Martin Jacobs*

Accounting for Changing Tastes: Approaches to Explaining Unstable Individual Preferences

DOI 10.1515/roe-2015-1007

Abstract: Standard economic theory assumes individual preferences to be fixed and exogenously given. This view has been challenged by numerous empirical observations. In reaction to those challenges, economic theory has been modified, mostly by including additional arguments into individuals' utility functions. Among the approaches that tackle preference change are new consumer theory, habit formation, interdependent and status preferences, social and emotional influences, and reference point-dependent preferences. Hence, while standard economics largely abide by their assumption of stable preferences, an array of alternative approaches is now available to account for changing tastes. Some of these approaches are old and have been discussed in the literature for many decades while others are younger. However, all approaches have in common that they, in some cases surprisingly, have not made it to standard microeconomics textbooks. This survey aims at putting the approaches in perspective. For each of them, empirical evidence as well as methodological issues are discussed.

Keywords: individual preferences, preference change, rationality, behavioral economics

JEL Classification: B41

1 Introduction

Standard economic theory usually assumes individual preferences to be exogenously given and stable. This assumption extends to both normative and positive theory.

Within normative theory, consistent behavior is demanded to ensure rationality. Consistency, in turn, is ensured by a stable preference relation. FRIEDMAN (1953) justified the claim for consistency with an argument of natural selection. Individuals who disobey the rules of rationality were exploited and over some

*Corresponding author: Martin Jacobs, Department of Economics, University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany, E-mail: jacobs@economics.uni-kiel.de

time crowded out by more rational actors. FRIEDMAN's argument is highly controversial (see e.g. ALCHIAN, 1950). However, much more basic thoughts also provide justification for at least some degree of rationality in behavior. For, if an individual acts completely unsystematically, one cannot in a meaningful way talk about that individual's preferences, but would have to conclude that any kind of individual preference were absent. Thus, the term "preference" itself implies a certain minimum degree of systematicness.

As for positive theory, numerous observations of real life behavior seem to confute stable individual preferences. Illustrative examples comprise e.g. fads or addictive behavior. Some economists, however, have passed on justifying their assumption of stable preferences not by its realism, but other reasons. FRIEDMAN employed the argument of a scientific division of labor: "[E]conomic theory proceeds largely to take wants as fixed. This is primarily a case of division of labor. The economist has little to say about the formation of wants; this is the province of the psychologist. The economist's task is to trace the consequences of any given set of wants" (1971, p.13). MICHAEL and BECKER (1973) argued that neither economics nor any other social science possess a well-developed theory of preference formation. Abolishing the assumption of stable preferences would therefore lead to a model that could explain any empirical observation by claiming that individual preferences have changed without providing any new insights as it stays unclear why and how preferences change.

This critique of arbitrariness, however, can also be applied to a model which assumes preference stability. As the *ceteris paribus* condition is never perfectly fulfilled in reality, any observed sequence of decisions can be rationalized *ex post* within the framework of stable preferences, provided that the assumed preference structure is complex enough.¹ Strictly speaking, the question of whether individual preferences are stable is a matter of definition. If a change in behavior is observed, one can either assume a change in the individual's utility function and thus preferences or one can look for a parameter that has changed over the same course of time and integrate this parameter into the utility function. By this procedure, the individual's preference structure, while becoming dependent on more exogenous factors, retains its stability. Even if such a parameter cannot be found, one can still assume that an unobservable relevant factor has changed. A prime example for such an unobservable parameter might be human capital as employed by STIGLER and BECKER (1977).

Standard economic theory assumes that individual choice is influenced only by prices and income. This assumption is clearly unrealistic. For example, if an individual's marginal rate of substitution between ice-cream and winter

¹ This point was also made by ANDERSEN ET AL. (2008) and SOBEL (2009).

boots and thus her chosen consumption bundle changes from January to July though prices and income remain constant, this will hardly be to the surprise of anyone. Yet, as FRIEDMAN (1953) pointed out, unrealistic assumptions in models are not *per se* problematic. The quality of a model should be judged from the exactness of its predictions concerning the purpose of its creation rather than from the reality of its assumptions. If a more complex model that yields more exact predictions exists, complexity and exactness are to be traded off against each other.

The purpose of this survey is to introduce and to assess such more complex approaches that include additional arguments into the individual's utility function and are in this way able to explain actual human behavior more exactly. Due to these models, MICHAEL's and BECKER's claim that there were no useful theories of preference formation seems indefensible from today's point of view. Note that these models largely retain the rationality assumptions imposed on preferences. Therefore, before proceeding to the models, some more remarks are made on the relationship between preferences and rationality in Section 2. Sections 3 and 4 deal with the attempts to defend the assumption of stable preferences using incomplete information or new consumer theory, respectively. Section 5 looks at adaptive preferences from a psychological and philosophical viewpoint. Sections 6–9 introduce approaches that account for habit formation, interdependent and status preferences, emotional effects, and institutional influences on individual preferences. Section 10 considers reference points in decision-making and prospect theory. Section 11 concludes.

