

Behavioral and Experimental Economics Do Inform Public Policy*

Arno Riedl
CESifo, IZA & Maastricht University

February 26, 2010

Abstract

Experimental and behavioral economics are well established branches in the economic science. This essay presents and discusses some results and behavioral regularities from these fields which are of potential and actual importance for public policy. After a brief introduction of what experimental and behavioral economics is, some important behavioral regularities as presentation and framing effects, pro-social behavior, and reciprocity are introduced and it is reported how they interact with prominent trading institutions, taxation, and social and individual well-being. Throughout some implications for public policy are discussed.

*This paper is based on my keynote lecture ‘Sociality and Institutions’ presented at the workshop on Behavioral Public Economics at the Venice Summer Institute 2009 and on my inaugural lecture ‘Facts and Fiction in Public Policy: How Behavioral and Experimental Economics Can Inform Public Policy’ at Maastricht University. I owe thanks to the participants of the Workshop on Behavioral Public Economics at the Venice Summer Institute 2009 and to the organizers of this workshop for their hospitality. I also thank an anonymous referee and the editor for their helpful and valuable comments on an earlier version of this paper which was circulated under the title ‘Behavioral and Experimental Economics Can Inform Public Policy: Some Thoughts’.

“After consulting my advisory board of experimental and behavioral economists, I am confident that the reframing proposed in the new public policy program will increase subjective well-being by 34 percent and pro-social orientation by 27 percent at almost no cost.” (taken from Amir et al., 2005, re-phrased by the author)

1 Introduction

Despite the largely unpredicted recent financial crises and accompanied economic downturn, most, if not all, recent public policy choices still rely on the traditional economic concept of rational economic man and woman, the *homo economicus*. Interesting examples in this respect are recent reforms in the Netherlands concerning the health insurance market and the markets for the supply of gas and electricity. In all cases important principles of the reforms were (i) more competition between suppliers, and (ii) more choice possibilities for the consumers. The main arguments in favor of such reforms are based on the received wisdom among economists that increased competition and an enhanced choice set for consumers are ultimately (consumer) welfare increasing.

To a large extent the presumed positive outcomes rely on the assumption that consumers will exercise their power to choose. For the energy market reform in the Netherlands, the expectations about the exercise of consumer power have been largely disappointed, however. At least if one measures this power by the fraction of consumers switching the supplier, because only a negligible fraction of consumers switched to an alternative electricity or gas provider. After the reform of the health insurance sector consumers seemed to take up their power to choose, though. It was reported that with the introduction of the new health insurance system in total 21 percent of consumers have changed their insurance company (de Jong et al., 2008). However, also this switching behavior looks less impressive if one takes into account that a majority of those individuals who switched the insurance company did this within a so-called collective agreement where the employer (or another collective) and not the individual chose the insurance company. Finally, only about 9 percent of all ‘switchers’ decided to do so on an individual basis. Furthermore, in 2008 the percentage of people switching their health insurance provider decreased to a mere 4 percent (NIVEL, 2009).

However, any well-trained economist will argue that (non-)switching behavior *per se* does not yet mean that consumers did not make the correct choices. Indeed, revealed preference theory states that those who did not switch simply reveal that they had already chosen their utility maximizing insurance package and/or energy supplier and, hence, had no reason to switch.

But let us be a little bit skeptical and ask if there is any way to assess if consumers

indeed made ‘good’ choices. Unfortunately, there is no study (at least I am aware of) that investigates this question for the mentioned recent reforms in the Netherlands in a systematic way. Yet, an evaluation of a similar reform of the Swedish social security system in 2000 may help us to get some clues about how such reforms may work out and whether consumers indeed make the good choices traditional economic theory assumes. Cronqvist and Thaler (2004)¹ investigated consumer choice behavior after the introduction of the new system. In this system participants are allowed to choose their own portfolios, but there is also one ‘default’ fund that is selected automatically for those who do not actively choose. The authors asked if, compared to the default, active choices are better choices. They conclude that “it would be hard to make the case on an *ex ante* basis that the actively selected portfolios were better than the default fund” (Cronqvist and Thaler, 2004, p. 427).² In addition, in the first three years as well as up to 2007 the actively chosen funds did also worse in terms of returns (see Thaler and Sunstein, 2008, p. 427). Another interesting observation is that after the Swedish government reduced the campaigns advertising active choice most people (90 percent) opt for the default fund and almost nobody makes any changes to the chosen portfolio or switches the chosen fund.³

This study strongly suggests that consumers do not (always) make the wise choices traditional economic models assume. However, there are too many unobservables (e.g., risk preferences, self-selection effects) that may influence behavior and one may not want to draw too strong conclusions from this field evidence. Therefore, in the remainder of this contribution I shall present ‘clean’ evidence that standard assumptions of economic models are indeed often violated and argue that neglecting the observed ‘non-standard’ behavioral regularities will lead to wrong predictions and worse public policy than necessary. The plan of the rest of paper is as follows. First, I will shortly discuss the scope of experimental and behavioral economics. Then I shall present important examples of violations of standard behavioral assumptions based on questionnaire studies and laboratory experiments. Thereafter, I shall link up these observations with questions concerning public economics and public policy. Finally, I shall present some ideas about interesting and important further research directions.

¹See also Thaler and Sunstein (2008, ch. 9).

²In particular, in comparison to the default fund, actively chosen portfolios contained a higher equity exposure and much more local concentration (e.g., almost 50 percent of the equities are from Swedish firms), required more active management, and had higher fees.

³When the system was introduced in 2000 two-thirds of participants actively selected a portfolio on their own. Interestingly, the proportion of people actively choosing their own portfolio decreased to 17.6 percent in 2001 and to only 8 percent in 2006.

2 The coherence of experimental and behavioral economics

2.1 Experimental economics

“The principle of science, the definition almost, is the following: The test of all knowledge is experiment. Experiment is the sole judge of scientific ‘truth’.”
(Feynman, 1964, chapter 1)

What precisely is experimental economics? As the name suggests it is the branch of economics that uses experiments to investigate human behavior in economic decision situations. Experimental economics is a method that brings real people to the laboratory (or the field) where they make real choices with which they earn (or lose) real money (Plott, 1982). An important feature of the method of economic experimentation, which distinguishes it from traditional empirical economic research, is that experiments allow the researcher to tightly control the environment in which people make choices. The controllable components of this environment comprise technologies, initial endowments, action spaces, timing of actions, accessible information, context and to some limited, but important, extent also preferences. Varying these elements in a controlled way allows *ceteris paribus* inference and the isolation of true causes of change in human behavior to an extent unattainable by other methods of investigation. Additionally, laboratory experiments can be replicated by other researchers under the same or different conditions, thereby assessing the robustness of obtained results. As economists, however, we know that there is nothing like a ‘free lunch’. An often raised concern about the experimental method is the presumed lack of external validity of the obtained results. This is indeed a serious concern, in particular, when one aims at using the experimental method for informing public policy. I shall therefore come back to this issue at the end of the paper.

There are two hard and fast principles which experimental economists subscribe to and which also differentiate economic experiments from most experiments in psychology and marketing. Firstly, in economic experiments the monetary earnings subjects receive depend in a transparent way on the choices made by the subjects. The reason for the application of this principle is that it is one story to merely tell what one would do in a particular situation, but another story to actually take a particular action if it is linked with monetary consequences.⁴ Secondly, deception of subjects is effectively banned. The main reason being that once deception is used it is likely that it will leak out. Subsequently the knowledge of being deceived will spread through the subject population, which seriously undermines an important advantage of experiments, namely having control over the

⁴For evidence that monetary incentives indeed make a difference, see, e.g., Camerer and Hogarth (1999) and Forsythe et al. (1994).

information and knowledge subjects have concerning the economic situation they are in. For a discussion of the effects and costs of using deception in experimental research see, e.g., Ortmann and Hertwig (2002) and Jamison et al. (2008).

2.2 Behavioral economics

“But also needed is imagination to create from these hints the great generalizations to guess at the wonderful, simple, but very strange patterns beneath them all, and then to experiment to check again whether we have made the right guess.” (Feynman, 1964, chapter 1)

Behavioral economics is not a synonym for experimental economics although they share some common grounds. Initially behavioral economics strongly relied on empirical evidence generated in psychological and economic experiments. Nowadays, behavioral economics defines itself broader as an approach incorporating evidence - not necessarily experimental - from psychology and other disciplines to explore the limits of existing models of behavior and create new simple ones that can explain actual behavior in a better way than current models are able to.

