Economic Ethics for Real Humans*

The Contribution of Behavioral Economics to Economic Ethics

Lisa Herzog**

Wirtschaftsethik für menschliche Wesen – Der Beitrag der Verhaltensökonomie zur Wirtschaftsethik

This paper discusses how economic ethics can profit from taking into account the results of behavioral economics. In contrast to the neo-classical mainstream of economics, behavioral economics does not presuppose the model of 'economic man', but explores the ways in which real human beings make economic decisions. The example of akrasia and its effects on old-age saving shows that behavioral economic research opens new fields for economic ethics. A central ethical aspect in this context is the question about the moral autonomy of economic agents. A Rawlsian approach shows that opt-out systems, which take into account typical behavioral tendencies, can, under certain conditions, be a way of combining the desiderata of supporting rational behavior and safe-guarding autonomy.

Keywords: Economic Ethics, Behavioral Economics, Economic Methodology, Akrasia, Libertarian Paternalism

1. Introduction

In Pushkin's novel *The Captain's Daughter* young Pjotr Andrejitsch starts his own life, after leaving his parents' estate, by losing hundred rubles when drinking and playing cards with a stranger. His old valet Saweljitsch is horrified: how could the young gentleman fool away his money like that? But Pjotr Andrejitsch insists on being his own master and pays his debts, although secretly repenting his carelessness (Pushkin 1992).

It can be painful to see others behave irrationally. Why do people not always act on the basis of good reasons, why do they violate their own principles? Moral philosophers have pondered upon this question since ancient times. Aristotle identified *akrasia*, 'lack of mastery', as culprit. While Socrates held that the one who knows what is right will always act accordingly, Aristotle skeptically remarked that this "contradicts things that appear manifestly" (Aristotle 1985: 1145b). A person afflicted with *akrasia* goes against reason due to some *pathos* ('emotion', 'feeling'). This weakness of the will, affecting behavior towards others as well as behavior towards oneself, leads to actions which the actor condemns beforehand and regrets afterwards.

---

* The article has been subject to a double blind peer review process. Date of submission: 6th September, 2007; revisited version accepted for publication: 14th June, 2008.

** Dipl.-Vw. Lisa Herzog, New College, Holywellstreet, UK-OX13BN Oxford, Great Britain, phone: +44-(0)1865-280000-x21122, e-mail: lisa.herzog@new.ox.ac.uk, fields of expertise: behavioral economics, economic ethics, political philosophy.

This paper deals with a special form of *akrasia* concerning oneself: the problem of abiding by (long-term) plans. In a world of growing complexity planning is unavoidable in order to allocate scarce resources like time or money. But a plan is only as good as the ability to stick to it – just think about New Year's pledges. What should be done with those who deviate from their own principles, whose behavior is suboptimal from their own long-term point of view? Taking into account areas such as health and saving behavior, this question is relevant not only to individual ethics, but also to political philosophy, and in particular to economic ethics.

In the past years economics, the dismal science, has started to occupy itself with phenomena such as *akrasia* – 'time-inconsistence preferences' in technical terms. This has taken place in the context of behavioral economics, an approach which emphasizes the "significance for economic analysis of the realism of one's modeling assumptions in terms of their behavioral and institutional dimensions" (Altman 2006: xvi). In particular, departures from the assumption of perfect rationality of traditional *economic man* are taken into account and examined systematically.

In this article I argue that economic ethics should take into account the research by behavioral economists, as it is of high relevance for its topics and its methodology. Section II contrasts the methodological principles of neoclassical economics and behavioral economics. In section III I discuss *akrasia* as an example of how behavioral economics brings into view topics which have been deliberately ignored by neoclassical economics, but which play a role from an ethical point of view. This carries chances as well as risks, a central aspect being human autonomy: How can one make sure that well-meant measures designed to "help consumers help themselves" (Lynch/ Wood 2006: 1) do not infringe personal liberty? Section IV proposes a pragmatic approach to the problem of 'models of man' in economic ethics and concludes.

2. Perfect models and imperfect humans – different approaches to 'the economy'

Conceived very broadly, economic ethics[1] asks how the economic life of our societies should be organized in order to realize moral values like a good life or human autonomy. A central part of economic ethics is the level of institutional ethics ('Ordnungsebene'): reflections about the way in which institutions (e.g. courts, the market, or government institutions) should be designed to achieve morally desirable results. Institutions can be understood as "collectively accepted system[s] of rules (procedures, practices)" which create "deontic powers" like rights, duties, obligations, etc. (Searle 2005: 21, 10). The importance of institutions results from the fact that, due to temporal and informational constraints, it is impossible to discuss the (ethical) implications of every single act; institutions offer standardized rules which can be applied to numerous cases. Institutions permit and shape human interaction (North 1990: 3) and allow

[1] I understand business ethics as part of economic ethics, focusing on considerations concerning the position and the behavior of companies. As I am dealing with a methodological issue, I do not mention business ethics explicitly; however, some of the points I make have implications for, or can be applied to, business ethics as well.

cooperation (Homann/Suchanek 2005: 36f.) – or they can in fact prevent it, if competitive behavior is desired for attaining welfare gains. By influencing behavior, institutions can also shape preferences, thus having substantive cultural consequences (see Bowles 1998 for an overview).

One might assume that economic ethicists, reflecting on institutions for the economic sphere, could ask economists about the 'laws' which govern the economy, in order to devise reasonable, enforceable institutions. However, the mainstream of economic theory in the last decades, neoclassical microeconomic theory, is not so much characterized by a topic ('the economy'), but by a specific approach: it analyses the behavior of *economic men*, rational individuals that maximize their utility under given constraints (cf. Hausman 1992: ch. 1). The question of how economic ethics should treat standard economic theory has been a point of fierce debate in recent years. E.g., while Karl Homann's school of economic ethics adopts *economic man* as heuristics, but opens up the notion of utility to include whatever individuals conceive as advantage, Peter Ulrich accuses neoclassical theory of being value-laden and defends the primacy of ethics over economics (Homann/Suchanek 2005: 26-29; Ulrich 1997: 12f., 102ff.; 2002: 2f.).