2 Preferences and rationality

Preferences are buried inside individuals and cannot be observed directly. Economists usually circumvent this problem by assuming that individuals act according to their preferences so that preferences are revealed by observable choices.

The approach of revealed preference comes with limitations and implicit assumptions, as pointed out by HAUSMAN (2012). Preferences can be elicited only when a choice is made, so preferences among objects that are unavailable for choice remain latent. When preference is inferred from an act of choice, it is implicitly assumed that the chooser correctly and completely perceived all relevant aspects of the choice situation. When preference is inferred from a combination of several acts of choice, it furthermore is assumed that context effects are absent. Despite those limitations, HAUSMAN endorsed the tight connection between preference and choice in economic theory.

In contrast, others have criticized revealed preference as too narrow a concept. For example, an individual could consider two alternatives incommensurable and nevertheless, and if only at random, choose one of them (see GRÜNE-YANOFF and HANSSON, 2009b; SEN, 1973). SEN (1977) criticized that revealed preference theory labels every man an egoist since by definition he cannot escape from maximizing his own utility. Whatever his utility function may look like, acting against his own preferences is defined impossible. According to SEN (1973, 1994), a single preference ordering does not suffice to correctly represent all aspects of individual preferences. Those aspects comprise self-interest as well as well-being, social norms, and decision. Norms might drive a wedge between self-interest and observed behavior. A similar point was made by SCHICK (1984). The idea that an individual owns several distinct preference orderings is corroborated by KAHNEMAN'S, WAKKER'S and SARIN'S (1997) distinction between "experienced utility" and "decision utility".

Furthermore, the implicit assumption of absent context effects has been argued to be often violated. SEN (1993, 1997) discussed both chooser- and menu-dependent preferences. Chooser dependence means that one's preference might reverse depending on whether a choice is made by oneself or by someone else. The reason for the reversal may lie in the individual's desire to abide with a social norm. A person who would never choose to take the biggest slice of a cake may be happy to be forced to take it. Menu dependence refers to a preference change caused by an extension or contraction of the choice set. Its reason may again be a social norm. A person choosing not to take the very last biscuit may have a different preference if there is plenty left over for others. Though, an individual's valuation for freedom of choice *per se* or additional information conveyed by the choice set may also play a role in that case. To clarify the last point, Sen gave the example of a man who is happy to accept the invitation of an acquaintance when he is offered some tea while he readily turns it down when he is offered either some tea or some cocaine.² Choice set-dependent preferences were also considered by TVERSKY and SIMONSON (1993) and KÓSZEGI and RABIN (2008). Another aspect of context-dependence occurs when preferences are influenced by variations of seemingly irrelevant aspects of the choice situation, e. g. framing effects. Such effects and approaches to include them in economic models are discussed extensively in Section 10 of this survey.

² In a similar vein, but more generally, HSEE ET AL. (1999) suggested that adding further options can lead to a preference reversal due to an increasing weight of difficult-to-evaluate attributes which gain most from the possibility of comparison when many alternatives are assessed jointly.

Besides implicit assumptions, revealed preference theory rests on a set of explicitly stated rationality axioms. Standard economic theory has developed different versions of revealed preference theory which connect to different sets of rationality axioms (see HOUTHAKKER, 1950; SAMUELSON, 1938). An overview and classification of them can be found in GRÜNE-YANOFF and HANSSON (2009a). Some axioms have to be seen as constitutive, i. e. obedience to them is conceptually necessary to be able to claim that a person holds preferences at all. Widely accepted constitutive axioms are asymmetry of preference ($APB \rightarrow \neg(BPA)$),³ symmetry of indifference ($A/B \rightarrow B/A$), reflexivity (A/A), and incompatibility of preference and indifference ($APB \rightarrow \neg(A/B)$). A person who disobeys constitutive rationality axioms can act totally unpredictably; therefore, one cannot construct a meaningful model to describe such behavior. Further axioms demanded by revealed preference theory concern transitivity and completeness of preferences and are controversial. The concept of transitivity includes two axioms: transitivity of preference ($APB \wedge BPC \rightarrow APC$) and transitivity of indifference ($A/B \wedge B/C \rightarrow A/C$). The necessity of transitivity axioms is discussed in the following paragraphs. Completeness means that, for any pair of alternatives, an individual must be able to state that she either prefers the one or the other alternative or that she is indifferent between the two ($APB \vee BPA \vee A/B \forall A, B$). This axiom is very demanding and seems unrealistic with respect to many complex situations people face in their lives. A more fundamental critique to the completeness axiom is the above-mentioned claim that two alternatives can be simply incommensurable.

