



Of morals, markets and mice: Be careful drawing policy conclusions from experimental findings!



Friedrich Breyer ^{a,*}, Joachim Weimann ^b

^a Department of Economics, University of Konstanz, P.O. Box 135, D-78457 Konstanz, Germany

^b Faculty of Economics and Management, Otto-von-Guericke-University Magdeburg, P.O. Box 4120, D-39016 Magdeburg, Germany

ARTICLE INFO

Article history:

Received 5 March 2015
Received in revised form 2 June 2015
Accepted 8 June 2015
Available online 12 June 2015

JEL classification:

D010
D030
D630

Keywords:

Erosion of morals
Market ethics
Double auction

ABSTRACT

Sometimes experimentalists draw far-reaching policy conclusions from their findings. However, this is dangerous if there is ambiguity with respect to the right interpretation of the results. A good example for this danger is the well-known study by Falk and Szech (Science, 2013), who conclude that “markets erode moral values”. If this were true, economists, who have emphasized the efficiency enhancing effects of markets for centuries, would have to reconsider their judgments fundamentally. In this note we demonstrate that the claim made by Falk and Szech is unfounded for at least four reasons and that their experimental results can be (and should be) interpreted, if anything, in the opposite way.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Experimental results often contradict conventional economic wisdom when participants in the lab do not seem to behave fully rationally or respond less to (financial) incentives than the theory of Homo Economicus would predict. Sometimes the experimentalist is then tempted to draw far-reaching policy conclusions, e.g. with respect to the efficiency of the price mechanism or the need to regulate certain markets. However, those conclusions can only be justified if the experiment in the lab is an exact reproduction of a real (market) institution, which is often impossible to achieve. If this limitation is not taken into account, economists risk throwing the baby out with the bath water.

The *Science* paper by Falk and Szech (2013a, 2013b) is a case in point. In their experiment, the authors employed a new design to evaluate the effects of different institutions on the willingness to pay of subjects for the life of a laboratory mouse and concluded that “markets erode moral values”. This article met with great interest in the media, at least in Germany. Several daily newspapers reported on it and the authors were invited to appear on TV shows for discussions with moral philosophers.¹ In addition, the Huffington Post reported on it with an article bearing the headline: “Markets Erode Morals, Let People Do Horrible Things: Study”.² The message seemed to confirm the widespread sentiment – originating from the financial crisis – that markets should be seen with great skepticism. Indeed, if it were true that markets erode moral values, economists, who have emphasized the efficiency enhancing effects of

* Corresponding author. Tel.: +49 7531 882568.

E-mail address: Friedrich.Breyer@uni-konstanz.de (F. Breyer).

¹ <http://www.cens.uni-bonn.de/team/board/armin-falk/press-releases-and-articles>.

² http://www.huffingtonpost.com/2013/05/13/markets-morals-study_n_3267995.html.

markets for decades, if not centuries, would have to reconsider their judgments fundamentally. In this comment we want to demonstrate that the experimental results described by Falk and Szech do not justify the far-reaching conclusions described above and that the experimental design used by Falk and Szech could be improved.

The experiment of Falk and Szech was conducted in Bonn, Germany, in May 2012 with 787 participants. The issue was the choice between receiving money (up to 20 euros) for allowing a laboratory mouse to be killed or forgoing the money and saving the life of the mouse. There were three different treatments. In treatment one, called “individual”, subjects could decide either to get 10 euros or to save the mouse. Treatment two, “bilateral market”, defined a more complicated structure in which the experimenter offered two players a total of 20 euros if they agreed to let a mouse be killed, and introduced a bargaining period in which the players essentially bargained how to divide the proceeds of this trade between them. In order to suggest that this bargaining procedure is a market transaction, the two players are named “seller” and “buyer” of the mouse, respectively, but this characterization is misleading since the roles of the players are completely symmetric. The third treatment, “multilateral market”, had the same structure, but with more than two players involved. The authors compare the “fractions of subjects who were willing to kill a mouse for monetary amounts below or equal to 10 euros” (Falk and Szech, 2013a, p.707, heading of Fig. 1).³ As this fraction was significantly higher in the latter treatments than in the “individual” treatment, the authors conclude that “market interaction displays a tendency to lower moral values, relative to individually stated preferences” (p.710, col. 2).

We have four major objections concerning the results and their interpretation by Falk and Szech, where our focus will be on a comparison of treatments one and two:

1. Contrary to the claim made by Falk and Szech, treatment one is a typical market situation, even more typical for real markets than treatment two.
2. The comparison of treatment one (individual) with treatments two and three (bargaining) does not allow clear conclusions because more than one treatment parameter was changed.
3. The statement made by the authors that “market interaction erodes moral values, relative to individually stated preferences” is unfounded because no independent preference statements by the subjects were elicited.
4. Even if it was true that markets erode moral values of individual participants, markets as social institutions could still lead to morally superior outcomes compared to other institutions.