Importantly, behavioral economics does not abandon the disciplining strict formality which distinguishes traditional theoretical economic modeling from ‘softer’ approaches in some other social sciences. It also does not try to fit a new model for each new behavioral anomaly or regularity but rather seeks for parsimonious models and themes that can be applied to many different domains. Behavioral economics is also not another sub-discipline, next to labor economics, public economics or all the other sub-economics, but understands itself as a modeling approach which should be applicable to a wide range of economic questions. The ultimate aim being “generating theoretical insights, making better predictions ..., and suggesting better policy.” (Camerer and Loewenstein, 2004, p. 3) In particular, the latter is also theme of this contribution.

3 Examples of behavioral regularities

3.1 Presentation and framing effects

Traditional economic reasoning is usually silent about possible effects of the presentation or framing of a decision situation. For instance, from the viewpoint of revealed preference theory it simply does not matter whether one has to make choices in sequence or simultaneously. Rational economic man and woman will always choose according to their true preferences. These preferences are assumed to be well-behaved, coherent and invariant with respect to superficial variations in the way a choice problem is presented. However,

one might start to wonder then why grocery and other stores are pricing their products that often with 9.90, 19.99 and similar - just a little bit below round number - prices. Is this just co-incidence and does it happen that the true marginal costs of all these products are exactly these prices? Probably not.

Let me pose the problem more concretely by discussing an example adopted from Simonson (1990) who was among the first to demonstrate with the help of an experiment that it may greatly matter whether consumers have to choose from an array of products simultaneously or sequentially. Suppose, one is entering a grocery store today and this store offers to choose one snack out of six different brands of snacks for free.⁵ The same happens in one week from today and in two weeks from today. Hence, in each subsequent week one is free to choose the most preferred snack for free. Now, consider another grocery store that also offers free snacks for the next three weeks, again one per week. However, in this store the rule for choosing the free snack is slightly different. Here one has to choose today which brand of snack one would like to receive today, which one to receive in one week, and which one to receive in two weeks. Thus, the only difference is that in the second store one chooses today for today and the next two times whereas in the first store one decides in each week on the spot.

When facing such offers, for a homo economicus, doing the right thing is a pretty straightforward decision. Just choose the most preferred brand. Note, that this does not mean that one is going to choose the same brand for each of the three weeks, because one might like variety in snacks. Importantly, however, whether the decision has to be made simultaneously or sequentially should not make a difference. Consequently, it should not matter if one is confronted with the choice sequentially three weeks in a row as in the first store or if one has to choose at once for all three weeks as in the second store. One might not choose the same snack in each week, however, the variety of snacks one chooses should not differ under the two conditions. Now, the question asked was if real consumers act in this way. At least, students in a laboratory study (Simonson, 1990) did not do that. In the sequential choice (grocery store 1) only 9 percent chose a different snack in each week whereas in the simultaneous choice for sequential consumption (grocery store 2) this was the case for 64 percent of participants. These are by no means small differences and even a skeptic should be ready to admit that these results are hard to reconcile with the assumption of stable and/or coherent preferences. In a follow-up study Simonson and Winer (1992) corroborated the laboratory findings in the field by using scanner data of actual yogurt purchases in a grocery store. They find that the variety of flavors chosen significantly increases with the number of purchases per occasion. The observed choices in the laboratory and the store strongly indicate that revealed preferences systematically depend on the way the choice set is presented, a dependence not accounted for in traditional

⁵If one does not like snacks at all it can be replaced by brands of other products, e.g. beer or yogurt.

economic models.⁶

A second prominent example which probably casts even more doubt on the assumption of coherent and stable preferences is the famous study, known as the ‘Asian disease’ problem, by Tversky and Kahneman (1981). Tversky and Kahneman conducted questionnaire studies with students (Stanford University and University of British Columbia) where they asked them to indicate their preference concerning different programs proposed to combat an unusual Asian disease. To study potential framing and presentation effects the problem was presented in two economically equivalent but presentational different formulations. Figure 1 reproduces the original text.

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:	
Problem 1: [$N = 152$]	
If Program A is adopted, 200 people will be saved.	If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.
Which of the two programs would you favor?	
Problem 2: [$N = 155$]	
If Program A (C) is adopted, 400 people will die.	If Program B (D) is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.
Which of the two programs would you favor?	

Figure 1: The ‘Asian disease’ problems of Tversky and Kahneman (1981)

Problem 1 presents the decision situation in a *positive* frame by emphasizing that lives can be *saved*. Problem 2, in contrast, presents the very same options in a *negative* frame by emphasizing that some people will have to *die*. Obviously, Programs A in Problem 1 and 2 are identical since in both cases 200 people will be saved and 400 people will die for sure. The same is true for Programs B, where in both problems 200 people will be saved and 400 people will have to die, in expectations. Hence, whatever people like more, there should be no significant difference in revealed preferences between the two problems.

⁶These observations are consistent with the concept of ‘choice bracketing’ and its specific consequence of ‘taste change’ (Read et al., 1999). The former refers to the fact that the way people make decisions, narrowly or broadly, affects their choices. The latter refers, specifically, to the effect that the choice people make today can change their tastes and, hence, influence their choices in the future. The emergent property of ‘diversification bias’ seems to be a robust phenomenon and is replicated in several studies (see Read et al., 1999, p. 178).

Figure 2 depicts the frequency of actual choices. When confronted with Problem 1 (positive frame) an overwhelming majority of 72 percent opts for program A that saves 200 people for sure whereas only 28 percent opts for the risky program where 200 people are saved only in expectations. When confronted with Problem 2 a dramatic shift in revealed preferences occur. Now, only a minority of 22 percent goes for the sure outcome of 400 dead people but 78 percent are ready to accept the risky choice where 400 people die only in expectations. It should be obvious that such a strong framing effect effectively inducing ‘revealed preference reversal’ is hard to square with the assumption of coherent and stable preferences.

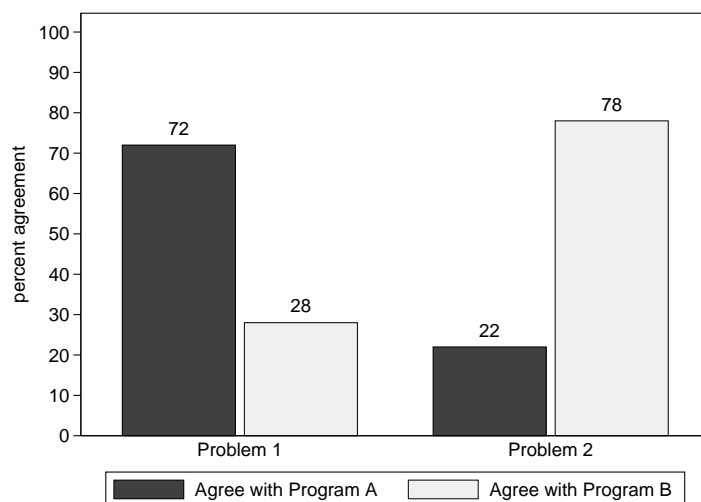


Figure 2: Actual choices in ‘Asian disease’ problems (Tversky and Kahneman, 1981)

The offered options in the Tversky and Kahneman study are both rather undesirable, in the sense that one has to choose between sacrificing more or less people. Unfortunately, this makes them representative for many decisions made in the public domain. Consider, for instance, public investment decisions, especially, investment in infrastructure concerning safety. The decision not to invest in more secure highways or railway infrastructure means to effectively decide to accept deaths that otherwise could have been avoided. Similarly, not investing into research for an influenza vaccine means to risk avoidable deaths in case of an outbreak of influenza. Many more examples could be given from airport security to school crossing guards (‘lollipop man’).⁷

“ ... in situations where self-interest and ethical values with wide verbal alle-

⁷A nice and rather harmless example where the management of a semi-public enterprise seems to take framing effects into account is the following: since only recently the Dutch railways do not use the word ‘delay’ any more when announcing that a train arrives late at the railway station. Instead, the phrase that “the train will arrive in a few minutes” is used.

giance are in conflict. Much of the time, most of the time in fact, the self-interest theory ... will win.” (Stigler, 1980, p. 176)

3.2 Morality and reciprocity

Although the neoclassical concept of utility is broad and flexible (critics might even say tautological and hence without content) and, therefore, in principle not restricted to narrow selfish preference orderings, most work and almost all applications in (public) economics assume that people are narrow-minded selfish material wealth maximizers. One might argue that, as long as models based on the assumption of narrow self-interest describe behavior of real people sufficiently well and, hence, make correct predictions that can be used to evaluate and validate public policy there seems to be no reason to abandon such models. This is precisely the argument endorsed by Milton Friedman, who argues that theories should be judged by the accuracy of their predictions but not by the accuracy of their assumptions (Camerer, 2005). However, if not only the basic assumptions are counter-intuitive (and empirically proven to be wrong) but also the models’ predictions incorrect or at least misleading then this approach is in deep trouble.