Neoclassical theory, however, is – by definition, as it were – blind to certain problems which would not occur if real human beings behaved like *economic man*. Behavioral economics, in contrast, searches for ways in which real humans depart from the picture neoclassical theory draws. Based on empirical research, it tries to pin down such 'anomalies' in a systematic way. To illustrate the difference it is helpful to consider how both approaches deal with *akrasia*.

2.1 The smooth world of standard microeconomic theory

In neoclassical economic theory one is by assumption in the best of all possible worlds concerning instrumental rationality: *economic man* always reasons correctly and acts perfectly rationally. Technically this is expressed by saying that the individual's preferences are stable, complete and transitive. Those preferences can be expressed in a utility function, and *economic man* will always act in a way that maximizes his utility function under given constraints, for instance those of a financial budget. Economists do not care very much about what enters the utility function. As L. Robbins famously said, "[e]conomics is not concerned with ends as such. (…) it asks how their [human beings] progress towards their objectives is conditioned by the scarcity of means" (Robbins 1932: 24). So from a purely formal point of view, the neoclassical approach is very wide, because all kind of things could be incorporated into the utility function of the individuals; the paradigm of 'rational choice' can thus also be used in other social sciences. As a matter of fact, however, the utility functions of most models in the mainstream of economic theory include only the agent's own monetary payoffs, and sometimes working time, or they describe relative evaluations between different goods. Given the preferences of individuals in a defined setting, economists can calculate the equilibrium towards which the system tends. They can search for frameworks, for instance tax systems, in which no resources are wasted and the utility functions of

the individuals concerned are maximized.[2]

Economic ethicists can resort to neoclassic economic theory if they come to the conclusion that the forces of the market, i.e. individual utility maximization, should be implemented at some point in the social system and that institutions should be designed accordingly. It has to be borne in mind, however, that this strategy yields the implicit assumption that people really maximize their 'neoclassical' utility in the given context.[3]

For economic ethicists the usage of neoclassical economics has important methodological implications. First and foremost, the neoclassical approach seems to be well-fitting with a fundamental principle of liberalism: the personal aims of every human being, i.e. the aims incorporated in the utility function, belong to the sphere of personal autonomy and are treated as given. Neither governmental authorities nor 'experts' of any kind are allowed to intrude into this sphere – this is expressed as 'consumer sovereignty' in neoclassical economics. Rather, the government should create a just and efficient institutional framework in which individuals can pursue their ends in the best way possible.

But on the other hand this means that a lot of potentially relevant questions are suppressed, as the intrapsychic structure of the individual's motivations and decisions is reduced to the mathematical calculus of utility optimizations. In particular, the question whether people might eventually do things that are individually irrational – ways in which they might harm themselves as seen from their own long-term perspective – cannot be asked.[4] This can be seen in the case of intertemporal decision making. The neoclassical approach uses the theory of 'discounted utility' (for an overview see Nyhus/Webley 2006: 297ff.): individuals consider future payoffs in their utility function, but give them less weight the further away they are. In particular, it is assumed that the value of future flows of money is discounted by individuals at a constant rate for each future period.[5] This implies that the relative importance of two subsequent years is the same, no matter how far away they are from the present. Under this assumption utility

---

2 I am aware, of course, that this rough-and-ready sketch of neoclassical theory is inadequate to capture the breadth of this school. It serves chiefly to accentuate the contrast to behavioral economics. For a detailed discussion of the merits and limitations of neoclassical theory see Hausman (1992).

3 This can be justified, for example, by arguing that other behavior can often be exploited in dilemmatic structures. This is the thrust of Homann's extensive use of *economic man* as heuristics for the analysis of social structures (e.g. Homann/Suchanek 2005: 23ff.). Homann/Meyer (2005: 96) conceive of *economic man* as a category systematically connected to dilemmas such as the prisoners' dilemma.

4 In contrast, the problems of social structures in which individually rational behavior leads to socially suboptimal results, such as the prisoners' dilemma etc., are widely dealt with, namely in game theory.

5 This assumption is justified by the stipulated existence of perfect capital markets at which agents can borrow or invest money at a constant rate of interest. The market rate is used for discounting: an amount of money to be received next year is less valuable than the same amount received now, because having it now means that one can invest it in the capital market and receive interest (Nyhus/Webley 2006: 299).

maximizing individuals will stick to investment or saving plans they make in the starting period, because it is optimal to do so from the point of view of every period (Strotz 1956: 172). Because of this assumption problems of self-discipline cannot arise; *akrasia* is no topic for standard economics.

Of course, (most) mainstream economists are aware that their assumptions about human behavior are not overly realistic. Different rationales for the use of *economic man* have been brought forward (see Schlicht 2003; Morgan 2006). One of the most influential defenses has been offered by Friedman (1953): the realism of the assumptions does not matter as long as a theory yields correct predictions. Interpreted as heuristic, *economic man* cannot be falsified by pointing to contradicting human behavior. However, it can be questioned whether it is a useful heuristic; and in particular, whether it should be the only heuristic used in analyzing economic behavior.

2.2 Experiment above paradigm – the approach of behavioral economics

The charge made by behavioral economists against neo-classical economics is that there are numerous phenomena that cannot be explained by the standard approach. By the help of experiments and empirical observations it has been shown that human behavior often differs systematically from what standard economic theory predicts.[6] In contrast to neoclassical theory, behavioral economists take seriously that "[f]actors such as social norms, morals, perceptions of justice, various attitudes, and particular beliefs can influence the way people behave, even sometimes if their behavior is not in their own immediate self-interest" (James 2006: 598).

