this is not reason to nullify or change study itself. This factor must be properly taken into account in conclusions. Pros of field studies is that we are testing behavior in perfectly or almost perfectly real world. Cons are changing conditions and no clear distinction between real impact and impact of other influences.

List and Imran (List, et al., 2010) are describing pros of fields experiments. Field experiments introduce exogenously timed variation in incentive structures. This opens up the possibility to identify the causal impact of monetary incentives on the behavior of individual workers, and on firm performance as a whole. Combining personnel files from human resource departments within the firm, with primary data collection that is inherent in field experimentation, allow researchers to examine the effect of monetary incentives on a range of margins of worker behavior, capturing both the intended and unintended consequences of incentive provision.

However, many economists have long been pessimistic that an experimental approach could offer such vivid illustrations of cause and effect in their field. For example, Samuelson and Nordhaus (1985) wrote in their introductory economics textbook:

The economic world is extremely complicated. There are millions of people and firms, thousands of prices and industries. One possible way of figuring out economic laws in such a setting is by controlled experiments. A controlled experiment takes place when everything else but the item under investigation is held constant. Thus a scientist trying to determine whether saccharine causes cancer in rats will hold "other things equal" and only vary the amount of saccharine. Same air, same light, same type of rat. Economists have no such luxury

when testing economic laws. They cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors. Like astronomers or meteorologists, they generally must be content largely to observe. (List, 2011)

The empirical gold standard in the social sciences is to estimate a causal effect of some action, but amidst the complexity of the real world, this is easier said than done. Economists have long worked on approaches that seek to separate cause and effect in naturally occurring data. A few decades ago, a standard approach was to use multiple regression analysis in an attempt to hold other factors constant. But economists have now taken to heart the old maxim that "correlation doesn't imply causation," and have in recent decades sought out a variety of other approaches. (List, 2011)

To be sure, we must work carefully when drawing conclusions based on the results of field experiments. Was the selection of participants into the treatment and experimental groups truly random? Do those who are not treated take some action as a result of being in the experiment, albeit in the control group, that they might not otherwise have taken? Is there something about the population being studied - perhaps risk-tolerance or persistence or belief that the treatment works - that warrants caution in generalizing the results to other populations? In the last few years, a lively literature has debated these and other issues that can arise in field experiments. (List, 2011)

Several experimental studies have provided evidence that suggest indifference curves have a kink around the current endowment level. These results, which clearly contradict closely held economic doctrines, have led some influential commentators to call for an entirely new economic paradigm to displace conventional neoclassical theory-e.g., prospect theory, which invokes psychological effects. (List, 2014)

## 1. Expected Utility Theory and Prospect Theory

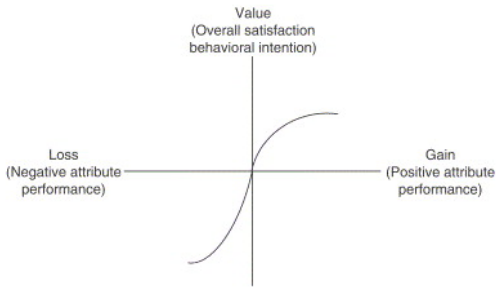
Expected Utility Theory (EUT) states that the decision maker (rational agent) chooses between risky or uncertain prospects by comparing their expected utility values, i.e., the weighted sums obtained by adding the utility values of outcomes multiplied by their respective probabilities. (Mongin, 1997) That basically means that rational agent know exactly which utility brings all prospects (or at least know which prospect is the best option) and choose the highest utility. Expected utility theory reigned for several decades as the dominant normative and descriptive model of decision making under uncertainty, but it has come under serious question in recent years (Daniel Kahneman, 1992). Kahneman and

Tversky presented model of choice called prospect theory, and explain the major violations of expected utility theory in choices between risky prospects with a small number of outcomes.

Expected utility theory define human as a rational agent with utility curve. This curve determines agent´s preferences and decisions in all situation and cases. Those definition of rational agent are based on some assumptions or axioms in this specific case. Primary those are independence axiom, completeness axiom, transitivity axiom and continuity axiom. Inter alia, the basic independence assumption, which is used in most theoretical and applied economic models to assess the operation of markets, has been directly refuted in several experimental settings (List, 2014) (Daniel Kahneman, 1991) (Daniel Kahmenan, 1990). These experimental findings have been robust across unfamiliar goods, such as irradiated sandwiches, and common goods, such as chocolate bars, with most authors noting behavior consistent with an endowment effect.1 Such findings have induced even the most ardent supporters of neoclassical theory to doubt the validity of certain neoclassical modeling assumptions. Given the notable significance of the anomaly, it is important to understand whether the value disparity represents a stable preference structure or if consumers' behavior approaches neoclassical predictions as market experience intensifies. (List, 2003)

Moreover Kahneman and Tversky (Daniel Kahneman, 1979) came with prospect theory, which take in case human´s psychology and describe how important is context, situation or risk in human´s decisions. Prospect theory works with curve which represent only current situation and evaluate gains differently than losses (shows value function).