In the following I shall describe two prominent simple games where models based on narrow material self-interest turned out to do a bad job in predicting actual behavior. These examples will show that fairness considerations and, in particular, an inclination towards reciprocal behavior are important constituents of human behavior. I shall distinguish between *negative reciprocity* and *positive reciprocity*. Negative reciprocity describes the tendency to respond to an unkind act with an unkind act, whereas positive reciprocity describes a kind response to a kind course of action.⁸ The following examples will make these differences clear.

3.3 Negative reciprocity - the (mini) ultimatum game.

The ultimatum game (Güth et al., 1982) can be interpreted as a negotiation or bargaining situation that is stripped down to its most important constituents. It is a situation involving two people where one individual can make a take-it-or-leave-it offer to the other individual who can, indeed, take it or leave it.⁹ Figure 3 depicts - for the sake of the argument - an even further boiled down version of the originally investigated ultimatum

⁸In a sense positive and negative reciprocity may be viewed as just two sides of the same medal, since a non-unkind act is obviously a kind act. However, different emotions may be involved (e.g., anger vs joy) with negative and positive reciprocity, which is likely to make the responses psychologically and physiologically different. Additionally, reference points of fairness are important for the judgment of kind and unkind behavior.

⁹Note that such (or similar) situations are not uncommon in everyday life. For instance, it is akin to shopping in Western supermarkets or shops where one is usually not negotiating the price of the product but rather takes it or leaves it (on the shelves).

game (adapted from Falk et al., 2003). There one player, say Peter, has received 10 euro which he has to split between himself and, say Rita, in a take-it-or-leave-it way. If Rita accepts the offer both receive money according to the proposal. If she rejects it, then neither Peter nor Rita receive any money. For simplicity, Peter is given only two possibilities how to split the money. He can make a rather unkind offer, “I take 8 and you get 2”, leaving most of the money for himself or he can decide to be kind and propose to split the money evenly, “I take 5 and you get 5”. Rita, faced with one of these offers, has to decide whether to accept the offer or to turn it down. Traditional economics assuming narrow selfishness tells us that, because more money is better than less money, Rita will accept any offer. In terms of the figure it means that Rita will be kind after a kind proposal (5,5) and will also be kind after an unkind proposal (8,2).

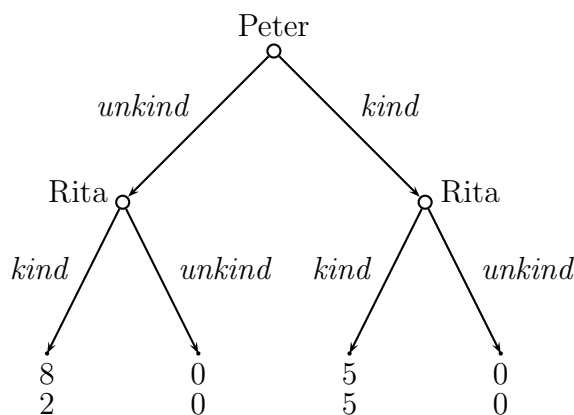


Figure 3: The mini ultimatum game

The empirical facts, however, deviate significantly from this prediction. A typical, qualitatively and quantitatively representative, result generated in many experiments is shown in Figure 4. It shows the percentage of accepted and rejected offers for both possible proposals. As one would expect the kind proposal of (5,5) is never followed by the unkind response of a rejection. The situation looks quite differently, however, if an unkind proposal of (8,2) is put on the table. Such a proposal is often followed by an unkind response, namely rejection. In the reported experiment this happens in more than 40% of the cases. It is important to see that the unkind response is costly, leaving both players without any monetary gain. This is precisely what makes it incompatible with traditional economic reasoning assuming narrow selfishness.¹⁰

¹⁰Actually, the rejection rates for offers of only about 20 percent of the whole pie are usually higher than the 40 percent reported here. A likely reason is that Falk et al. (2003) applied the so-called strategy method where subjects have to decide upon acceptance and rejection before they know the actual choice. That is, they make their decision in a cold (emotional) state, whereas responses to actual offers are made in hot (emotional) states (see Loewenstein, 1999).

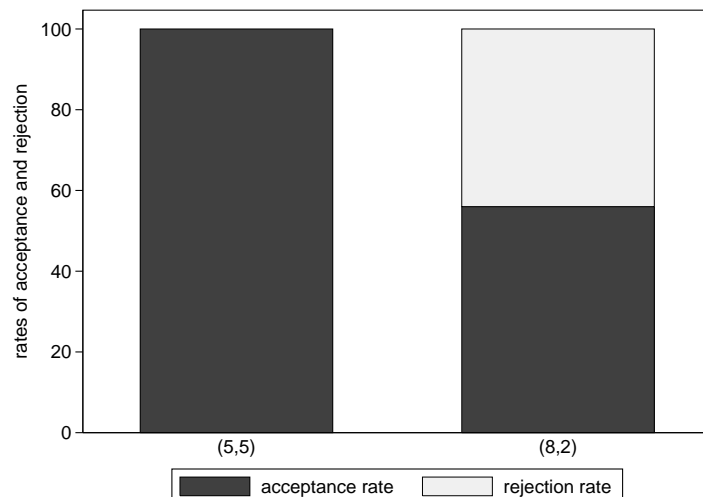


Figure 4: Acceptance and rejection rates in mini ultimatum game (data source: Falk et al. (2003))

3.4 Positive reciprocity - the gift-exchange game.

Probably the best known example for the existence of a predisposition towards positive reciprocity stems from a game coined gift-exchange game.¹¹ In economics, the basic idea behind this game dates back (at least) to Akerlof (1982), who argued that gift-exchange is an important constituent of labor contracts that are genuinely incomplete. The incompleteness of the contract refers to the fact that the effort exerted by an employee is often not verifiable because it cannot be observed by the employer and/or not be enforced by a third party. Akerlof's theoretical model, however, relied on assumptions about economic behavior of employees that are at odds with the assumption of narrow selfishness. Namely, that employees respond to higher wages positively, in the sense that higher wages make them exert higher and more costly effort.

In Fehr et al. (1993)¹² this idea is put to a test in the experimental laboratory. In fact, the implemented situation more generally represents any kind of patron-client or principal-agent relationship where contracts cannot be (perfectly) enforced. More concretely, consider the following situation (based on Riedl and Tyran, 2005): A number of people is divided into a set of 'employers' ('buyers') and a (larger) set of 'employees' ('sellers'). The rules of the game are as follows. If an employer hires an employee who provides effort, e , and receives a wage w , then the employer's earnings π are $30 + 10e - w$. That is, the employer earns a lump-sum of 30 plus 10 times the effort e exerted by the employee minus the wage w paid. The wage is between 1 and 100 ($1 \leq w \leq 100$) and

¹¹A game very similar in nature is the so-called trust game (Berg et al., 1995).

¹²See also Fehr et al. (1997, 1998).

the effort between 1 and 10 ($1 \leq e \leq 10$). The earnings u to the employee is then the wage, w , he receives minus a ‘cost of effort’, $c(e) = e$ plus some lump-sum payment of 4 ($u = w - e + 4$). The sequence of actions is as follows. The employer first offers a ‘contract’ specifying a wage w . When the contract is signed the employee receives the wage *before* exerting any effort. Only, after the wage is paid out the employee decides on his effort, where he is completely free to choose any level of effort. Importantly, each employer-employee interaction is anonymous and essentially one-shot. That is, there is no possibility for reputation building or retaliation.

What will be the outcome of such an interaction? Consider first the employee who has received a wage and has now to decide on the effort level. Under the assumption of narrow selfishness, any employee will choose the effort level with the lowest cost, $e = 1$, no matter how high or low the wage received is. In effect, choosing a higher effort level only decreases his total earnings. A rational and selfish employer will perfectly anticipate this behavior and, hence, offer the lowest acceptable wage, $w = 5$, which is still accepted.¹³

However, in the experiment this predicted outcome is very rarely observed. Moreover, gift exchange is observed as a very strong behavioral regularity. Figure 5 illustrates this. The figure plots, on the vertical axis, the (average) effort chosen by the employees on the average wage offered by the employers, on the horizontal axis. Under the assumption of narrow self-interest employees should always choose the lowest effort level implying no positive (or negative) relationship between effort and wages. This is reflected by the straight line running through $e = 1$. The actual data, however, show a strikingly different pattern: effort is clearly and strongly increasing with the wage received. In the figure this is reflected by the increasing line, which connects average effort levels for wages smaller 20, between, 21 and 40, 41 and 60, and larger than 61. Hence, in conclusion this and many other studies (for a recent review see Fehr et al., 2009) clearly show that people respond positively reciprocal. An important side effect of the observed gift exchange is that it increases efficiency (in terms of surplus maximization) and decreases inequality in earnings, in comparison to the benchmark outcome predicted under traditional behavioral assumptions.