One of the phenomena analyzed by behavioral economics is *akrasia*, under the technical terms 'time inconsistency' or 'hyperbolic discounting'.[7] Empirical research has found that the model of discounted utility delivers a poor description of reality (see Frederick et al. 2002; Nyhus/Webley 2006). Many people do not use a constant rate of discount when evaluating temporal sequences; in particular, they overvalue the presence in relation to the future ('present-biased preferences'). This implies that the evaluation of any plan depends on the point of time at which an agent finds himself. In consequence, options are evaluated differently depending on how close or far off they are in time; behavior, accordingly, becomes 'dynamically inconsistent' (Strotz 1956).[8]

---

6 For an overview see Camerer/Loewenstein 2003, for recent developments see Rabin 2002 and Altman's 2006 excellent handbook.

7 Due to a mathematical specification of the discounting function, see below footnote 8.

8 A common mathematical model for such behavior is a temporal sequence of 'selves', each with a discounting function geared to its period. The discounting function is hyperbolic, which amounts to a strong emphasis on the present period in contrast to later periods (Laibson 1997). This model can be interpreted as saying that at any moment the individual is 'rational' about the relative value of different future periods, but 'irrational', or 'myopic', about the value of the present period compared to the future, as he overvalues immediate consumption. Another model is proposed by George (2001): he distinguishes between first-order and second-order preferences which can contradict each other. For an overview over further approaches see Nyhus/Webley (2006: 303- 305).

A vivid illustration of this phenomenon is given by experiments in which subjects were asked whether they preferred $110 in 31 days or $100 in 30 days, and whether they preferred $110 tomorrow or $100 now. Often, people prefer $100 now, but also $110 in 31 days (Frederick et al. 2002: 361) This means that if they now take the decision to choose $110 in 31 days, but then have the possibility to choose again on day 30, they will decide against their original plan and prefer $100 at once; this is a so-called 'preference reversal'. The perception of temporally stretched bundles of consumption is thus distorted – one is prone to make decisions one would not make from a more distant perspective and which one is likely to regret afterwards – just as during the excitement of the evening Pjotr Andrejitsch thinks that playing cards is so wonderful that it is worth it, even if the expected outcome is rather negative, and he deeply regrets having entered the game the next morning.

This short description of the economic treatment of *akrasia* illustrates the approach of behavioral economics. The crucial methodological decision is to weaken the assumption of rational egoistic utility maximization. Instead, behavior which is irrational from the point of view of *economic man* but which is found to be very real in empirical studies is taken into account and analyzed systematically. Most behavioral economists try to find more general utility functions or calculus procedures which can incorporate both the observed 'irrational' *and* the 'standard' rational behavior, with the latter being a special case.[9] This allows for mathematical treatment and empirical testing of models, sometimes even for quantifications of behavioral parameters.

Many insights of behavioral economics are well-known by psychology and sociology as well as by folk psychology and anecdotal evidence. Standard economic analysis has deliberately ignored such phenomena in order to keep a unified paradigm easily amenable to sophisticated mathematical analysis. At the moment there exists a lively debate between mainstream economists and behavioral economists about the pros and cons of both sides (e.g. Rabin 2002), and new models and concepts are developed that try to deliver both accurate empirical predictions and a sufficient level of generality. They provide instruments for a more differentiated analysis of economic problems, e.g. the provision of public goods.

3. Taking 'blind spots' into the picture – the case of *akrasia*

Why might economic ethicists be interested in behavioral economics? There are four central arguments. Firstly, there are numerous cases in which behavioral economics yields better predictions than standard economic theory, and can thus be said to be the better theoretical framework, especially for areas "outside the strictly economic (market) sphere" (Frey 1997: 118). In this context, behavioral economics is increasingly concerned with analyzing in which cases *economic man* is – or is not – a useful approximation (see for instance Camerer/Fehr 2006 with the telling title "When does

---

[9] Formally, this keeps behavioral economics within the rational choice paradigm. This paradigm is 'softened', however, because diverse factors can be included into the utility function and humans can be modeled as a sequence of individually maximizing temporal selves. It becomes a matter of practicability whether constructing models in which – mathematical – maximization takes place is worthwhile – and a matter of taste whether one still calls the result 'rational choice'.

'Economic Man' Dominate Social Behavior?"). Thus, behavioral economics can help to find better institutional arrangements by informing about typical behavioral tendencies (cf. also Hollstein 2005: 430).[10] This can also include taking into account empirical findings on perception of justice, which have considerable implications for the acceptance of institutions (see Irlenbusch 2003).

Secondly, behavioral economics deals with ethical matters, e.g. altruistic preferences, the relevance of which for economic ethics has recently been emphasized by Panther (2005). Furthermore, it occupies itself with questions of rationality, e.g. overconfidence, which are important for our understanding of (ethical) decision making. Ethicists and political philosophers in general and economic ethicists in particular might be interested in this research and in the relevance of these tendencies in economic contexts.

A third aspect is that behavioral economics is less strictly bounded to an *a priori* model of man, but tests empirical hypotheses about human behavior. Although many theories do deliver models in terms of utility functions, those utility functions are by far richer than the standard ones. For example, if the utility of other persons can enter the utility function of the economic agent (see Fehr/Schmidt 2006 for an overview) it cannot be said any more that economic agents are by assumption selfish – the critique has to become more differentiated.[11] It is especially important that new topics that have been 'blind spots' in neoclassical theory, but which cause tangible social problems in the real world, can be addressed by behavioral economics.

Last but not least it is likely that behavioral economics will increasingly be used to consult policy makers. In a special issue of the US-American *Journal of Public Policy & Marketing* about this topic, Lynch and Wood write in the Editors' Statement that "policy makers must make the investment to learn more about consumer behavior theory" (2006: 6). In a similar vein, Amir et al. (2005: 450) reflect about how to "get policymakers to listen" to behavioral economists. Bernheim/Rangel (2005) explore the possibilities of public economics, focussing on policy problems concerning saving, addic-

---

10 It is worth mentioning that this is not only a topic for microanalysis, but also relevant on a macroeconomic level; as has recently been argued in a widely noticed speech on the relevance of norm following for macroeconomic theory by Nobel Prize winner George Akerlof (Akerlof 2006).