The adequacy of transitivity axioms has been discussed extensively. ARMSTRONG (1939) criticized the transitivity assumption for indifference. A person may be confronted with a range of alternatives. Let each pair of two “neighboring” alternatives be almost identical such that the two practically cannot be distinguished from each other. The chooser will therefore be indifferent between any two neighboring alternatives. Transitivity now implies that she also has to be indifferent between all alternatives, whereas it can easily be imagined that she may have a preference for the one or the other if two options from the opposing ends of the selection range are compared. SUGDEN (1985) and ANAND (1993) provided further examples of preference patterns that appear reasonable from everyday experience, yet violate transitivity. SEN (1977) argued with respect

³ In this and in the following logical expressions, “*P*” and “*I*” refer to strict preference and indifference, respectively. Thus, *APB* reads “*A* is preferred to *B*”, and *A/B* reads “Neither is *A* preferred to *B*, nor *vice versa*”. “ $\neg$ ” denotes logical negation, “ $\wedge$ ” logical conjunction, and “ $\vee$ ” logical disjunction.

to rationality axioms in general that a real man who obeys all those axioms would behave peculiarly and appear to others as a “rational fool”.

A justification of the transitivity assumption is the so-called “money pump argument” which is connected to FRIEDMAN’S reasoning mentioned in the introduction. The typical story goes like that: let an agent own an intransitive preference ordering according to which she prefers B over A and C over B, but A over C. The argument says that she can now be offered a sequence of transactions in the end of which she is left with her initial endowment, but has lost some money. Let the agent be endowed with A. In this situation, she will be willing to pay a sufficiently small amount of money to exchange A for B as she prefers the latter over the former. Following a similar reasoning, she will then be willing to pay for an exchange for C, and, in the last step, for exchanging C for A, her initial endowment. Repetition of such cycles would eventually lead to the agent being crowded out of the economy. The money pump argument is highly controversial. CUBITT and SUGDEN (2001) demonstrated that exploitability by a money pump does not necessarily follow from an intransitive preference pattern. SUGDEN (2004) furthermore argued that, even if consumers do have in principle exploitable intransitive preferences, competition between traders in a market will ensure that consumers are not exploited in equilibrium. The main line of criticism draws upon the implicit assumption of context-independent preferences in the argument (see e.g. ANAND, 1987, 1993; SCHICK, 1986). If the agent sees an exchange offer in the context of already made exchanges, she may recognize the attempt of exploitation and abstain from further transactions. More specifically, the argument does not take account of the choice set dependence involved in bringing the intransitivity about. Intransitive, choice set-dependent preferences as described above do not lead to exploitation unless the exploiter manages to manipulate the choice set again and again, which is unrealistic when agents are not restricted to one trading partner. Allowing for choice set dependence, it does not follow from the preference pattern in the example that the agent’s preferences are intransitive when she is confronted with all three options at a time. Regret theory which is dealt with in Section 8 as one approach to describe human behavior more realistically also introduces choice set dependence and abandons the transitivity assumption.

Empirical evidence for rationality axioms is mixed. Inspection of consumption time series data rarely indicated violations of revealed preference theory (see KOO, 1963; LANDSBURG, 1981). However, such studies involve some methodological problems. Firstly, due to for the most part steadily increasing incomes in the data sets, there may have been few opportunities for consumption patterns violating revealed preference theory to emerge. Secondly, the observations range

over such long periods of time that one may claim that preference change cannot be ruled out as an explanatory factor for observed violations (for an explanation and discussion of the test methods, see GRÜNE, 2004). Laboratory experiments that avoid these problems, in contrast, have found substantial violation rates (see SIPPEL, 1997). The money pump argument found experimental support in a study by CHU and CHU (1990) in that exploitation of intransitive preferences led subjects to commence behaving consistently. All in all, however, experimental evidence is greatly at odds with rationality axioms. Studies revealed a lot of phenomena that are inconsistent with different axioms, among which are such today well-known patterns as the paradoxes discovered by ALLAIS (1953) and ELLSBERG (1961), and the reversal of preference depending on whether preferences are elicited via direct choice or pricing (see GREETHER and PLOTT, 1979; LICHTENSTEIN and SLOVIC, 1971; LINDMAN, 1971). An overview of experimental violations of rationality axioms can be found in CONLISK (1996).

Some methodological problems are inherent in empirical preference tests (see GRÜNE, 2004). The second problem mentioned in the above paragraph applies in general. As preferences are unobservable, preference stability and adherence to rationality axioms can only be tested together. When a sequence of choices is observed that is not in accordance with standard theory, the reason can be either preference instability or irrationality (or both). In a procedure that claims to test for rationality, stability is implicitly assumed. Similarly, a test for stability assumes rationality. If a parameterized model is put to the test, misspecification constitutes another possible explanation for unexpected results. Following the ideas of Sen described above, still another explanation may be that an agent indeed has rational and stable preferences, but has chosen to act against her preferences. Even if a test does not reveal unexpected behavior, stability and rationality of preferences cannot be taken for granted, for observationally consistent behavior could also be the result of mere coincidence.

A lot has now been said about rationality of preferences, although the topic of this survey shall be their (in-) stability. The reason for that is that most of the approaches discussed in the following sections, while introducing additional arguments of utility functions, retain the assumptions on preference rationality. The purpose of this section was to show that these assumptions have to be handled with care and should be kept in mind during the rest of this survey. Of course, assuming complete irrationality cannot be a fruitful way to explain behavior. However, some economists have proposed a middle course under the label of “bounded rationality”. The ideas and implications for preference stability connected to that approach are discussed in the next section.