4 Behavioral regularities, public economics, and public policy

One might wonder what all of this has to do with public economics and public policy. In the following I shall argue that the described behavioral regularities can indeed be of eminent importance for public economics and policy.

¹³An employee without a contract receives a payment of 4.

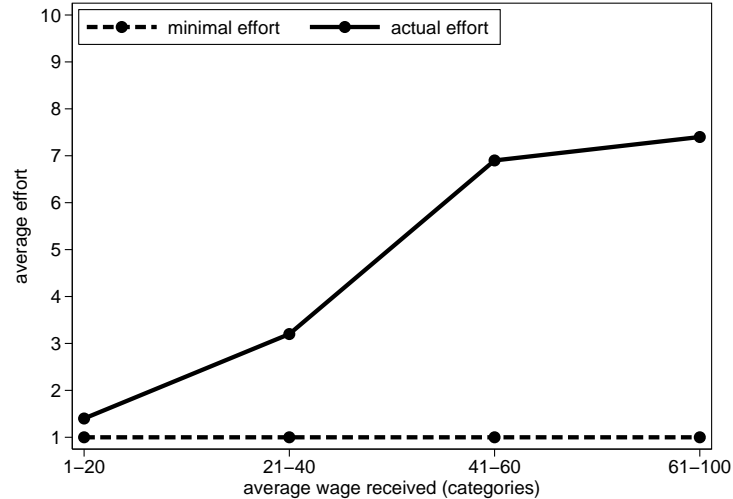


Figure 5: Gift exchange in an experimental labor market (source: Riedl and Tyran (2005))

4.1 On the interaction between trading institutions, morality, and tax-shifting.

An interesting and important example where economic institutions and inclinations toward reciprocity interact in a non-trivial way is the case of *tax liability side equivalence*. Tax liability side equivalence is a basic tenet in public economics. It states that the statutory incidence, that is who is legally responsible to pay a tax, is irrelevant for economic incidence, that is who actually bears the tax burden. In the words of one of the authorities in public economics, Richard A. Musgrave,

“it is a matter of indifference whether a general tax on transactions is assessed on the seller’s or on the buyer’s side of the market” (Musgrave, 1959, p. 351).

Importantly, under traditional economic assumptions this holds true independently of the trading environment (monopoly, oligopoly, competition, or bargaining) provided that prices can in principle adjust freely. Interestingly enough, however, much of the public debate about tax burden (and subsidy benefit) in the media and the political arena is concerned with statutory instead of economic incidence. This rises the question whether the public reasoning or the professional economic reasoning is incorrect. For public policy the answer to this question is obviously important because it determines which groups of the society are actually going to carry the burden of a tax, and, hence, what the distributional and allocational consequences of a tax are. Neglecting the behaviorally true tax burden and relying (only) on normative prescriptions that are based on incorrect behavioral assumptions may, therefore, have very undesirable political and economical consequences.

To uncover the behaviorally true tax burden, ideally one would like to shift the statutory tax burden from one side of the market, e.g. buyers, to the other side of the market, e.g. sellers, *leaving everything else equal*. Naturally, such situations do not occur regularly in the field. Fortunately, laboratory experiments are an ideal method to do precisely this. The following shortly reports on three sets of experiments testing tax liability side equivalence under three important economic institutions: competitive markets with complete contracts, bargaining, and gift exchange markets.

Three studies (Borck et al., 2002; Kachelmeier et al., 1994; Ruffle, 2005) experimentally investigate tax (and subsidy) incidence equivalence in competitive markets under various trading mechanisms. The results of these experiments can simply be summarized: tax liability side equivalence holds in competitive experimental markets independent of the trading institution. Traditional theoretical and behavioral tax incidence coincide. Kerschbamer and Kirchsteiger (2000) give an interesting twist to these results by studying tax liability side equivalence in a simple bargaining environment. They modify the ultimatum game (similar to the one described above) such that in case of acceptance of a proposal one of the negotiators has to pay a tax from her gross earnings. In one treatment the proposer has to pay the tax. In a second treatment the statutory tax burden is placed upon the responder. If tax liability side equivalence holds the distribution of earnings should be the same under both tax treatments. However, if statutory tax incidence implies also a moral obligation to actually bear part of the tax burden then tax liability side equivalence breaks down. It is well known that in bargaining situations reciprocity and distributional fairness crucially influence behavior and economic outcomes. Hence, one might expect that tax liability side equivalence is especially vulnerable in such situations. Indeed, Kerschbamer and Kirchsteiger (2000) find that the side of the transaction that is legally obliged to pay the tax also bears a disproportional part of the economic tax burden. The observed incompleteness of tax shifting is also economically significant. When the tax liability was taken from the responder and placed upon the proposer the offered net (after tax) income for responders *increased* between 20% and 24%. In contrast, standard tax liability side equivalence predicts no change in net income at all. Therefore, this is a clear case where trading institution and moral behavior interact in a way traditional theory fails to predict.

Perfectly competitive markets and pure (two person) bargaining situations are at the two extremes of actually existing economic exchange institutions. Although both are not uncommon they are probably not the most frequent trading institutions. Rather, a mixture of both seems to be the most common one. Gift exchange markets or markets with incomplete (or even no) contracts, as described above, incorporate both elements: competitive market interaction and bilateral bargaining. Riedl and Tyran (2005) investigate tax liability side equivalence in such markets. In one set of experiments buyers are legally obliged to pay a tax and in another set of experiments sellers have the legal obligation to

pay the tax. If tax liability side equivalence holds then there should be no difference in the outcomes of real variables between the two tax regimes. However, if moral obligations, as in the study of Kerschbamer and Kirchsteiger (2000), are important then the side on which the tax is levied should also bear a relatively larger economic burden of the tax. In gift exchange markets three important variables can be investigated: net prices, effort and net earnings. Tax liability side equivalence may fail on each of these dimensions with different consequences for the final distribution of income.

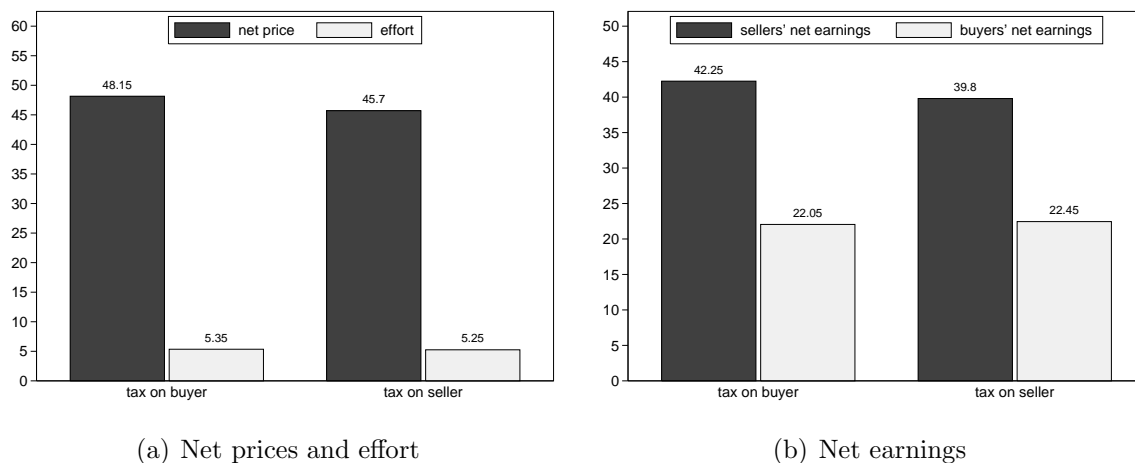


Figure 6: Prices, effort, and earnings under the two tax regimes (source: Riedl and Tyran (2005))

Figure 6 depicts average net prices and exerted efforts in panel (a) and average net earnings in panel (b) for both, taxes levied on buyers and taxes levied on sellers. From the figure it is obvious that there are no large differences between the two tax regimes on all three dimensions of comparison. Furthermore, the observed small differences are neither statistically nor economically significant. This is an important but also somehow puzzling result given the observation that tax liability side equivalence breaks down in pure bargaining and not when it is coupled with a competitive market. It strongly suggests that the details of market interaction and perceptions about moral obligations to pay a tax interact in a non-trivial way with real and economically significant results in terms of tax shifting.