11 For example, much of the debate between Peter Ulrich and Karl Homann in the recent years was about the degree of liberty to act morally individuals have when competing in free markets (e.g. Homann 2004: 18; Ulrich 1997: 111ff., 161ff.). Research in behavioral economics about preferences for fairness has delivered valuable insights for this debate. Fehr/Schmidt (1999) have developed a model that explains why agents' fairness considerations, which can be shown to exist in ultimatum games, play a very limited role in competitive environments (which can also be shown empirically). The same model can explain how fairness considerations can help solve common good problems when members are given an opportunity to punish shirkers (for an overview over models that take into account payoffs to others and other factors see Fehr/Schmidt 2006). So when it comes to concrete cases, such research can help to evaluate whether the institutional arrangement *allows* people to realize preferences for fairness or not; this is valuable information for economic ethicists who reflect about the normative question of whether they *should* be realized in a given context.

tion and public goods. This new development should not be ignored by economic ethicists.

It is important to bear in mind, however, that using behavioral economics for the purpose of economic ethics carries certain risks. Whereas in neoclassical theory agents do by assumption never harm themselves, from the point of view of behavioral economics this is well possible. For example, a person with a strong sense of fairness might go to great lengths in order to punish a shirker. It is often difficult to judge from an outside perspective whether this is the expression of a true preference (which might still be morally wrong) or an impulsive reaction which should be judged 'irrational'.[12] Intending to be 'closer to reality', behavioral economics thus brings to the fore moral questions, e.g. questions of justice and autonomy. In what follows, I will illustrate this problem by means of the example of *akrasia* concerning saving for the retirement, which is discussed by means of an informative pilot study.

3.1 A saving program for myopic consumers

Akrasia concerning intertemporal allocation of consumption can be a problem especially for those who are responsible for providing their own income in old age. If consumers with self-control problems are aware of them, and sophisticated enough, they might find ways to play tricks on themselves, like Odysseus who had himself tied to the mast when passing by the sirens (Schelling 1978; Elster 1979). It is unclear, however, how many of those affected by *akrasia* manage to do so. Familiar with empirical results concerning myopia, Thaler/Benartzi (2004) have developed the program "Save More Tomorrow" (SMarT®).

For perfectly rational agents the problem of old age saving is easy: individuals will roughly smooth consumption over their entire life, saving in years with higher income and spending the assets in the years after retirement when they have no earned income (Modigliani/Brumberg 1954; Friedman 1957). Although there may be exceptions, I will assume in the following that this is a reasonable assumption: it enables persons to live a responsible and autonomous life and to realize their own conception of a good life. Empirical studies show, however, that this is more easily said than done. In addition to the question of how to calculate an adequate savings rate there are problems of self-discipline, in particular procrastination because of inconsistent time preferences. So many people save less than they actually want (Thaler/Shefrin 1981; Shefrin/Thaler 1988; Bernheim 1995). Macroeconomic data confirm that the U.S. personal savings rate is in fact very low (Thaler/Benartzi 2004: 182).

In the SMarT® program several insights from experimental behavioral research are combined to help people overcome *akrasia*. Employees get an opportunity to bind themselves several months before the program starts; this helps to overcome the present bias. Linking the increase of the savings rate to a pay increase makes it easier for employees to start saving, because a nominal loss in disposable income is avoided.

---

12 Additional evidence that it is not easy to draw a line between 'true preferences' and 'moral action' comes from recent findings in neural economics. DeQuervain et al. (2004), for example, find that humans derive satisfaction from punishing shirkers. I am grateful to an anonymous referee to point this out to me.

This accommodates 'loss aversion', a tendency to weigh (nominal) losses more heavily than gains (cf. Tversky/Kahneman 1981). Every year, the contributions to the savings plan are automatically increased until a defined maximum is reached. It is important to note that "(…) the employee can opt out of the plan at any time" (Thaler/Benartzi 2004: 166). The authors assume, however, that once in the program, many employees would stick to it, because of another behavioral tendency, namely status quo bias or inertia (ibid.: 165-170).

Note that from the point of view of textbook economics the existence of a program like SMarT® should not make any difference: if it is optimal for individuals to raise their savings rate in the way proposed by the program, they will do it anyway. In reality, however, the program does make a difference. Thaler/Benartzi report three implementations in different companies; in all cases the participants raised their savings rate significantly (for details see ibid.: 173, tables 2, 3 and 4).

What is special about the SMarT® program in contrast to other policy recommendations by economists? As the authors write, they "take seriously the possibility that some of these low-saving workers are making a mistake" (ibid.: 166). The program uses the mechanisms that normally impede adequate saving, in particular procrastination and inertia, and employs them in order to attain higher saving rates. Thaler/Benartzi suggest that programs such as SMarT® might be implemented more widely on an opt-out basis as part of a 'libertarian paternalistic approach'. Following Sunstein/Thaler (2003: 185), they define 'libertarian paternalism' as follows:

> "Libertarian paternalism is a philosophy that advocates designing institutions that help people make better decisions but do not impinge on their freedom to choose."

3.2 A Rawlsian perspective

From the point of view of economic ethics the idea of 'libertarian paternalism' belongs to the level of institutional ethics. At first glance, a program like SMarT® seems a wonderful idea: it creates more (old age) welfare for people without directly diminishing the welfare of any others. It is likely that many philosophical schools will agree on that.[13] The one chosen here is a contractarian perspective in a Kantian spirit, which emphasizes the need to protect the autonomy of humans as morally responsible persons as well as the solidarity of humans as social beings. Rawls's well-known idea of the veil of ignorance can serve as a starting point for a moral evaluation (Rawls 1971).