**Picture 1 - Value function, source: (Daniel Kahneman, 1992)**



Reference point is a moment where we are when we are making decision. Strange shape of value function represents fact that human´s mind take losses much more intensively than equivalent gains. That means function do not represent absolute result but relative result relative to reference point. Here we are talking about dependent decision making. All possible results can be divided between gains and losses. By defining reference point there is possibility to create different curve for gains and losses at once.

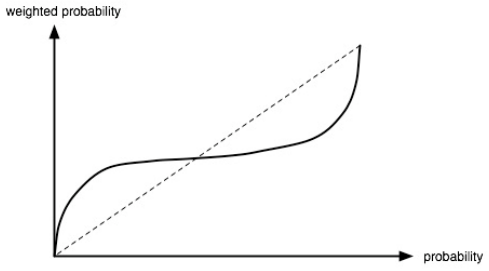
All effects describes in prospect theory has significant impact to human´s behavior but they are not incorporated in standard economics. Moreover those conclusions can be applied to a wide field of economics (including labor market).

Next principles consequent from prospect theory is no perfect perception all possible future states of the world. Prospect theory looks at decision making process in two steps:

1. The editing phase
2. The evaluation phase

The editing phase encompasses what are known as framing effect (effect of different context). The evaluation phase involves the decision process of choosing among options; this decision is influenced by two processes, one related to subjective value, the other to perceptual (Daniel Kahneman, 1979). With this decision making process there is significant distortion against reality and all of future states of the world get relatively different weight than by rational agent (shows weighing function).

**Picture 2 - Weighing function, source: (Daniel Kahneman, 1992)**




In the work Kahneman and Tversky (Daniel Kahneman, 1979) were performed research, within which were defined more general approaches to understanding risk and decisions making. Effects which entering in the decision-making process has been reported in the evaluation phase. Generally those effects can be divided into 2 groups: heuristics and influence of the context. As simplifying heuristics, shortcuts and systematic errors of human

decision-making have been identified more. The influence of the context is collectively referred as the influence of context or framework effect.

## 2. Work productivity experiment

One of the key elements of behavioral economics is measuring productivity at different incentives, stimulus and motives. If we accept fact that work performance is not influenced only by reward we can weigh up many potential things which can change human motivation.

### 2.1 Work performance experiment

The purpose of the experiment was to test how could be performance affected if we will reward same work first by fixed fee and second by variable fee. Simultaneously, same quantity of work was rewarded by same fee.

This experiment (organized in The Laboratory of Experimental Economics at The University of Economics in Prague) was designed for as easy measurement as possible. That is clear laboratory approach which guarantee easily evaluable data with minimum noise. On the other hand there is larger distance between behavior of participants in the experiment and real live behavior. Therefore the conclusion must be made in view of this fact.

Technically, there were organized 4 sessions where was undergrad students asked to rewrite some digitalized statistic[1] to Excel. Those statistics was very similar, therefore there was good comparison between works done. Two of those sessions have been rewarded by fix fee and two sessions have been rewarded by variable fee. Variable fee have been derive from average performance of fix fee groups (for same work same reward). Every session lasted 20 minutes long. (Petr Obergruber, 2014)

Average work performance of group 1 and 2 (fix fee groups) were 376.7 and average work performance of group 3 and 4 (variable fee groups) were 495, 0. That means that variable fee group achieve more than 30% better result than fix fee group. In that particular case, variable fee has shown as more effective way as work reward. (Petr Obergruber, 2014)

[1] There are large numbers of statistic which are digitalized by Google, but only by scanner. Therefore there is hard manipulation with them. Excel form has obvious pros. Moreover, behavioral economics experiments are considered more relevant if the measured work is needed or demanded and not just made up.

Those single results might indicate the advantage of variable pay (wages) but the answer is not so similar.

To expand the relevance of those results on longer time period than 20 minutes. At the end, every participant has been asked for 3 part questions*: Would you be willing to participate in similar experiment, if it lasted:*

1. One hour
2. One working day
3. One working week

By this results, people would not be willing extend their work in variable fee group as much as in fix fee group. One of possible explanation could be, that fix fee basically means higher reward, if I do not want work too hard.