3 Incomplete information and bounded rationality

Decision-making under incomplete information can lead to outcomes regarded suboptimal from *ex post*. Thus, when new information becomes available either exogenously or by endogenous learning processes, re-evaluation of alternatives leading to changed behavior may take place. It depends on the kind of incompleteness of information whether individual preferences should be considered stable or unstable in such situations.

The concept of bounded rationality can be traced back to SIMON (1955). Its fundament lies in the idea that cognitive processing capacities of human beings are to be seen as scarce resources which have to be economized. If deliberation costs exceed the expected gains from a completely rational decision, acting under bounded rationality, e. g. by employing heuristics, will be optimal *ex ante* (see CONLISK, 1996; FREY and EICHENBERGER, 1994). CONLISK recognized that following this argument strictly leads to an infinite regress problem. When one thinks about how much cognitive capacity to employ in one's decision-making, second-order costs accrue as scarce cognitive capacity is again needed for coming to this decision. To find the optimal level of second-order costs, further deliberation conferring third-order costs is needed, and so on. Evolutionary psychology provides a loophole, claiming that evolutionary forces have chosen the level of rationality of humans that is adaptive to their evolutionary environment. It is to be considered, however, that evolutionary psychology also suggests that evolution has shaped human minds more generally, i. e. it has determined which problem solution techniques to employ in certain situations. From this point of view, there is no reason to assume rational calculus as a natural state or starting point (see AUMANN, 2008; COSMIDES and TOOBY, 1994). If one nevertheless takes rational behavior as the reference, bounded rationality can be regarded as a special case of incomplete information since cognitive restrictions preclude an individual from completely recognizing and correctly processing all relevant aspects of the environment or her own preference structure.⁴

The assumption that a decision maker perfectly knows all possible courses of action, all aspects of the consequences of those actions, and the probabilities with which certain actions will yield certain consequences is unrealistic for many real world problems. Obtaining additional information about either unknown feature of the problem may be possible, but will incur information

⁴ The similarities between the approaches of bounded rationality and incomplete information were stressed by CONLISK (1996, pp. 690 f.).

gathering costs. Taking those costs into account, SIMON (1955) proposed that the decision maker may decide to settle for an outcome he finds satisfactory instead of checking all possible alternatives. When new information becomes available in some way, the individual may recognize a superior option and change his behavior. New information can become available through an exogenous shock as well as endogenously through learning effects connected to past decisions. For example, as to experience goods, consumption is necessary to acquire information about product quality (see NELSON, 1970).

Abolishing the assumption of complete information allows for an explanation of changing behavior within a model of stable preferences. Surely, changed consumption behavior, abiding with revealed preference theory, indicates altered preferences about consumption goods, but it can easily be argued in line with new consumer theory that individuals' preferences should be defined with respect to qualities conferred by goods instead and that preferences in that sense remained stable (see also Section 4). Yet, the matter can become much more intricate if one allows for cognitive and psychological processes that blur the division line between preferences and information conditions. The cognitive constrained individual confronted with a complex environment may gather new information in the form of experience only unconsciously and may therefore be unable to decide whether her changed behavior originates from new information or changed preferences. COHEN and AXELROD (1984) suggested that it may be adaptive for such constrained agents to treat new information as new preferences. The opposite may be true for an individual who tries to reduce her cognitive dissonances (see FESTINGER, 1957). As changing preferences may endanger her self-image of a rational, consistently acting person, she may unconsciously ascribe her changed behavior to changed environmental conditions, although changed preferences are the real origin. Thus, she may engage in a selective search for new information that can justify her behavior.⁵

Imitation of others can be another reasonable reaction to incomplete information or constrained cognitive capacity (see CONLISK, 1980). An extreme form of imitation is herd behavior in which individuals may completely ignore their private information and follow the mass. BANERJEE (1992) and BIKHCHANDANI, HIRSHLEIFER, and WELCH (1992) constructed models of rational herd behavior which pay special attention to the fragility of the phenomenon. In their models, only slight new information becoming available can cause herd behavior to start

⁵ Classical dissonance theory assumes that dissonance reduction can occur only after a binding decision has been made. Yet, more recent research suggests that individuals process information in a biased manner even before they decide in order to be able to justify their choice more easily *ex post*. See BROWNSTEIN (2003).

or stop. Thus, fads that seem irrational at first sight can be explained as reactions to altering information conditions without resorting to changing preferences. Rational herd behavior as predicted by these models could also be observed in laboratory settings (see ANDERSON and HOLT, 1997; HUNG and PLOTT, 2001).

Incomplete information need not necessarily refer to environmental conditions, but an individual can also be incompletely informed about her own preferences. If cognitive capacity is scarce, it is to be expected that keeping a consistent or even complete preference ordering for all decision problems that may possibly arise entails significant storage costs. Weighing costs and expected benefits will then probably lead humans to hold only imperfect preferences. MARCH pointed out: “Limitations of memory organization and retrieval and of information capacity affect information processing about preferences just as they affect information processing about consequences. [...] Human beings have unstable, inconsistent, incompletely evoked, and imprecise goals at least in part because human abilities limit preference orderliness” (1978, p. 598).