One conclusion to be drawn from these studies is that researchers as well as policy makers should not solely rely on traditional economic theorizing when assessing the economic burden of a tax. Additionally, the institutional environment and its interaction with moral and reciprocal inclinations is crucially affecting the actual economic outcome. Where the precise borders for predictive accuracy of the standard economic model lie is still an open empirical question, though. If we want to understand under which circumstances tenets like tax liability side equivalence indeed hold or - more importantly - have to be adapted

we need a research program that systematically evaluates such pieces of economic wisdom.

4.2 Identifiable victims and hidden taxes.

Real outcomes can be non-trivially influenced not only by the interaction between institutional design of trading institutions and behavior of economic agents but also already through pure presentation effects. In a 1968 article Thomas Schelling noticed that “[t]he life you save may be your own” and that “the death of a particular person evokes anxiety and sentiment, guilt and awe, ... [but that] ... most of this awesomeness disappears when we deal with statistical death” (Schelling, 1968). This is probably the first account by an economist pointing towards how different we perceive identifiable and statistical victims. Loewenstein et al. (2006) take up this issue and examine the public policy consequences of this human inclination of having stronger feelings towards an identifiable victim than towards a statistical victim. They argue that from a welfare economics point of view “people may be insufficiently sympathetic towards statistical victims”. Mainly psychological research strongly supports the claim that individual concrete cases have a much more powerful motivational effect than statistical cases. This seems to be the case even if the statistics are objectively more informative than the individual case. A typical example in this respect is that opinions about the abuse of welfare payments is shaped much more strongly by individual experience than by objective statistics.¹⁴

An important public policy implication of the identifiability effect is that for politicians ‘hidden taxes’ tend to be much more popular than other taxes. For example, the value added (or sales) tax is for most consumers (including economists) simply part of the purchase price of a commodity, and, hence, has no identifiable victim. This concealment may make it politically easier to raise value added taxes rather than more direct taxes. How serious and economically important the misperception of hidden taxes is, is convincingly shown by Chetty et al. (2009). These authors conduct a field experiment in a grocery store where in one treatment the sales tax is made salient by explicitly showing it on the price tag whereas in another treatment only the tax inclusive price is shown. The study finds a significant 8 percent decrease in purchases and sales revenues when the tax is made salient. This salience effect is corroborated with field empirical data using variations on taxes on beer between different U.S. states. An important conclusion of this study is that it is crucial to distinguish between tax elasticities and price elasticities when thinking about tax policy; a distinction not necessary in traditional public economics.

Other prominent examples of hidden taxes are withholding income taxes which make people think that the money transferred to the tax authority is not their own and corporate income taxes which makes people believe that shareholders pay the tax although it is

¹⁴For clean evidence from the laboratory (dictator game giving) as well as the field (housing for needy), see Small and Loewenstein (2003).

mostly the factor labor that actually carries the tax burden (McCaffery, 1994; McCaffery and Baron, 2006). In the mentioned examples the lack of identifiability makes the taxes themselves as well as an eventual increase of them much more acceptable than it would be the case for non-hidden taxes.¹⁵

Experimental and behavioral research has uncovered a number of other behavioral regularities that are particularly important for taxation economics, like misperception of the progressivity of the tax system, confusion of marginal and average income tax rates or non-rationally high tax compliance rates (see, e.g., de Bartolome, 1995; Liebman and Zeckhauser, 2004; Slemrod, 2006). In his contribution to this issue Slemrod (2010) discusses some of these problems in more detail.¹⁶

4.3 Endogenous preferences and competition.

One central assumption in economics is that people have “fixed lifetime preferences” (Bernheim and Rangel, 2007). In particular, this means preferences are assumed not to change across states of nature or institutional constraints. In this perspective, preferences are exogenously fixed and independent from the environment an individual is immersed in. In contrast to this traditional view, Bowles (1998) argues vividly in favor of *endogenous* preferences. He claims that our preferences are not well defined and stable but rather strongly dependent on the environment we have to deal with. However, all the evidence he puts forward in support of his claims is either indirect or is open to alternative interpretations.

In a recent study, Brandts et al. (2008, 2009) experimentally test the influence of trading institutions on subjective well-being and (social) preferences directly. In a series of experiments subjects are divided into two groups. One group interacts in a competition free environment whereas subjects in the other group interact in a competitive environment. The hypothesis is that the experience of competition versus no competition *per se* leaves its traces in subjects’ well-being (in the sense of Kahneman et al. (1997)’s ‘experienced utility’) and in their social preferences (that is, their ‘social disposition towards others’). To test this hypothesis methods from social psychology are combined with experimental economics. At the beginning of the experiment subjects are asked to perform a social value orientation test, which measures subjects’ social preferences by letting them allocate real money between themselves and some anonymous other person. The same test is conducted also after subjects had experienced a competitive or a non-competitive environment. Hence, social preferences are measured before and after the experience with a

¹⁵The psychological appeal of hidden taxes is nicely summed up by the aphorism of Russell Long, one of the most powerful and influential tax legislators as chairman of the U.S. Senate Finance Committee “Don’t tax him, don’t tax me, tax the man behind the tree.” (quoted after Small and Loewenstein, 2003)

¹⁶For earlier accounts on taxation and behavioral economics, see McCaffery and Slemrod (2006) and Kirchler (2007).

particular trading institution. If traditional economic reasoning would be correct, then for the social orientation of participants it should not matter whether they have experienced a competitive or a non-competitive environment. This is not what is found. Firstly, there is a difference in the social orientation of subjects experiencing different institutions, and secondly, within the competitive institution the social orientation differs strongly between subjects that experienced it in different roles (i.e., being on the long or the short side of the trade relationship).

Figure 7 gives a visual impression of the differences in social preferences. It shows the change in social orientations from before to after the experience with the competitive and the non-competitive environment, respectively.¹⁷ Interestingly, under both trading

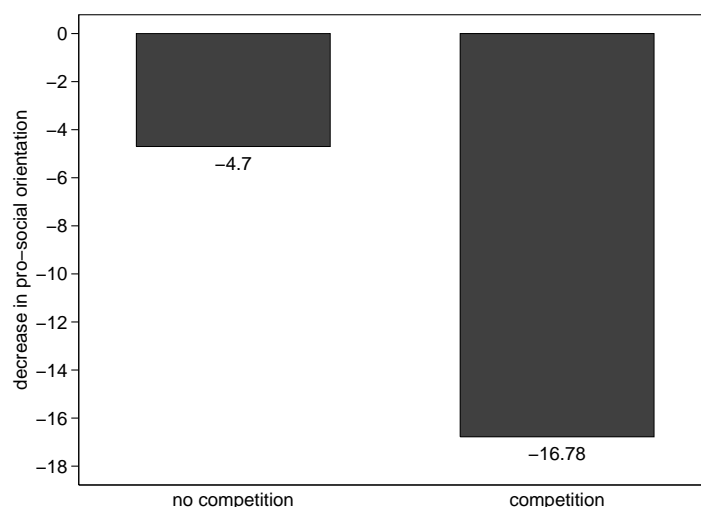


Figure 7: Change in social preferences in competitive and non-competitive environment (source: Brandts et al. (2009, 2008))

regimes pro-social orientation deteriorates. Importantly, however, compared to experience with the no competition regime the decrease in pro-social orientation is more than three times larger when subjects experience competition. This clearly indicates that preferences are indeed not - or at least only partly - exogenously given and strongly shaped by subjects' institutional experience. An additional result of this study is that not only social orientation deteriorates under competition but that subjects' also suffer a loss of experienced utility (i.e. subjective well-being). Similar findings are reported in a meta-study by Bowles and Polania Reyes (2009). The authors survey the data of 51 experiments and find that explicit economic incentives aimed at the narrowly materially selfish individual often not only have the intended effect of providing information and suggesting socially appropriate behavior but also the unintended and unwanted effect of comprising intrinsic

¹⁷For clarity of presentation only the average of the strongest change in social preferences among subgroups across all investigated competitive environments is shown.

motivation and self-determination. An important conclusion drawn from reviewing these studies is that economic incentives and social preferences seem more often complements than substitutes.