It is worth mentioning that our critique does not aim at the *external validity* of the experimental results – which is a quite general problem with laboratory experiments. Quite to the contrary, we are questioning their *internal validity*, i.e. we doubt that the authors have interpreted their own results correctly.

2. Objection 1: the interpretation of the data

What is a typical market situation? Economists distinguish between different market forms characterized by the degree of competition. But in nearly all of these forms and, in particular, in competitive markets, consumers behave as price takers. This means they only decide whether or not to buy a good at a given price. Bargaining over prices takes place in some very particular markets such as those for houses, used cars and other second-hand goods in which most consumers trade rather infrequently. In the majority of all markets, consumers do not have any strategic leeway, but they just adapt their behavior to the given price. This also holds for decisions with an explicit moral component. Imagine someone who wants to buy a chocolate bar and finds a variety of “regular” chocolate and “fair traded” chocolate at the supermarket. Then she has to decide whether or not to pay the higher price for the fair-trade product knowing that the extra charge is a transfer payment to producers in developing countries. Once again, there are no negotiations and there is no bargaining; there is simply an individual decision. Thus treatment one (individual) in the Falk and Szech paper describes a typical market decision of a consumer who behaves as a price taker. The experimenter acts as the seller, and the offer is: paying a price of 10 euros for saving the life of a mouse. The price is fixed and the subject as the buyer only has to decide whether to buy the life of the mouse or not. This is exactly the same decision our chocolate consumer has to take when she chooses between regular and fair traded chocolate.

Falk and Szech run the “individual” treatment also in a slightly different way. In the “price list treatment” subjects were offered a list of increasing prices and they had to decide at what price they would prefer to get the money instead of saving the life of the mouse. The results of this treatment are identical with the “individual” treatment. This finding makes our point even more transparent:

The price list treatment makes use of a well-known procedure which is normally used to get information about the true willingness to pay for a particular item. The advantage of the price list procedure is that it is incentive compatible: given that the price is chosen at random, it is the best answer to state the true willingness to pay. And this is the best answer, because the subject is *in the position of a price taker*. It is rational to state the true preference because she cannot influence the price which is randomly chosen. Thus, it is not surprising that the price list treatment reproduces the results of the individual treatment. In both treatments subjects behave as they do in most of the markets they are confronted with during their lifetime: they respond to a given price.

On the other hand, economists know that small group bargaining situations differ tremendously from the typical market situation described above. The crucial point here is that in a typical bargaining situation the players have private information that they could use

³ The expression “willing to kill a mouse” is misleading because subjects did not have to kill the mouse themselves. The correct wording would have been “willing to let a mouse be killed”.

for *strategic behavior*. For example, each bargainer knows his own reservation price but not the corresponding price of his counterpart. If the strategic leeway given by this private information is used, then it is not certain that the bargaining process will end up in an efficient solution (Myerson and Satterthwaite, 1983). Even in very simple two-person bargaining situations such as the famous ultimatum game (Güth et al., 1982) experimental economists have observed very surprising results that differ strongly from what would be expected from a market outcome. Therefore, treatments two and three in the Falk/Szech paper do not describe typical market situations. Bargaining does not usually take place in competitive markets but in bilateral monopolies such as collective wage agreements.

We do not dispute that the double auction (DA) is indeed an institution that is able to reproduce market *outcomes*. This is why we argue (in our fourth point) that the outcomes of the institutions should be compared and not the stated willingness to pay. And the outcomes of the treatments “price list” and “double auction” do not differ. Furthermore, we still do not think that the DA used in the *experiment* can be treated as a typical market situation. First, although auctions have a growing importance in e-business, most of these auctions are not double auctions but have other designs (eBay uses a second price auction). Second, the DA used by Falk and Szech has a very special feature. It can be interpreted as a complicated public-good game in which the price of the public good (the survival of the mouse) is fixed at 20 euros but the willingness to pay of each player is private information. In contrast to more traditional public-good games, the default in this game is that the public good is provided, whereas the players can cancel the provision if they can agree on the division of the savings. It is crucial for the interpretation of the results that this is neither a usual DA nor a typical market situation.

If we accept the criterion of comparison proposed by Falk and Szech, namely the share of subjects agreeing on the death of a mouse for 10 euros or less, which is 45.9% in treatment one and 72.2% in treatment two, and if we accept for the moment that killing mice violates a moral norm, the right interpretation tends to be opposite to the one given by the authors: in everyday market situations, moral norms play a more prominent role than in less frequent bargaining situations.

3. Objection 2: the comparison of the treatments

Economists use a simple method to investigate things experimentally. They design treatments that differ in only *one* specific aspect from each other and then compare behavior in these treatments. If subjects behave differently between treatments, it can be reasoned that the treatment differences are important for individual behavior. Obviously, such reasoning is impossible if more than one design element is changed at the same time. This is the case in the Falk and Szech experiment. There are at least three important design elements that change from treatments one to two and three.