With such restrictions, learning effects concerning an individual’s own preferences are possible. It is unclear whether one should speak about preference change when such learning effects are realized. SIMON stated that the “consequences that the organism experiences may change the pay-off function” (1955, p. 113); he thus treated the issue within a framework of changing preferences. Alternatively, one may claim that each man owns a stable preference structure *a priori*, but his preferences are first unknown to himself and have therefore to be discovered by experience. MIRRELES took this standpoint when he asserted: “It should be emphasised that uncertainty about one’s tastes, and consequent openness to suggestion, [...] is not evidence that firm preferences are absent. One does not know what visiting the Taj Mahal is going to be like: but, when one is there, uncertainty about tastes is much diminished” (1982, p. 69). This position is now known as the “discovered preference hypothesis” due to PLOTT (1996). The proponents of the hypothesis argue that individual preferences will in the end appear stable and in accordance with standard theory, i. e. obey rationality axioms, if people are provided with enough incentives, feedback, and opportunities for deliberation and learning. Experimental evidence for the issue is mixed. While variations of the experimental design led to fewer violations in some studies (see COX and GREYER, 1996; HOFFLER and ARIELY, 1999; VAN DE KUILEN, 2009; VAN DE KUILEN and WAKKER, 2006; WEST, BROWN, and HOCH, 1996), violations persisted in other experiments although the tasks were simple or opportunity for further deliberation was allowed for (see CUBITT, STARMER, and SUGDEN, 2001; SLOVIC and TVERSKY, 1974). Note that without

any specification how much feedback, incentives or learning opportunities exactly are needed for true preferences to be discovered, the discovered preference hypothesis cannot be falsified empirically.

Another approach assumes that preferences do not exist *a priori*, but are, at least in part, constructed only when they are needed for decision-making (see BETTMAN, LUCE, and PAYNE, 1998; PAYNE, BETTMAN, and JOHNSON, 1992; SLOVIC, 1995). Inspection from the viewpoint of bounded rationality suggests that holding fixed preferences will pay off for choice problems an individual is frequently confronted with whereas case-by-case construction of preferences will be less costly for uncommon problems. As far as preferences are constructed, it should not be claimed that they are stable. One might indeed be tempted to expect that also preferences constructed at different points in time should be the same, provided that the underlying decision problems are identical. However, the *ceteris paribus* condition is never perfectly fulfilled in reality. According to PAYNE, BETTMAN, and JOHNSON (1992), the process of preference construction does not follow an invariant algorithm. Rather, the individual is equipped with a repertoire of preference formation techniques, and the decisions how to access information and which method of preference formation to use in a certain situation highly depend on context effects. The concept of evolutionary psychology mentioned above corroborates this view since it also proposes that it is adaptive to employ context-dependent reasoning. The approach of constructed preferences is able to explain the broad experimental evidence of framing effects and preference reversal phenomena (see e.g. LICHTENSTEIN and SLOVIC, 1971; TVERSKY and KAHNEMAN, 1981). However, allowing preferences or the process of preference construction to depend on inconspicuous aspects of the choice environment entails the danger of arbitrariness, which is discussed in Section 10.

4 De gustibus non est disputandum

In their seminal paper *De gustibus non est disputandum* STIGLER and BECKER (1977) stuck up for the perpetuation of the preference stability assumption. They exemplarily showed how phenomena of *prima facie* preference instability like fads or addictive behavior can be explained within a stable preference framework. For this purpose, they deployed new consumer theory as described by LANCASTER (1966) (see also BECKER, 1965, 1974). According to new consumer theory, households do not directly derive utility from the consumption of goods purchased on markets. Rather, utility U is derived from basic characteristics and

qualities Z_j reminiscent of Bentham's causes of pleasures and pains. These commodities, in turn, are produced by the household using market goods x_{kj} , time t_j , and human capital H as inputs. Formally,

$$U = U(Z_1, \dots, Z_n), \quad (1)$$

where

$$Z_j = f_j(x_{1j}, \dots, x_{mj}, t_j, H). \quad (2)$$

The resources of a household thus are not only market goods, but also time and a human capital endowment. The latter two are priced with the individual wage rate w and a shadow price p_H that reflects the costs of a marginal increase. BECKER (1974) termed this total budget a household's "social income" S ,

$$S = \sum_k p_k x_k + wT + p_H H, \quad (3)$$

where p_k , x_k , and T denote the market price of good k , the quantity of good k owned by the household, and total time available to it. While the time endowment is fixed, investment in additional human capital is possible. Optimization thus requires that the price-weighted indirect marginal utilities of all market goods as well as human capital be equated. New consumer theory entails some advantages compared to standard theory. For example, the inclusion of basic characteristics allows for predictions whether two goods are complements or substitutes.