The parties behind the veil of ignorance are especially concerned about the fate of the least-advantaged group in society, because they are risk-averse and they fear ending up in this group. Famously, this is why they use the *maximin principle* to deduce the *difference principle*. Now let us assume that the parties in the original position know that many people in the society are plagued by *akrasia* and that this makes it difficult, if not impossible, for them to behave according to their long-term preferences. The parties are assumed to be afraid of ending up in this group, and they will therefore try to raise

---

13 E.g., a utilitarian approach is very likely to welcome such a program. However, it might also advocate a coercive system if this yields higher (total or average) outcomes.

the welfare level of this group. They could decide to implement institutional structures that will make it less likely that myopia will cause those individuals to enjoy less welfare than they actually could, which would be Pareto-inefficient. However, the choice of institutions is restricted by Rawls's first principle of justice which states that the most extensive basic liberty possible is to be granted to anybody. What the parties are likely to end up with, therefore, is an opt-out system, just as Thaler/Benartzi suggest (Thaler/Benartzi 2004). This does not confine personal autonomy, yet it helps those who are threatened by *akrasia* to reach higher welfare levels.[14]

3.3 Safeguarding autonomy

There remains, however, a discomforting feeling: albeit for a good purpose, 'libertarian paternalism' deliberately exploits human weaknesses. In contrast to the fictitious parties behind Rawls's veil of ignorance who might discuss such programs, real humans might not (or only faintly) be aware of their own psychological tendencies. Mechanisms that target those tendencies might be put to less salutary ends as well; after all, they do not primarily address humans' free will and capacity of reflection, but conceive of them as trapped by psychological mechanisms and governed by situational forces. Does one not play with the fire when using such mechanisms?

In a way, the answer is simply: yes, one does. However, it is not a fire one lights, but one that is burning anyway – the behavioral tendencies in question exist in the 'default setting' as well. Better education may curb them to a certain extent and help people master them, but denying their existence seems unrealistic.[15] In many situations the question is not so much whether people tend to stick to the status quo, for example, but rather which status quo they stick to: one set by historical and sociological contingencies or one deliberately set. Therefore, ethicists should not close their eyes in view of such phenomena, but rather try to handle them in an ethically responsible way.[16] From a contractarian perspective, the crucial point in any policy that wants to help people help themselves is to protect the personal autonomy of those concerned.

---

14 This argument finds further support from self-determination theory (see, for instance, Moller et al. 2006). While "understanding autonomy or the experience of choice as a basic psychological need" (ibid.: 104), these authors emphasize that too much choice or the felt pressure to make a decision can undermine the experience of autonomy. Instead, an autonomy-friendly environment is essential: "[a]utonomy-supportive communications are intended to let people engage in mindful considerations of what is right for them" (ibid.: 109). It should go without saying that for policies such as SMarT® a respectful, friendly communication is of highest importance.

15 It should be noted, however, that education and other means are not always very successful at 'debiasing' humans; the success depending on the kind of bias and the method of debiasing, cf. Fischhoff 1982, Larrick 2004.

16 *A fortiori*, as other actors, e.g. the credit card industry, are likely to have less qualms about knowingly putting to use psychological mechanisms of consumers – in ways which can be ethically questionable (cf. Bertrand et al. 2006: 8, 20). The delicate balance about how much the liberty of companies trying to sell their goods should be confined in order to protect consumers from their own myopic wishes is a topic worthwhile to be pursued by economic ethicists as well. Lynch/Zaubermann (2006), for example, discuss the legitimacy of tempting rebates and similar offers which build on consumers' inconsistent temporal preferences. This leads into the area of marketing ethics (see Whysall 2000 for an overview).

A first caveat to be mentioned is that it would be a categorical mistake to deduce compulsory measures from the fact that there are detrimental behavioral tendencies. If the problem at hand is really caused by behavioral tendencies only, it should be possible to solve it by putting to use behavioral factors, e.g. by changing the default setting, but leaving open a possibility to opt out. The introduction of binding schemes, by contrast, needs to be justified differently, e.g. by the fact that the incentives in a given situation carry the features of a prisoners' dilemma, present in many areas of social insurance. This is especially relevant to cases in which it can be rational for individuals – maybe only for a minority – to behave differently. As Jon Elster has argued in a similar context, any institution that has to do with the internal forces within the individual has to make sure that agents must not be forced to accept offers of 'help' if it can be rational for them to reject them (Elster 1979: 81). Any participation must be based on – at least implicit – consent. Besides, the 'exit option' has to be clearly communicated; it must neither be hidden in small print nor be complicated by bureaucratic hurdles.

A further requisite is nicely summed up by Elster in the following sentences: "Exploiting intrapsychic mechanisms that are unknown to the individual can never be justified" (Elster 1979: 82). Whenever possible, the parties involved should have an open discussion about the problems involved and in which the behavioral mechanisms are explained to everybody, and then agree *ex ante* on desired behavior and design institutions accordingly.[17] As long as a program like SMarT® is offered on an opt-in basis it is likely that it will be chosen only by those employees who know about their own character traits and realize that the program uses exactly those traits to make them save more. If enrolment is automatic this might not be the case. To honor the moral autonomy of those enrolled it is necessary to let them know how the program works; it should be openly explained that it is successful because of tendencies like myopia and inertia.[18]

Of course, agents have to be informed about all details of the program, e.g. about the way in which the savings rate is calculated. Communication of those features will make a public discourse possible, which can help to prevent, or at least detect, cases of abuse or manipulation. The fact that a program builds on a-rational factors to overcome weakness of the will does not make it inaccessible to rational and critical discourse; on the contrary, it is *a fortiori* desirable that an open discussion takes place. Open communication can also serve as a litmus test for the righteous intentions of those who implement the program. In his 1795 *Perpetual Peace*, Kant identifies publicity ('Publizität') as condition for any just policy (Kant 1983: 140ff.). For a program in the spirit of 'liberal paternalism' to be ethically acceptable, it must be possible to publicly justify why psychological mechanisms are used. This can serve to distinguish between

---

17 I am grateful to an anonymous referee for pointing this out to me.

18 This might be done by a letter saying something like the following: "It is a common experience to many people that although they intend to save for the retirement they postpone the start of a saving plan from month to month (…) This program is designed to overcome these problems by a scheme of automatic increases of the savings rate (…) Most people tend to abide by the program once they have agreed to take part (…)."

programs that really want to help people overcome own weaknesses and realize their long-term preferences, and such programs that have other, less acceptable aims.[19] Furthermore, it must be openly communicated if there are indirect aims to be reached by such measures (e.g. raising the national saving rate in order to provide more capital to the industry). It would be morally objectionable not to let enrolled workers know that there are further motivations for the implementation: they would be treated as means to somebody else's ends, instead of being given free choice whether to accept this side effect.