In public and political debates reforms aiming at more competition are often backed with the received wisdom of traditional economics that consumers will be better off in the presence of more competition. This argument refers to the notion of consumer surplus which (in practice) measures welfare in purely material terms. This surplus may indeed increase with more competition and less regulation (although in the experiment discussed above even this is not the case). However, the ultimate aim of welfare economics and public policy should be the maximization of the citizens' well-being. In this sense, the findings of deteriorating social orientations and suffering subjective well-being puts question marks behind the supposedly purely positive effects of more competition and less regulation. Unfortunately, there is no generally accepted measure of subjective well-being developed, yet, and much more research into the measurement and determinants of well-being is necessary.¹⁸

The evidence reported above also points to the important and not yet well understood interplay between heterogeneous social preferences and institution design and formation. There is mounting evidence that people differ quite substantially with respect to their social preferences (e.g. Andreoni and Miller, 2002; Engelmann and Strobel, 2004; Fisman et al., 2007; Bellemare et al., 2008), and there is also recent evidence that even subtle institutional differences may alter behavior substantially (see, e.g., Falk and Kosfeld, 2006; Reuben and Riedl, 2009). However, evidence how these interact is only very recently emerging (Kosfeld et al., 2009).

4.4 Presentation effects and public policy.

Finally, by closing the circle, let me give two examples where research into presentation and framing effects can very concretely inform public policy and thereby increase general well-being. The first one is taken from Amir et al. (2005) and highlights the importance of whether a decision task is presented simultaneously or sequentially. In many places police lineups are used to identify suspects of crime. In such a lineup eyewitnesses of crimes attempt to recognize one person in a group of suspects standing next to each other (Figure 8). Evidence from psychological research casts serious doubts on this much used method of identification. Lindsay and Wells (1985) designed a clever experiment and showed that the likelihood of false identification of an innocent subject is much higher under simultaneous than under sequential lineup. At the same time the frequencies of correct

¹⁸Recently, some progress is booked concerning the measurement of subjective well-being. Kahneman and Krueger (2006) introduce a measure based on self-reports of peoples' emotional states (see also Brandts et al., 2008).



Figure 8: Typical police lineup

recognition of the guilty suspect did not differ significantly between the two conditions. This, together with subsequent research, clearly shows that in identification tasks - as police lineups - the approach to present items or individuals one at a time leads to better, less biased, judgments than the practice of presenting several individuals simultaneously. The straightforward and unambiguous policy recommendation is, therefore, to abandon simultaneous lineups in favor of sequential lineups. Actually, at least two U.S. American states have adopted the alternative methodology.

The organization of organ donation is another important domain where it has been shown that framing effects can crucially alter behavior and, in this case, make the difference between life and death. Johnson and Goldstein (2003) investigate and compare two different organ donation systems around the world. They distinguish between two so-called no-action defaults. No-action defaults are the choices implicitly imposed on individuals who do not take an active decision. In the case of organ donation the most widely used default decisions are "presumed-consent" and "explicit-consent". Presumed-consent means that people are assumed to be organ donors as long as they do not actively indicate otherwise. Explicit-consent means that individuals have to actively register for being a donor, otherwise they are not. The authors investigate the effective consent rates for being a donor across the two defaults experimentally as well as by cross-country comparisons. Traditional economic theory assumes that preferences are fixed and known to the economic agent, which implies that for effective consent rates it should not matter whether one or the other default option is taken. Figure 9 shows the facts. The four leftmost bars depict effective consent rates for four countries (Denmark, the Netherlands, UK, and Germany, respectively) which apply the explicit-consent as default. The bars to the right show effective consent rates for countries with presumed-consent as default. The

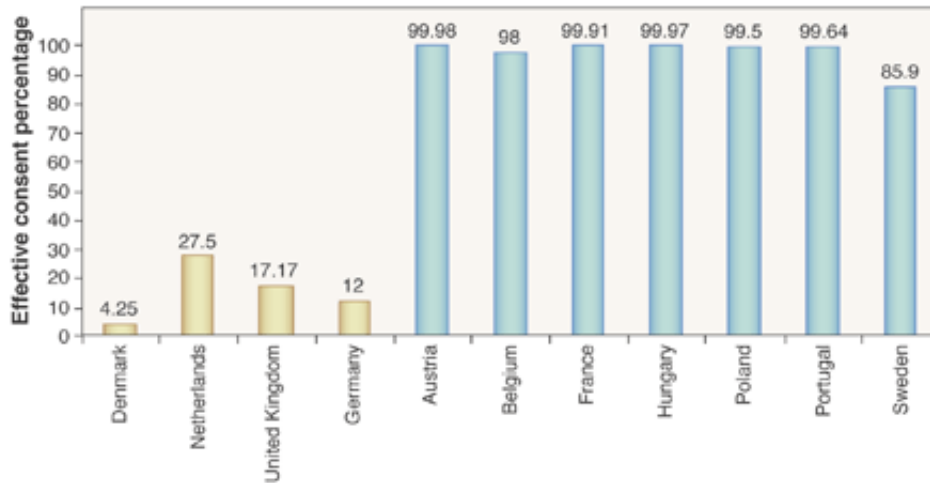


Figure 9: Effective consent rates in countries with explicit consent (four leftmost bars) and presumed consent (seven rightmost bars) (source: Science 302 2006 p.1338)

differences are striking. Whereas in the explicit consent countries the effective consent rates are maximally 27.5 percent (the Netherlands), the minimum consent rate in countries with presumed consent is 85.9 percent (Sweden). These differences are surely too large to be explained by effort or transaction costs of actively opting in and opting out in the explicit consent and presumed consent countries, respectively. To exclude also this potential explanation the authors conducted an experiment where effort and transaction costs were virtually zero. In the experiment the difference in effective consent rates between the two default options is slightly smaller than in the cross-country study but still a long way off from being equal. Interestingly, a neutral framing without any default led to the same effective consent rate than the presumed consent default. This allows the conclusion that the unbiased ‘true’ preferences concerning organ donation are better elicited with presumed consent than with explicit consent. That some policy makers are - at least sometimes - aware of the fact that default options make a difference is nicely reflected in arguments brought forward during a recent discussion about reforming the donor registration system in the Netherlands. In 2005 a coalition of parties in the parliament actually did bring forward a motion that would have changed the explicit consent default into a presumed consent default. At the end, however, there was not enough political support for such a radical change of the donor registration system.¹⁹ In any event, this example highlights how insights from behavioral and experimental research are related to important

¹⁹A clear case of a presentation and framing effect is also reported in an article of the Dutch newspaper *NRC Handelsblad* (September 2 & 3, 2006, pp. 41-42) about the Dutch immigration and naturalization service (IND). The newspaper reports that if employees of the IND reject an application for a ‘temporary residence permit’ they have to motivate their decision in written form. For the case of ‘hardships of asylum seekers’ the IND employees have to motivate the decision if they accept the application. The result being that cases of hardship are hardly ever positively assessed whereas temporary residence permits are relatively easily issued.

political decisions and can inform public policy more accurately than traditional economic reasoning would be able to.

5 Further directions: the field and the brain

A common argument of skeptics against the use of laboratory experiments, in general, and as a policy advice instrument, in particular, is its supposed lack of external validity. This is indeed an important concern because if regularities observed in the laboratory do not carry over to the field any conclusions and public policy advice drawn from these experiments could be dangerously misleading. The potential problem of lack of external validity is not unique to economic (or psychological) experiments, however.²⁰ In physics, the feather and the stone which fall with the same speed in vacuum but with different speeds in ‘real life’ is a well-known illustrative example. As the air resistance in the terrestrial atmosphere affects the fall velocity of the two objects, in economic situations many factors one can control for in the laboratory but not in the real world can influence behavior and blur or even wipe out behavioral regularities observed in the laboratory. One way to tackle this problem, which is unique to the experimental method, is to add pieces of real life context to the ‘dry’ laboratory environment *in a systematic way*. In this way one can trace if and how such ‘pieces of reality’ alter behavior.²¹

Another important way to check for external validity is to replace the usual student subjects with subjects who are experienced with the decision situation at hand and/or are more representative than students. Such experiments have been conducted for a variety of decision situations (see, e.g., Fehr and List, 2004; Egas and Riedl, 2008, among many others). The general upshot from these experiments is that ‘experts’ often do not make significantly different decisions than students in the same situation, although there are sometimes subtle and surprising differences. For instance, Alevy et al. (2007) investigate the behavior of financial market professionals regarding information cascades and find that “professionals are less Bayesian than students” (ibid., p. 161) but report only little evidence for differences in cascade formation. Haigh and List (2005) investigate the difference in myopic loss aversion (MLA) between students and professional traders and find that “traders exhibit behavior consistent with MLA to a *greater* extent than students” (ibid., p. 523, emphasis in original). Hence, sometimes behavioral regularities found with students are even amplified with non-student subjects.

A third way to test external validity is to conduct ‘field experiments’. Field experiments might be seen as an extreme combination of the two already described ways of adding

²⁰Note that theoretical reasoning is confronted with exactly the same potential lack of external validity.