STIGLER and BECKER used the concept of new consumer theory to rationalize *prima facie* inconsistent behavior while they maintained the stable preference assumption. Assuming that the consumption of addictive goods changes the level of human capital, addictions can be categorized into being beneficial or harmful depending on whether they increase or decrease human capital. For beneficial addictions, increasing consumption over time can be explained by a falling shadow price for producing the underlying characteristic as human capital grows. As for harmful addictions, the shadow price increases over time, but this does not necessarily imply a reduced consumption of the addictive good. On the contrary, if the individual's demand for the underlying characteristic is relatively inelastic, increasing consumption is needed to substitute human capital in the production process. Customary behavior and traditions are explained by a stock of specialized human capital people have built up over time. Abiding by traditions thus allows for cheaper production of utility-creating characteristics. To rationalize fashions and fads, Stigler and BECKER again used new consumer theory. If distinction is a utility-producing characteristic, the consumption of a market good entails an external effect on

other consumers since the ability of the good to produce distinction is diminished. The evolving dynamics can lead to fads for market goods while the preferences for distinction remain stable, which was also shown by KARNI and SCHMEIDLER (1990); PESENDORFER (1995) and POSTLEWAITE (1998) took this idea one step further when they endogenized the demand for distinction and social status being signaling devices in matching games or instruments for success on imperfect markets.

STIGLER and BECKER did not provide a proof for their claim of stable preferences. They stated that, as individual preferences are in the last instance an empirical matter, such a proof is impossible. It is therefore open to discussion whether modeling preferences as stable or unstable is more reasonable. Some criticism of the STIGLER-BECKER approach is concerned with its implication of “rational addiction”. Since individuals have stable preferences and rationally maximize their utility by their actions, addicts should be happy with their addiction. Everyday experience, however, WINSTON (1980); SCHELLING (1984) and AKERLOF (1991) argued, suggests that they are caught in a permanent conflict between their short-term and long-term interests. In their view, the willingness of some addicts to be hospitalized in rehabilitation centers indicates that they are dissatisfied with their situation. BECKER (1992) answered this critique by referring to decision-making under uncertainty. While addicts may indeed be unhappy *ex post*, they still could have rationally maximized their expected utility *ex ante*, not knowing by how much the consumption of a further unit of an addictive good would intensify their addiction. Further criticism pertained to the arbitrariness of the Stigler-Becker approach. While explaining changing behavior with unstable exogenous preferences is of course arbitrary, the concept of STIGLER and BECKER may be no better in this respect. COWEN stated: “Asserting a change in tastes possesses a certain degree of arbitrariness, but the changes in the household production function [...] may be just as arbitrary. Postulating that music listening changes the ability of music to produce relaxation does not appear radically different from postulating that music listening changes the taste for music” (1989, p. 129). To overcome this arbitrariness, a structured model is needed that explains which factors influence preferences or human capital in which ways. This survey discusses several of those factors. In its pure form, the STIGLER-BECKER approach does not provide any such structure. Therefore, “the levels of human capital, the functional form of the production functions, and, indeed, the nature of generalized consumption goods can be freely selected by the modeler” (SOBEL, 2005, p. 401). This renders the model practically irrefutable, which means that it does not constitute any progress compared to claiming that preferences can change arbitrarily.

5 Adaptive preferences and ethical issues

In models by NG and WANG (2001) and WELSCH (2005), economic agents are able to adapt their preferences strategically towards environmental conditions. The aim of preference modification is utility maximization with new preferences. However, adaptation costs incurred by preference change have to be taken into account. Although the problem is inherently dynamic, both models are static. Following the model of NG and WANG, the individual's utility U is gross utility V minus preference adaptation costs K , where the former depends on the quantity of goods consumed x and an individual preference parameter a , and the latter can be assumed increasing in the extent of preference change compared to a former point in time Δa , i. e.

$$U = V(x, a) - K(\Delta a). \quad (4)$$

Obviously, the marginal positive effect of preference change on gross utility has to equal marginal adaptation costs in optimum.

How can such a model be interpreted? Preference adaptation allows the individual to increase her utility by putting more value on goods that are relatively cheap or in another way easy to achieve. NG and WANG suggested an interpretation of the preference parameter as degree of materialism in one's conviction. Relatively rich people might tend to become more materialistic than the poor through preference adaptation. In a similar vein, OXOBY (2003, 2004) modeled an endogenous process of social stratification that originates from adaptive preferences. Surely, the assumption that people can freely and consciously manipulate their preferences to enhance their utility level is unrealistic. Nevertheless, unconscious psychological factors may be at work that modify preferences, at least within a certain range,⁶ as if they were manipulated strategically. Cognitive dissonance theory (see ARONSON, 1997; FESTINGER, 1957) predicts such adaptation processes. Changing one's preferences may often prove the easiest way to accommodate one's self-image of an intelligent and successful person with reality. Several studies are concerned with the consequences of endogenous preferences of this sort for economics.⁷ AKERLOF and

⁶ ELSTER (2004) criticized that economic models of preference adaptation contain only costs, but no constraints on preference change. He insisted that the individual cannot modify her preferences beyond certain boundaries determined by her initial preferences and her information.

⁷ Alternatively to this view of malleable preferences, the aim of maintaining a positive self-image can be incorporated into economic analysis with exogenous preferences. Positive self-image then becomes just another good the individual values and trades off against other goods. See BREKKE, KVERNDOKK, and NYBORG (2003).