But there is one more possible explanation: higher performance could be timely limited and people are able (or willing) to work on the identified level of performance only for short time. (Petr Obergruber, 2014)

### 2.2 Reselling common goods and market equilibrium experiment

In this study (List, 2003), researchers were gathered primary field data from two distinct markets to test whether individual behavior converges to the neoclassical prediction as market experience intensifies. Studied behavior was primary:

1. Trading patterns of sports memorabilia at a sports card show
2. Trading patterns of collector pins in a market constructed by Walt Disney World

List examined explicit statements of value in actual auctions on the floor of a sports card show. All of these markets are natural settings for an experiment on the relationship between market experiences, as they provide natural variation across individual levels of expertise.

Examining trading rates of sports memorabilia in an actual marketplace, List (List, 2003) observed an inefficiently low number of trades by naive traders, consistent with prospect theory. This evidence suggests behavior does change as market experience is accumulated, but two important issues remain. First, do consumers learn to overcome the endowment effect in situations beyond specific problems they have previously encountered? Second, given that List (List, 2003) was not primarily interested in testing the major theories,

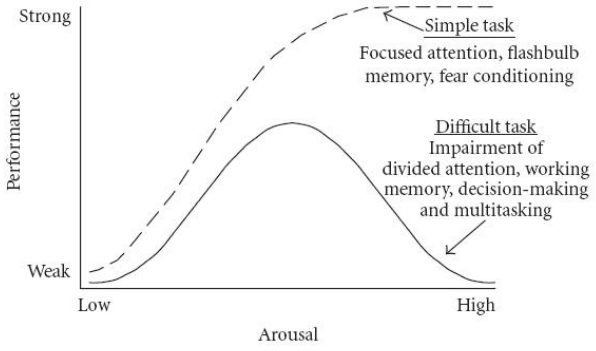
his results are open to interpretation. For example, his data may not properly delineate between prospect theory and neoclassical theory because experienced agents may have planned on reselling the good.

As in the first described experiment, we see results which cannot be explains the only way. The fact that we gathered statistically significant data does not guarantee perfect, complete or only explanation.

### 2.3 Thinking versus Doing – performance under stress experiment

This quite famous study (Racheli Barkan, 2010) is primary focused on people with jobs which are mostly mentally challenging (means difficult task in the picture) more than physical (means simple task in the picture). However, those are much harder testable than people with physically challenging work. As Ariely or Kahneman and Tversky (Daniel Kahneman, 1992) showed there is strong regularity in the performance of people according to characteristic of their workload (picture 3). This regularity is commonly known as Yerkes-Dodson law.

**Picture 3 - Yerkes-Dodson law, source: (David M. Diamond, 2006)**




Researchers' ambition in this experiment was tests first-class bankers. Bankers claim themselves as "special people" which work better under stress. With context of Yerkes-Dodson law for difficult tasks they basically said that their performance do not decreasing with increasing stress (or generally arousal) performance is increasing too.

Unfortunately for this there are intractable problems for the realization of a suitable experiment. First, we can hardly replicate stress condition which matches with bank job stress level. Normally we could offer high payment to participants which increase stress level. But in the case of bankers that payment should have been so high that it was unreachable for

researches. Second, even more important, bankers refused to participate in a similar experiment.

Without the ability to test bankers, researches looked for another source of data that could help understand how highly paid, highly specialized professionals perform under great pressure. The choice fell on basketball players, more specifically on "clutch players". Clutch players are paid much more than other players, and are presumed to perform especially brilliantly during the last few minutes or seconds of a game, when stress and pressure are highest. (Racheli Barkan, 2010)

Researchers found that the nonclutch players scored more or less the same in the low-stress and high-stress moments, whereas there was actually a substantial improvement for clutch players during the last five minutes of the games. So far it looked good for the clutch players and, by analogy, the bankers, as it seemed that some highly qualified people could, in fact, perform better under pressure. (Racheli Barkan, 2010)

But, when the first effort was to test banker's performance there are several more problems which must be solved. As researches looked after other possible explanation they found simple fact: there are two ways to gain more points in the last five minutes of the game. An NBA clutch player can either improve his percentage success (which would indicate a sharpening of performance) or shoot more often with the same percentage (which suggests no improvement in skill but rather a change in the number of attempts). So we looked separately at whether the clutch players actually shot better or just more often. As it turned out, the clutch players did not improve their skill; they just tried many more times. Their field goal percentage did not increase in the last five minutes (meaning that their shots were no more accurate); neither was it the case that nonclutch players got worse. (Racheli Barkan, 2010)

Moreover, the first effort was to test bankers not basketball players. But in the context of Yerkes-Dodson law we can predict with high probability that bankers will not improve their performance under high stress because their work is even more mental than playing basketball – we can assume with strong assumptions that banker´s work is better explainable with the "difficult task" curve of Yerkes-Dodson law than with the "Simple task" curve.