4. Conclusion

Akrasia is only one area of research in behavioral economics that is important for economic ethics. Other topics are of a similar relevance and have partly been taken up,[20] e.g. norm following, social preferences, habit formation, the role of intentions, or crowding out of intrinsic motivation by monetary incentives. They have to be considered carefully, because it is to a far lower degree clear that deviations from the model of neoclassical economics are harmful to the agent himself or to others. In many cases one might rather have to question adequacy of the narrow economic notion of utility or rationality than consider ways of changing people's behavior. This may be complicated in detail, and demand a large amount of knowledge of local conditions and people's perception of situations. Not to take such behavior into account, however, can lead to serious inefficiencies, e.g. when intrinsic motivations are crowded out by monetary incentives (Frey 1997: part II).[21]

One might ask whether it would not be desirable to develop an overarching, psychologically realistic model of man, integrating sociological, psychological, political and cultural aspects, on which economic ethics could build. Apart from questions of realizability, however, such a model would presumably make it very difficult to deliver results of some degree of generality. Bruno Frey, a leading behavioral economist, suggests a more pragmatic alternative: "[i]t leaves the partial models including some specific psychological effects as they are, and relegates the task of choosing the 'appropriate' model to the problem at hand" (Frey 1997: 124). Depending on the problem to be

---

19 E.g. could a sentence like the following be added: "Research has shown that a majority of people are tempted to spend too much on present consumption instead of saving. Therefore we offer the possibility (…)". Contrast this with "Research has shown that hungry people tend to buy more junk food than they really wish from a long term perspective. We want to sell junk food to you, therefore we try to catch you when you are extremely hungry (…)".

20 As mentioned above, Panther (2005) discusses the consequences of research on altruistic preferences for economic ethics. Possible impacts that research on reciprocity has on public policy are discussed by Falk (2001). From a slightly different perspective (i.e. evolutionary economics and sociology), Beschorner (2000) discusses the role of habitual actions as a factor in business ethics.

21 Another prominent example of an 'inefficiency' is the discrepancy between the willingness to donate organs and the actual rates of donations in countries with an 'opt-in', or 'explicit-consent' system for organ donation (cf. Johnson/Goldstein 2003), and the debate about opt-in versus opt-out defaults. This problem is structurally slightly different from the *akrasia*-case: rather than helping people to help themselves, the question is how one can help people realize their good intentions to help others, without restricting the liberty of those who do not want to donate.

solved, the concepts and findings of behavioral economics, social psychology and other disciplines can be applied in addition to the purely neoclassical concepts. Frey hopes that "putting the problem [instead of some model] first", "could turn economics from a technique driven science into an art" (ibid.).[22] This empirically informed 'art' is a highly attractive partner for cooperation with economic ethicists, and ethical theory might in turn inspire it.[23] The approach of behavioral economics might thus be a way for fruitful interdisciplinary work for ethicists, social scientists and economists. For example, the question about the moral qualities of capital markets can be asked in a much more differentiated way if the structure of specific markets is analyzed more thoroughly, including phenomena such as biases, overconfidence and herding (see e.g. Dorsey 2004).

Beyond this relevance for more or less 'applied' matters in economic ethics, behavioral economics can also have an impact on the systematic level of economic ethics. This affects the question of 'who is to be blamed'. If individual ethics is taken as starting point for economic ethics, the responsibility for undesirable facts must lie with one or more – 'evil' – individuals. With neoclassical economics as the methodological starting point, there is an important extension: beyond the fact that people might have evil preferences, the structure of social interaction can be dilemmatic, leading from innocuous preferences to socially unwanted results. In the latter case it is the institutional framework that needs to be changed – this has been especially emphasized in Homann's institutional ethics (Homann/Suchanek 2005: 31ff.). Behavioral economics delivers an additional insight: seemingly minor behavioral tendencies or situational factors, ignored by traditional economics, might be responsible for unwanted consequences. It is an important insight that 'big' problems need not necessarily have 'big' causes; and as Bertrand et al. (2006: 10) write, "large manipulations can sometimes have negligible effects, whereas apparently small manipulations can have a dramatic influence".[24] The authors suggest using this insight in fighting poverty, e.g. by getting

---

22 Suchanek (2005) criticizes the use of behavioral economics for economic ethics because he wants to keep the methodological purism of *economic man plus restrictions* (see also Homann/Meyer 2005: ch. 4). With an open notion of utility and a regard for 'relevant' restrictions in different situations, however, this paradigm already admits many elements which do not belong to traditional *economic man*, but are systematically explored by behavioral economics (e.g. other-regarding preferences). It seems arguable, therefore, whether this criticism of empirical economics (Homann/Suchanek 2005: 381ff., Homann/Meyer 2005: 101) is justified, as Homann and his followers admit that research into what counts as utility and about how the model of economic man can be applied is very useful for their approach (Homann/Meyer 2005: 100; Suchanek 2005: 101) – this is in fact what many behavioral economists do.

23 To foreclose possible misunderstandings: the point about 'art' is not subjective arbitrariness, but the acknowledgment that there is no strict (mathematical) procedure to which one could stick. Rather, a certain faculty of judgment is necessary to deliver good results, e.g. for seeing which non-monetary factors might enter people's utility functions. So 'art' is understood not so much in the aesthetic sense alone but rather in the sense of, say, Paul Bocuse's art of cooking.

24 This insight does not offer an excuse for those responsible – after all, they might have had the responsibility to take into account such circumstantial factors. However, economic ethics is not so much interested in the moral (and legal) judgment of potential evildoers (this rather belongs to moral and legal philosophy), but in bringing about better institutions, and for that task the analysis of causal connections is more important. To take a simple example from business ethics: stat-

more US-American working poor into welfare programs to which they are entitled. Overcoming seemingly "minor" obstacles, such as "small hassle factors" (e.g. long and complicated application forms), perceived identities and shame, can be expected to have a large impact (ibid.: 16ff.). As this example also shows, knowledge of general tendencies of behavior, such as *akrasia*, often has to be combined with detailed knowledge of local cultures, habits and interpretations. Instead of armchair economics, dialogue with the persons concerned will be necessary to find clever solutions to social problems.