DICKENS (1982) showed that with cognitive dissonance state intervention can enhance welfare by establishing safety regulations. RABIN (1994) modeled agents who hold moral preferences. If an activity is considered immoral, agents can react to this by reducing the respective activity or adopting the (false) belief that the activity were in fact acceptable. Both acting against his beliefs and holding false beliefs incurs costs to the agent. KONOW (2000) applied a similar model to explain behavior in modified dictator games. GAO and SCHMIDT (2005) explained irrational behavior by agents' intention to maintain self-value. HAHNEL and ALBERT (1990) investigated a general equilibrium setting. A Pareto-efficient general equilibrium still exists with endogenous preferences if all agents are perfectly informed about the process of preference formation. However, the equilibrium is unstable since if a distortion in the price system occurs, people adapt their preferences towards distorted prices, thereby aggravating the distortion.

Adaptability of preferences raises ethical problems. If preferences are stable, they are clearly defined. If one likes to see preferences that originate from different information conditions as changing, it still seems clear that the preferences developed when information is complete should be the benchmark.⁸ But, with malleable attitudes as considered in this section, which preferences shall be regarded as "true" or ethically justified? Is the anti-materialistic conviction of a poor true if it depends on him being poor and could be abandoned by giving him enough money? ELSTER (1982) warned that unconscious preference adaptation endangers freedom of action as people are seduced to resign themselves to goals which can easily be attained. Ethical judgment is problematic since there are two opposing effects of increased utility on the one hand and decreased autonomy on the other. SEN (1977) suggested that preferences of higher order (i. e. a preference ordering of one's different preference orderings) be needed to express ethical judgments. The idea of such meta-preferences is not alien to philosophers (see e. g. JEFFREY, 1974); FRANKFURT (1971) even considered them essential for freedom of will and thus the concept of a person. A much more relaxed and probably naive view towards preference adaptation and even preference manipulation by the state was taken by HARSANYI:

The modern methods of propaganda and education offer a powerful tool for producing substantial changes in people's tastes at relatively low costs. For this reason, to-day the *Economic Problem* of a community [...] does not consist merely in finding the best uses for its scarce resources so as to gratify people's actual wants to the highest possible degree. Rather, it includes also the question of how these scarce resources should be divided

8 This view was taken by e. g. HARSANYI (1982, p. 55) and ELSTER (1982, pp. 235 f.).

between productive operations for *satisfying* people's actual wants and measures for *changing* these wants. It must be decided again and again whether a given amount of resources benefits the consumers more if it is used for gratifying their present tastes [...] or if it is used for educating these consumer's tastes, for altering their irrational likes and dislikes [...]. (1954, p. 213)

HARSANYI's standpoint gave rise to heavy criticism, particularly by Marxist authors, that the floodgates were opened to manipulation of individual preferences by group interests. This criticism is dealt with in Section 9.

6 Habit formation

The term habit formation refers to an individual's past or anticipated future consumption influencing her present utility. The concept stands in contrast to standard theory which assumes utility functions to be time-separable, going back to an elegant model of intertemporal utility by SAMUELSON (1937).⁹ Yet, its elegance and simplicity rather than its realism probably were the cause for this model to become a part of standard theory. Samuelson himself pointed to its limitations, and HICKS stated that "[i]t is nonsense to assume that successive consumptions are independent" (1965, p. 261). The meaning of the time-separability assumption is illustrated by FREDERICK, LOEWENSTEIN and O'DONOGHUE, who stated that "consumption independence says that a person's preference between an Italian and Thai restaurant tonight should not depend on whether she had Italian last night, nor whether she expects to have it tomorrow" (FREDERICK, LOEWENSTEIN, and O'DONOGHUE, 2002, p. 357).

In a general model of habit formation, an individual's utility U_t in period t depends on her consumptions x_s in the present as well as in past and future periods, i. e.

$$U_t = U_t(x_{t-n}, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_{t+m}). \quad (5)$$

A simpler form of the habit formation model disregards effects of anticipated future consumption and collects the effects of past consumption in a state variable, the so-called habit stock h_t , which is often defined as a discounted sum of past consumption levels. Hence,

$$U_t = U_t(x_t, h_t), \quad (6)$$

⁹ An extensive discussion of this model can be found in FREDERICK, LOEWENSTEIN and O'DONOGHUE (2002).

with

$$h_t = \sum_{j=1}^n \beta^j x_{t-j}, 0 < \beta < 1. \quad (7)$$

Different specifications of equation (6) exist. Utility may depend on the difference between current consumption and habit stock (for an overview, see POLLAK, 1970). In this case, the habit stock can be interpreted as a required minimum consumption level. Another specification treats the ratio of current consumption and habit stock as argument of the utility function (see e. g. ABEL, 1990), thus attenuating the effect of past consumption on present utility. In any case, an increase in the habit stock, by raising the consumption level an individual physically or psychologically needs or is accustomed to, decreases the utility derived from a given amount of present consumption, i. e. $\partial U_t / \partial h_t < 0$.