## Conclusion

Behavioral and experimental economics has quite different approach than mainstream economics. Inter alia we see on the three experiments above that statistic is a tool which makes link between science research and the story about human behavior. And this story about behavior is usually the final point of experiment or behavioral theory. In mainstream economics we often see statistical results and simple statement based on this result but there are not always whole story. We see mostly *what happens* but not *why it is happens*. This is big difference in statistic use between experimental and behavioral economics on the one side and mainstream economics on the other side[2]. The motivation to explain only *what happens* is not incomprehensible. Not only this approach is more accurate but it is also less open to interpretation. And what we are testing hypothesis with strong mathematical or statistical background and many variables, there could be simply no behavioristic explanation at all. Similar hypothesis (or theories) could count with so many imputes that human behavior is aggregated in few coefficients or variables. The forecasting ability is often tested on simulation or historical data.

But there is also strong motivation to explain *why is it happens*. This is primal goal of behavioral economics as we can see in one of the basic of behavioral economics – prospect theory. Behavioral economics is not only about finding regularities in human behavior (which can be irrational in some cases), but also find source of this behavior (psychological, sociological, evolutional or else). Obviously mainstream economics usually cannot work with those sources, mainly because it works with rational agent or some other sort of mathematical model[3]. Strong bind between behavioral and experimental economics comes from rejection of mathematical modeling of human behavior - instead, those two related fields try to explain human behavior in other ways. Behavioral economics in most cases use combination of psychology, sociology and other scientific disciplines (including mainstream economy, but behavioral economics does not prefer economics results before other disciplines).

At first glance it could seems that explanation *why is it happens* comes with more pros than *what happens* approach. But on the three experiments above we can see that *why is it happens* approach has problem with unclear conclusion, open to interpretation conclusions or

[2] By this paragraph author do not want to say mainstream economics does not consider action and reaction – such as *with increasing demand price rises as well*. The point is that hypothesis or theories with strong mathematical apparatus often overlooks some part of story neglects it or simply cannot assemble it.

[3] There are some exceptions as well. The most famous is probably the idea of monetary illusion.

conclusions which are simply wrong due to wrong usage of statistic – the results are attributed to wrong source. Those are problems which *what happens* approach does not faces.

On experiments mentioned above, we can see that we cannot easily test some behavior without treatment conditions and exceptions or some other sort of specific conclusion. On the first experiment conclusion from experiment could indicate clear advantage of variable fee before fix fee. But with effort to expand those results we see problem with the first intuitive conclusion. On second experiment we see that researches show perfect results but they left conclusion open to discussion. The third experiment again shows conclusion which on first look confirms hypothesis with "special people" which works significantly better under high level of stress and conclusions of Yerkes-Dodson does not apply to them. But after revision and examination of all possibilities we can see facts which correspondent with Yerkes-Dodson law.

## References

Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. Behavioral Law and Economics, 211-231.

Kahneman, D., Knetsch, J., & Thaler, R. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. Journal of Economic Perspectives, 193-206.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertainty Journal of Risk and Uncertainty, 297-323.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 263-263.

Diamond, D. (2007). The temporal dynamics model of emotional memory processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law.(Report). Neural Plasticity.

List, J. (2003). Does Market Experience Eliminate Market Anomalies? The Case of Exogenous Market Experience.

List, J. (2014). Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace.

List, J. (2011). Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off. SSRN Journal SSRN Electronic Journal.

Mongin, P. (1997). Duhemian Themes in Expected Utility Theory. French Studies In The Philosophy Of Science, 303-357.

Obergruber P., Hrubcová G. (2014). Behavioral economics and labor market. Behavioral economics and labor market. In MÍKOVÁ, I. (ed.). Proceedings of the 5th Scientific Conference for PhD. Students and Young Reserarches. 1, 2014, Vol. 1.

Racheli Barkan, Yosef Solomonov, Michael Bar-Eli, Dan Ariely. (2010). Clutch Players at the NBA. *manuscript - Duke University.*

**Contact**

Petr Obergruber
University of Economics, Prague
W. Churchill Sq. 4
130 67 Prague 3
Czech Republic
petr.grub@gmail.com