By the way, in Pushkin's novel, young Pjotr Andrejitsch's foolishness turns out to have answered a good purpose: Iwan Surin, to whom he lost his hundred rubles, is going to save his life later in the novel. After all, the freedom to make mistakes and to learn from them, especially when one is young, is part and parcel of human development. So to some degree it might well be included into the autonomy ethicists want to preserve in economic matters.

References

Akerlof, G. (2006): The Missing Motivation in Macroeconomics, Preliminary Draft (November 15, 2006): Presidential Address, American Economic Association, Chicago, IL, January 6, 2007 (available at (13/07/2007): http://www.aeaweb.org/annual_mtg_papers/2007/0106_1640_0101.pdf).

Altman, M. (2006): Introduction, in: Altman, M. (Ed.): Handbook of Contemporary Behavioral Economics: Foundations and Developments, New York: M.E.Sharpe, xv-xxii.

Amir, O. et al. (2005): Psychology, Behavioral Economics, and Public Policy, in: Marketing Letters 16(3/4), 443-454.

Aristotle (1985): Nicomachean Ethics (NE), translated by Terence Irwin. Indianapolis: Hackett Publishing Company.

Bertrand, M./ Mullainathan, S./ Shafir, E. (2006): Behavioral Economics and Marketing in Aid of Decision Making Among the Poor, in: Altman, M. (Ed.): Handbook of contemporary behavioral economics: foundations and developments. New York: M. E. Sharpe, 8-21.

Bernheim, B. D. (1995): Do Households Appreciate Their Financial Vulnerabilites? An Analysis of Actions, Perceptions, and Public Policy, in: Tax Policy and Economic Growth, Washington DC: American Council for Capital Formation, 1-30.

Bernheim, B. D./ Rangel, A. (2005): Behavioral Public Economics: Welfare and Policy Analysis with Non-Standard Decision Makers, NBER Working Paper 11518 (available at (13/07/2008) www.nber.org/papers/w11518).

Beschorner, T. (2000): „Achte auf Deine Gewohnheiten…"– Eine institutionenethische Betrachtung aus Sicht der Evolutorischen Ökonomik, in: Zeitschrift für Wirtschafts- und Unternehmensethik, 1(1), 64-83.

Bowles, S. (1998): Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions, in; Journal of Economic Literature 36(1), 75-111.

ing that corporate culture is an important factor in preventing bribery does not at all discharge those guilty of bribery, but it gives a clue as to what a manager who wants to foreclose corruption might do to achieve this end.

Camerer, C./ Loewenstein G. (2003): Behavioral Economics: Past, Present, Future, in: Camerer, C./ Loewenstein, G./ Rabin M. (Eds.): Advances in Behavioral Economics, Princeton: Princeton University Press.

Camerer, C./ Fehr, E. (2006): When Does 'Economic Man' Dominate Social Behavior?, in: Science, 311(5757), 47-52.

DeQuervain, D. J.-F./ Fischbacher, U./ Treyer, V./ Schellhammer, M./ Schnyder, U./ Buck, A./ Fehr, E. (2004) The Neural Basis of Altruistic Punishment, in: Science, 305(5688),1254-1258.

Dorsey, W. (2004): Behavioral Trading: Methods for Measuring Investor Confidence and Expectations, New York: Texere.

Elster, J. (1979): Ulysses and the Sirens. Studies in Rationality and Irrationality, Cambridge: Cambridge University Press.

Falk, A. (2001): Homo Oeconomicus Versus Homo Reciprocans: Ansätze für ein Neues Wirtschaftspolitisches Leitbild, Institute of Empirical Research in Economics, University of Zurich, Working Paper No. 79.

Fehr, E./ Schmidt, K. M. (1999): A Theory of Fairness, Competition, and Cooperation, in: The Quarterly Journal of Economics, 114(3), 817-868.

Fehr, E./ Schmidt, K. M. (2006): The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories, in: Kolm, S./ Ythier, J. M. (Eds.): Handbook on the Economics of Giving, Reciprocity and Altruism, Vol. 1, Amsterdam/London: Elsevier.

Fischhoff, B. (1982): Debiasing, in: Kahneman, D./ Slovic, P./ Tversky, A. (Eds.): Judgement under Uncertainty: Heuristics and Biases, New York: Cambridge University Press, 422-444.

Frederick, S./ Loewenstein, G./ O'Donoghue, T. (2002): Time Discounting and Time Preferences: A Critical Review, in: Journal of Economic Literature, 40(2), 351-401.

Frey, B. (1997): Not Just for the Money: An Economic Theory of Personal Motivation, Cheltenham: Edward Elgar Publishing.

Friedman, M. (1953): The Methodology of Positive Economics, in: Essays in Positive Economics, Chicago: University of Chicago, 3-43.

Friedman, M. (1957): A Theory of Consumption Function, Princeton: Princeton University Press.

George, D. (2001): Preference Pollution: How Markets Create the Desires We Dislike, Ann Arbor: University of Michigan Press.

Hausman, D. (1992): The Inexact and Separate Science of Economics, Cambridge: Cambridge University Press.

Hollstein, B. (2005): Gerechtigkeit – Ein aktueller Topos in der ökonomischen Theorie? Workshopbericht, in: Beschorner, T./ Hollstein, B./ König, M./ Lee-Peuker, M.-Y./ Schumann, O. J. (Hrsg.): Wirtschafts- und Unternehmensethik. Rückblick – Ausblick – Perspektiven, München/Mering: Hampp, 425-430.

Homann, A. (2004): St. Galler und Münchener Wirtschaftsethik: Gemeinsamkeiten und Differenzen, in: Ulrich, P./ Breuer, M. (Ed.): Was bewegt die St. Galler Wirtschaftsethik, St. Gallen: Institut für Wirtschaftsethik der Universität St. Gallen, 18-21.