²¹For a recent discussion of this and other pro and cons regarding the use of laboratory experiments in economics, see Falk and Heckman (2009).

‘realism’ to experiments. In field experiments researchers try to add the largest possible amount of ‘pieces of reality’ and use non-student subjects in their ‘natural’ environment. Harrison and List (2004) provide a typology of field experiments and define them, crudely speaking, as experiments where one is “recruiting subjects in the field rather than in the classroom, using field goods rather than induced valuations, and using field context rather than abstract terminology in instructions” (Harrison and List, 2004, p. 1009-1010).

Without doubt, field experiments are an interesting and important development in economic research, but they also have their disadvantages. For instance, in comparison to the use of induced valuations, the use of field goods contributes to a loss of control regarding participants’ true valuations of commodities, the use of special non-student subjects in their special circumstances does not necessarily allow to draw conclusions beyond the investigated group in the investigated circumstances, and field experiments do usually not allow for replications as precise as in the laboratory. Therefore, especially in research that is concerned with policy advice, field experiments are best viewed as an important complementary research method.

In the ideal case, an economic policy reform is evaluated with *all possible* scientific methods before a political decision is made. That is, theoretically, experimentally in the lab- and the field, and with traditional applied econometrics. It should be obvious that a thorough scientifically sound examination of a policy reform that reduces the risk of implementing false policies is much cheaper (at least in expectations) than the costs of an actually implemented false policy. A rare example of a first - albeit incomplete - attempt of such a scientific approach to policy issues is the evaluation of the so-called “Plan Van Elswijk” in the Netherlands, which proposes a radical reform of the (Dutch) financing system of unemployment benefits. In the evaluation of this plan simulation studies, laboratory experiments and a small field experiment were conducted (see van Winden et al., 1999, 2000; Riedl and Winden van, 2001, 2007, 2008, and the references therein). These studies produced a rather clear picture of the likely benefits and disadvantages of the proposed reform.²² An important lesson learned from these studies is that laboratory experiments can indeed provide valuable information even for such complex questions as unemployment benefit reform plans. Thus, one may safely conclude that laboratory experiments together with field experiments and theoretical reasoning will play an important role in public policy advice in the future.²³

A further - quite different - but potentially equally relevant and in the long run probably even more important - but also more debated - stream of research is the recent

²²At the end, despite the clear-cut results policy makers chose to interpret the results differently than most involved researchers did and, to a large extent, ignored the scientifically achieved outcomes.

²³An interesting recent example where laboratory experiments informed politics is given in Jacob K. et al. (2005). For a survey of experiments for economic policy in the context of industry regulation, see Normann (2004) and Hinloopen and Normann (2009).

combination of neuroscience with experimental and behavioral economics. This new research branch - *neuroeconomics* - uses knowledge about brain mechanisms to study the biological foundations of behavioral regularities observed in the laboratory and the field. Kevin McCabe, one of the pioneers in this new field of the behavioral sciences defines it as follows:

“Neuroeconomics is an interdisciplinary research program with the goal of building a biological model of decision making in economic environments. [More specifically, it] is the study of how the embodied brain interacts with its external environment to produce economic behavior. Research in this field will allow social scientists to better understand individuals’ decision making, and consequently to better predict behavior.” (McCabe, 2003, p. 294)

At first sight this may sound abstract and relatively remote from public policy issues. Indeed, in the above definition, the first part refers to the ‘pure scientific’ element of neuroeconomics. The second part, however, reflects the potential of this approach for public policy making. Having *good models* of human behavior is crucial for making *good predictions* of human behavior in economic situations. Neuroeconomics has the potential to significantly contribute to this quest for better models of economic decision making. When evaluating the potential of this new emerging field one may want to recall the history of game theory and laboratory experiments in economics. In its beginnings game theory was largely dismissed as being too academic and thought to be useful at most for war strategists but surely not for the analysis of ordinary economic interactions. Nowadays game theory is used for policy advice at a large scale in many countries. Recall, e.g., the commercial frequency auctions a few years ago where game theorists played a crucial role in designing auction formats.) Later, experimental and behavioral economics was smiled at by many economists because it was the received wisdom in economics that “economists ... cannot perform ... controlled experiments” (Samuelson and Nordhaus, 1985) a view that changed quite a bit over the years as the following quote by the very same authors seven years later testifies. There they state that experimental economics is an “exciting new development.” (Samuelson and Nordhaus, 1992). It is not unlikely that neuroeconomics awaits the same ‘fate’. Bernheim (2009) and Rustichini (2009) are recent critical appraisals of this new development in economics.

In any event, both directions - towards the experimental field and towards the brain - can be expected to be one of the most lively areas of research in economics, and both are likely to produce results that will lead to better models, better predictions, better advice, and ultimately - hopefully - also to better informed public policies.

6 Conclusion

In the course of this contribution I discussed a few areas in experimental and behavioral economics that are of importance for public economics and public policy. Naturally, many at least equally important issues were not even slightly touched upon. For instance, voluntary contributions to public goods (see, e.g., Gächter and Herrmann, 2009), time inconsistent intertemporal decision making, i.e., retirement decisions and pension systems (see, e.g., Frederik et al., 2002), decision making under risk and uncertainty, i.e., health insurance and social security in general (see, e.g., Kahneman and Tversky, 1979; Wakker et al., 2007; Heinemann et al., 2008), or the role of moral property rights in policy reform (see, e.g., Gächter and Riedl, 2005) only to name a few. In the future, also the fundamental issue of ‘welfare analysis beyond revealed preferences’ is awaiting public economics research, and more generally, economic theory. First important steps in this direction are made by Bernheim and Rangel (2009) and by Herings and Rohde (2006).

In recent years, the ‘facts’ of behavioral regularities have been shown to be incompatible with the traditional ‘wisdom’ of economic theory. This has led to the development of new theoretical approaches and models. It seems clear that for good public policy we need both good accounts of the behavioral facts and a theoretical knowledge that gives us the tools to deal with the upcoming challenges in an accurate way. This is all the more important as politicians are increasingly more willing to listen to the advice of behavioral and experimental economists. In fact, the 2005 quote that opens this paper is overtaken by reality, as the *Time Magazine* reports that behavioral economists are advising U.S. president Barack Obama (Grunwald, 2009).

References

- Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics*, 97(4):543–569.
- Alevy, J. E., Haigh, M. S., and List, J. A. (2007). Information cascades: Evidence from a field experiment with financial market professionals. *Journal of Finance*, LXII(1):151–180.
- Amir, O., Ariely, D., Cooke, A., Dunning, D., Epley, N., Gneezy, U., Koszegi, B., Lichtenstein, D., Mazar, N., Mullainathan, S., Prelec, D., Shafir, E., and Silva, J. (2005). Psychology, behavioral economics, and public policy. *Marketing Letters*, 16(3/4):443–454.
- Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.
- Bellemare, C., Kröger, S., and van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, 76(4):815–839.
- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10:122–145.
- Bernheim, B. D. (2009). On the potential of neuroeconomics: A critical (but hopeful) appraisal. *American Economic Journal: Microeconomics*, 1(2):1–41.
- Bernheim, B. D. and Rangel, A. (2007). Behavioral public economics: Welfare and policy analysis with nonstandard decision-makers. In Diamond, P. and Vartiainen, H., editors, *Behavioral Economics and Its Applications*, pages 7–84. Princeton University Press, Princeton.
- Bernheim, B. D. and Rangel, A. (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1):51–104.
- Borck, R., Engelmann, D., Müller, W., and Normann, H.-T. (2002). Tax liability side equivalence in experimental posted offer markets. *Southern Economic Journal*, 68:672–682.
- Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature*, 36:75–111.

- Bowles, S. and Polania Reyes, S. (2009). Economic incentives and social preferences: A preference-based lucas critique of public policy. working paper 2734, CESifo.
- Brandts, J., Riedl, A., and Winden, F. v. (2008). On competition and well-being. an experimental investiagtion into rivalry, social disposition, and subjective well-being. Maastricht University.
- Brandts, J., Riedl, A., and Winden, F. v. (2009). Competitive rivalry, social disposition, and subjective well-being: An experiment. *Journal of Public Economics*, 93(11/12):1158–1167.
- Camerer, C. and Loewenstein, G. (2004). Behavioral economics: past, present, future. In Camerer, C., Loewenstein, G., and Matthew, R., editors, *Advances in behavioral economics*, pages 3–51. Princeton University Press, Princeton and Oxford.
- Camerer, C. F. (2005). Behavioral economics. Lecturer at World Congress of the Econometric Society, 2005.
- Camerer, C. F. and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19:7–42.
- Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177.
- Cronqvist, H. and Thaler, R. H. (2004). Design choices in privatized social-security systems: Learning from the Swedish experience. *American Economic Review*, 94(2):424–428.
- de Bartolome, C. A. M. (1995). Which tax rate do people use: Average or marginal? *Journal of public Economics*, 56(1):79–96.
- de Jong, J. D., van den Brink-Muinen, A., and Groenewegen, P. P. (2008). The Dutch health insurance reform: switching between insurers, a comparison between the general population and the chronically ill and disabled. *BMC Health Services Research*, 8(58):1–9.
- Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society - B*, 275(1637):871–878.
- Engelmann, D. and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review*, 94(4):857–869.
- Falk, A., Fehr, E., and Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26.