First, the price setting mechanism is changed. While in treatment one the price is fixed, in the other two treatments the price is established in a collective bargaining process in which asymmetric information and strategic behavior play an important role (see above). Second, the number of people who decide whether the mouse dies is changed. This is important because, as we know from other experiments, groups decide differently from individuals (Kerr and Tindale, 2004; Kocher and Sutter, 2005). One reason for this may be that for collective decisions the (moral) responsibility is diluted.

A third change pertains to the payout rule. While in treatment one the question was asked only once and there was certainty that the amount earned in that round would be paid out, subjects in treatments two and three played the same game 10 times and only one randomly chosen round was used for payout. Therefore, participants in a particular bargaining round knew that there was only a 10% chance that they would actually earn the amount agreed on in that round.

If the change of institutions covers the change of totally different aspects of a decision situation, it is crucial to be able to identify which of these aspects are responsible for the change in behavior. Falk and Szech interpret their results in exactly this sense because they identify the *market environment* as being responsible for the change in the stated willingness to pay. By doing so, they exclude other “specific reasons”, that might be of importance (number of people involved in the decision, the framing, the bargaining mechanism, the random payout mechanism, the fact that one subject is called the “owner” of the mouse and so on). How important these other specific reasons are is emphasized by Falk and Szech themselves when they point out that the killing rate in the multilateral treatment of the DA was much higher than in the bilateral one.

4. Objection 3: the ignorance of individual preferences

The conclusion by Falk and Szech that “markets erode moral values” is problematic for a third reason. Their proposed benchmark for moral values is formed by “individually stated preferences” (Falk and Szech, 2013a, p.710). But individual preferences for the life of animals were not elicited directly but inferred from the WTP revealed in treatment one. But this is not a statement on the moral acceptability of killing mice as such. It is conceivable that a respondent finds killing mice morally wrong (and objects to animal experiments) but is not willing to pay his own money for the survival of one particular mouse because he thinks that this is the responsibility of the owner of the mouse. On the other hand, one might find killing mice morally acceptable but is charmed by the picture of the “cute” mouse shown by the experimenter and is therefore willing to pay to save this particular mouse.

It would have been preferable to ask subjects before the experiment directly if they found killing mice morally wrong. Many young people grow up with cats as family pets and are therefore used to seeing how mice are killed. In this context it is important to note that subjects were not informed before the experiment that these particular mice were surplus mice from a genetic laboratory that would have been killed anyway. Thus it is not at all clear that Falk and Szech's own apparent moral value that killing mice is unethical is shared by participants of this experiment.

5. Objection 4: an alternative criterion for comparing social institutions

Falk and Szech seem to propose to judge the ethical quality of an institution like the market by their effects on the behavior of *individual* participants. They are thus in agreement with moral philosophers such as Michael Sandel. Being a virtue ethicist, Sandel (2012) proposes to ask whether creating a market for certain goods changes their character or the way people think about them, and to judge the ethical quality of the market by the answer to this question.

In contrast, economists usually take a consequentialist viewpoint and base their moral evaluation of institutions such as markets or bilateral bargaining arrangements on the *outcomes* brought about by these institutions. In this case, the relevant outcome is not what offers subjects make or accept in the process of bargaining, but how many mice are killed. Yet by this measure, treatments one and two do not differ significantly since the shares of mice killed are 45.9% and 47.7%, respectively. Thus no matter what labels are used to characterize the institutions created by Falk and Szech in their lab experiment, there is no significant difference in the measure of moral value that is most common in economics.

6. Concluding remarks

We conclude that the claim made by Falk and Szech (2013a) that “markets erode moral values” is not at all supported by the results of their experiments. Their results demonstrate nicely that in the case of a typical market decision, consumers simply follow their preferences and that under these conditions moral values play an important role. In the case of bilateral bargaining, preferences-driven behavior is superimposed by strategic reasoning and therefore more subjects reveal a willingness to pay less than 10 euros for the life of a mouse. Furthermore, as we do not know the *moral persuasions* of the subjects in this experiment, nothing can be said about the erosion of their moral values. And finally, if what people *do* and what the outcome of the coordinating mechanism is count rather than what they *say*, there is no difference between these settings.

Acknowledgments

The authors are grateful to Mathias Kifmann, Ronnie Schöb and an anonymous referee for valuable comments.

References

- Falk, A., Szech, N., 2013a. *Morals and markets*. *Science* 340, 707–711.
- Falk, A., Szech, N., 2013b. Supplementary Materials for “Morals and Markets”. www.sciencemag.org/cgi/content/full/340/6133/707/DC1.
- Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388.
- Kerr, N.L., Tindale, R.S., 2004. Group performance and decision making. *Annu. Rev. Psychol.* 55, 623–655.
- Kocher, M.G., Sutter, M., 2005. The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *Econ. J.* 115, 200–223.
- Myerson, R.B., Satterthwaite, M.A., 1983. Efficient mechanism for bilateral trading. *J. Econ. Theory* 29, 265–281.
- Sandel, M., 2012. *What Money Can't Buy: The Moral Limits of Markets*. Farrar, Straus and Giroux, New York.