Homann, K./ Meyer, M. (2005): Wirtschaftswissenschaftliche Theoriebildung, Braunschweig: Westermann.

Homann, K./ Suchanek, A. (2005): Ökonomik – Eine Einführung, 2. Auflage Tübingen: Mohr Siebeck.

Irlenbusch, B. (2003): Auf der Suche nach Gerechtigkeit – Eine empirische Herangehensweise, in: Zeitschrift für Wirtschafts- und Unternehmensethik 4(3), 351-371.

James, S. (2006): Taxation and the Contribution of Behavioral Economics, in: Altman, M. (Ed.): Handbook of Contemporary Behavioral Economics: Foundations and Developments. New York: M. E. Sharpe, 589-598.

Johnson, E. J./ Goldstein, D. (2003): Do Defaults Save Lives? Science 302(5649), 1338-1339.

Kant, I. (1983): Perpetual Peace, and Other Essays on Politics, History, and Moral Practice, translated by Ted Humphrey, Indianapolis: Hackett Publishing Company.

Laibson, D. (1997): Golden Eggs and Hyperbolic Discounting, in: Quarterly Journal of Economics, 112(2), 443-77.

Larrick, R. P. (2004) Debiasing, in: Koehler, D. K./ Harvey, N. (Eds.), Blackwell Handbook of Judgment and Decision Making, New York: Blackwell Publishing, 316-337.

Lynch, J. G. Jr./ Wood, W. (2006): Special Issue Editors' Statement: Helping Consumers Help Themselves, in: Journal of Public Policy & Marketing, 25(1), 1-7.

Lynch, J. G. Jr./ Zauberman, G. (2006): When Do You Want It? Time, Decisions, and Public Policy, in: Journal of Public Policy & Marketing, 25(1), 67-78.

Modigliani, F./ Brumberg, R. (1954): Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data, in: Kurihara, K. (Ed.): Post-Keynesian Economics, New Brunswick: Rutgers University Press, 388-438.

Moller, A. C./ Ryan, R. M./ Deci, E. L. (2006): Self-Determination Theory and Public Policy: Improving the Quality of Consumer Decisions Without Using Coercion, in: Journal of Public Policy & Marketing, 25(1), 104-116.

Morgan, M. S. (2006): Economic Man as Model Man: Ideal Types, Idealization and Caricature, in: Journal of History of Economic Thought, 28(1), 1-27.

Nyhus, E. K./ Webley, P. (2006): Discounting, Self-control, and Saving, in: Altman, M. (Ed.): Handbook of Contemporary Behavioral Economics: Foundations and Developments. New York: M. E. Sharpe, 297-325.

North, D. C. (1990): Institutions, Institutional Change and Economic Performance, Cambridge et al.: Cambridge University Press.

Panther, S. (2005): Wirtschaftsethik und Ökonomik, in: Beschorner, T./ Hollstein, B./ König, M./ Lee-Peuker, M.-Y./ Schumann, O. J. (Hrsg.): Wirtschafts- und Unternehmensethik. Rückblick – Ausblick – Perspektiven, München /Mering: Hampp, 79-94.

Pushkin, A. (1992): The Captain's Daughter and Other Stories; Translated from the Russian by Paul Debreczeny, New York: Random House.

Rabin, M. (2002): A Perspective on Psychology and Economics, UC-Berkeley Working Paper No. E02-313.

Rawls, J. (1971): A Theory of Justice, Cambrige, MA.: Belknap Press of Harvard University Press.

Robbins, L. (1932): The Nature and Significance of Economic Science, London: Macmillan.

Schelling, T. (1978): Egonomics, or the Art of Self-Management, in: American Economic Review 68(2), 290-94.

Schlicht, E. (2003): Der *homo oeconomicus* unter experimentellem Beschuß, in: Held, M./ Kubon-Gilke, G./ Sturn, R. (Eds.): Experimente in der Ökonomik. Jahrbuch 2: Normative und institutionelle Grundfragen der Ökonomik, Marburg: Metropolis, 291-313.

Searle, J. R. (2005): What Is an Institution?, in: Journal of Institutional Economics, 1(1), 1-22.

Shefrin, H./ Thaler, R. H. (1988): The Behavioral Life-Cycle Hypothesis, in: Economic Inquiry 26 (4), 609-43.

Strotz, R. H (1956): Myopia and Inconsistency in Dynamic Utility Maximization, in: Review of Economic Studies, 23(3), 165-80.

Suchanek, A. (2005): Die Relevanz der experimentellen Ökonomik für die Wirtschaftsethik, in: Beschorner, T./ Hollstein, B./ König, M./ Lee-Peuker, M.-Y./ Schumann, O. J. (Hrsg.): Wirtschafts- und Unternehmensethik. Rückblick – Ausblick – Perspektiven, München/ Mering: Hampp, 95-103.

Sunstein, C. R./ Thaler, R. H. (2003): Libertarian Paternalism Is Not an Oxymoron, in: University of Chicago Law Revue, 70(4): 1159-99.

Thaler, R. H./ Shefrin, H. (1981): An Economic Theory of Self-Control, in: Journal of Political Economy, 89(2), 392-406.

Thaler, R./ Benartzi, S. (2004): Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving, in: Journal of Political Economy, 112(1), 164-187.

Tversky, A./ Kahneman, D. (1981): The Framing of Decisions and the Psychology of Frames, in: Science, 211(4481), 453-463.

Ulrich, P. (1997): Integrative Wirtschaftsethik. Grundlagen einer lebensdienlichen Ökonomie, Bern et al.: Haupt.

Ulrich, P. (2002): Ethische Vernunft und ökonomische Rationalität zusammendenken. Ein Überblick über den St. Galler Ansatz der Integrativen Wirtschaftsethik. St. Gallen: Institut für Wirtschaftsethik der Universität St. Gallen.

Whysall, P. (2000): Marketing Ethics – An Overview, in: The Marketing Review, 1(2), 175-195.