- Falk, A. and Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326:535–538.
- Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5):1611–1630.
- Fehr, E., Gächter, S., and Kirchsteiger, G. (1997). Reciprocity as contract enforcement device: Experimental evidence. *Econometrica*, 65:833–860.
- Fehr, E., Goette, F., and Zehnder, C. (2009). A behavioral account of the labor market: the role of fairness concerns. *Annual Review of Economics*, 1:355–384.
- Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does fairness prevent market clearing? an experimental investigation. *Quarterly Journal of Economics*, 108:437–459.
- Fehr, E., Kirchsteiger, G., and Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42:1–459.
- Fehr, E. and List, J. A. (2004). The hidden costs and returns of incentives - Trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5):743–771.
- Feynman, R. O. (1964). *The Feynman Lectures on Physics*. Addison-Wesley, Reading, MA.
- Fisman, R., Kariv, S., and Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, 97(5):1858–1876.
- Forsythe, R., Horowitz, J. L., Savin, N., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6:347–369.
- Frederik, S., Loewenstein, G., and O’Donogue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, XL:351–401.
- Gächter, S. and Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1514):791–806.
- Gächter, S. and Riedl, A. (2005). Moral property rights in bargaining with infeasible claims. *Management Science*, 51(2):249–263.
- Grunwald, M. (2009). How Obama is using the science of change. *Time Magazine*, April 2, 2009.

- Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4):367–388.
- Haigh, M. S. and List, J. A. (2005). Do professional traders exhibit myopic loss aversion? an experimental analysis. *Journal of Finance*, LX(1):523–534.
- Harrison, G. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, XLII:1013–1059.
- Heinemann, F., Förg, M., Jonas, E., and Traut-Mattausch, E. (2008). Psychologische restriktionen wirtschaftspolitischer reformen. *Perspektiven der Wirtschaftspolitik*, 9(4):383–404.
- Herings, P. J.-J. and Rohde, K. I. M. (2006). Time-inconsistent preferences in a general equilibrium model. *Economic Theory*, 29:591–619.
- Hinloopen, J. and Normann, H.-T., editors (2009). *Experiments and Competition Policy*. Cambridge University Press, Cambridge.
- Jacob K., G., Offerman, T. J. S., and Schram, A. J. H. C. (2005). Using first-price auctions to sell heterogeneous licenses. Working Paper, University of Amsterdam.
- Jamison, J., Karlan, D., and Schlechter, L. (2008). To deceive or not deceive: the effect of deception on behavior in future laboratory experiments. *Journal of Economic Behavior and Organization*, 68:477–488.
- Johnson, E. J. and Goldstein, D. (2003). Do defaults save lives? *Science*, 302:1338–1339.
- Kachelmeier, S. J., Limberg, S. T., and Schadewald, M. S. (1994). Experimental evidence of market reactions to new consumption taxes. *Contemporary Accounting Research*, 10:505–545.
- Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1):3–24.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.
- Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, 112(2):375–405.
- Kerschbamer, R. and Kirchsteiger, G. (2000). Theoretically robust but empirically invalid? an experimental investigation into tax equivalence. *Economic Theory*, 16:719–734.

- Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge University Press, Cambridge.
- Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, 99(4):1335–1355.
- Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. Technical report, Harvard University. unpublished manuscript.
- Lindsay, R. C. L. and Wells, G. L. (1985). Improving eyewitness identifications from line-ups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, 70(3):556–564.
- Loewenstein, G. (1999). A visceral account of addiction. In Elster, J. and J, S. O., editors, *Getting hooked: Rationality and addiction*, pages 235–264. Cambridge University Press, Cambridge, England.
- Loewenstein, G., Small, D. A., and Strnad, J. (2006). Statistical, identifiable, and iconic victims. In McCaffery, E. J. and Slemrod, J., editors, *Behavioral Public Finance*, pages 32–46. Russell Sage Foundation Press, New York.
- McCabe, K. (2003). Neuroeconomics. In Nadel, L., editor, *Encyclopedia of Cognitive Science*, volume 3, pages 294–298. Nature Publishing Group, Macmillan Publishers Ltd.
- McCaffery, E. J. (1994). Cognitive theory and tax. *UCLA Law Review*.
- McCaffery, E. J. and Baron, J. (2006). Thinking about tax. *Psychology, Public Policy, and Law*, 12(1):106–135.
- McCaffery, E. J. and Slemrod, J., editors (2006). *Behavioral public finance*. Russel Sage Foundation, New York.
- Musgrave, R. A. (1959). *The Theory of Public Finance*. McGraw-Hill, New York.
- NIVEL (2009). Nivel jaarboek 2008.
- Normann, H.-T. (2004). Experiments for economic policy making. Technical report, ENCORE, University of Amsterdam.
- Ortmann, A. and Hertwig, R. (2002). The costs of deception: evidence from psychology. *Experimental Economics*, 5:111–131.
- Plott, C. R. (1982). Industrial organization theory and experimental economics. *Journal of Economic Literature*, 20:1485–1527.

- Read, D., Loewenstein, G., and Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, 19:171–197.
- Reuben, E. and Riedl, A. (2009). Enforcement of contribution norms in public good games with heterogeneous populations. discussion paper, Maastricht University.
- Riedl, A. and Tyran, J.-R. (2005). Tax liability side equivalence in gift-exchange labor markets. *Journal of Public Economics*, 89:2369–2382.
- Riedl, A. and Winden van, F. (2001). Does the wage tax system cause budget deficits? *Public Choice*, 109:371–394.
- Riedl, A. and Winden van, F. (2007). An experimental investigation of wage taxation and unemployment in closed and open economies. *European Economic Review*, 51(4):871–900.
- Riedl, A. and Winden van, F. (2008). Input versus output taxation in an experimental international economy. unpublished manuscript, Department of Economics, Maastricht University.
- Ruffle, B. J. (2005). Tax and subsidy incidence equivalence theories: Experimental evidence from competitive markets. *Journal of Public Economics*, 89(8):1519–1542.
- Rustichini, A. (2009). Is there a method of neuroeconomics? *American Economic Journal: Microeconomics*, 1(2):48–59.
- Samuelson, P. A. and Nordhaus, W. D. (1985). *Economics*. McGraw-Hill, New York, 12 edition.
- Samuelson, P. A. and Nordhaus, W. D. (1992). *Economics*. McGraw-Hill, New York, 15 edition.
- Schelling, T. C. (1968). The life you save may be your own. In Chase, S. B., editor, *Problems in Public Expenditure Analysis*. The Brookings Institute, Washington DC.
- Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, 27:150–162.
- Simonson, I. and Winer, R. S. (1992). The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research*, 19:133–138.
- Slemrod, J. (2006). The role of misconceptions in support for regressive tax reform. *National Tax Journal*, 59(1):57–75.

- Slemrod, J. (2010). Old George Orwell got it backward: some thoughts on behavioral tax economics. *FinanzArchiv/Public Finance Analysis*, 66:forthcoming.
- Small, D. A. and Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, 26(1):5–16.
- Stigler, G. J. (1980). Economics or ethics? Tanner Lectures on Human Values.
- Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: improving decisions about health, wealth, and happiness*. Yale University Press, New Haven and London.
- Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458.
- van Winden, F. A. A. M., Riedl, A., Wit, J., and van Dijk, F. (1999). Experimenteel economisch onderzoek naar het Plan Van Elswijk. Eindrapport, CREED, Universiteit van Amsterdam.
- van Winden, F. A. A. M., Riedl, A., Wit, J., and van Dijk, F. (2000). Experiment: het Plan Van Elswijk. *Economisch Statistische Berichten*, 10-3-2000:197–199.
- Wakker, P. P., Timmermans, D. R. M., and Machielse, I. (2007). The effects of statistical information on risk and ambiguity attitudes, and on rational insurance decisions. *Management Science*, 53(11):1770–1784.