If consumers are able to anticipate the long-run detrimental effect of increased consumption on utility due to habit formation, an inert reaction to income increases is to be expected. SUNDARESAN (1989) found that a model which allows for habit formation can produce much more smooth consumption paths than can be explained by the permanent income hypothesis. CAMPBELL and DEATON (1989) analyzed US consumption data and detected both an excess smoothness of consumption to unanticipated income changes and an excess sensitivity to anticipated but not yet realized changes. Both phenomena, the authors argued, could be explained by a habit formation model. There is by now, all in all, extensive empirical evidence in support of the existence of habit effects in consumption data. Early evidence stems from HOUTHAKKER's and TAYLOR's (1966) projection of US consumer demand. They found habit effects in the past data for the majority of commodities they considered. Some commodities showed an opposite stock effect, which means that increased consumption in the recent past leads to less demand in the present. KAPTEYN, WANSBEEK, and BUYZE (1978) observed habit effects in Dutch holiday expenditure data. A number of more recent studies used household level data (often panel data), thereby allowing for a more rigorous test of habit formation as the data aggregation could have led to spurious correlation.¹⁰ Highly significant habit effects were found in data from a range of different countries (see CARRASCO, LABEAGA, and LÓPEZ-SALIDO, 2005; GUARIGLIA and ROSSI, 2002; HEIEN and DURHAM, 1991; NAIK and MOORE, 1996). In some other studies, however, habit effects could not be observed (see DYNAN, 2000; MEGHIR and WEBER, 1996).

¹⁰ See e. g. HEIEN and DURHAM (1991) and NAIK and MOORE (1996) for a discussion of methodological issues.

Habit formation can also help explain an array of other empirical phenomena the most prominent of which may be the equity premium puzzle. CONSTANTINIDES (1990) suggested that a habit stock of circa 80 percent of the current consumption level in connection with plausible values for the coefficient of risk aversion could rationalize that investors are willing to accept so much lower returns for secure forms of investments. This result was corroborated by further analyses of capital market data (see ABEL, 1990; BRAUN, CONSTANTINIDES, and FERSON, 1993; CAMPBELL and COCHRANE, 1999; FERSON and CONSTANTINIDES, 1991). Other phenomena which can be dealt with by habit formation models comprise life-cycle portfolio choice (see POLKOVNICHENKO, 2007), over-insurance and propensity for low deductibles (see BEN-ARAB, BRIYS, and SCHLESINGER, 1996), responses to changes in monetary policy (see FUHRER, 2000) and macroeconomic shocks (see RAVN, SCHMITT-GROHÉ, and URIBE, 2006), evidence that growth causes savings (see CARROLL, OVERLAND, and WEIL, 2000), and intertemporal choice phenomena (see WATHIEU, 1997).

Do individuals whose utility functions incorporate habit formation deal with their preferences rationally or myopically? A long-lasting and still unresolved debate has been concerned with this issue. In its early stage, POLLAK considered it a “fundamental assumption of the habit-formation model [...] that the individual does not take account of the effect of his current purchase on his future preferences and future consumption” (1970, p.761). Yet, most of the models mentioned in the above two paragraphs, in order to reproduce the empirically observable phenomena, employ agents who act rationally and anticipate their preference change. For models of optimal growth under habit formation as by RYDER and HEAL (1973) and BOYER (1978) to make sense, rationality at least on the level of a social planner has to be assumed.¹¹ Further theoretical models with rational habit formation include LLUCH (1974); BOYER (1983), and IANNACCONE (1986); BECKER and MURPHY (1988) aimed at rationalizing addictive behavior in their model of “rational addiction”. They employed the formal approach of STIGLER and BECKER presented in Section 4, but that approach, as SPINNEWYN (1981) showed, is equivalent to a model of rational habit formation. Any model of rational habit formation can be transformed into a model with time-separable utility by redefining prices and initial wealth, such that a decrease of initial wealth due to the inherited habit stock and an additional shadow price of consumption due to the habit effect are taken into account.

¹¹ In fact, BOYER (1978) explicitly assumed rationality also on the individual level.

A decreasing share of smokers in the US population in the years after evidence about the detrimental health effects of smoking became widely public was interpreted by BECKER and MURPHY (1988) as evidence that individuals were indeed able to react rationally to new information even if addictive behavior is concerned. PASHARDES (1986) and MUELLBAUER (1988) discussed methods to distinguish rational and myopic habit formation in empirical data. While PASHARDES found evidence for rational behavior in British consumption data, MUELLBAUER's US data point to myopia. Further empirical support for the hypothesis of myopic habit formation has come from other disciplines including psychology, medical science, and happiness research. Reviews are provided by RABIN (1998) and LOEWENSTEIN, O'DONOGHUE, and RABIN (2003). The evidence presented there shows that in various situations individuals are able to foresee the direction in which their preferences will change, but systematically and sometimes tremendously underestimate the speed or magnitude of the change. For example, experimental studies by KAHNEMAN and SNELL (1992) and LOEWENSTEIN and ADLER (1995) found that subjects were unable to anticipate their change of preference for certain food and music following from extensive consumption as well as their changing valuation for a coffee mug due to an endowment effect. A survey by SNELL, GIBBS, and VAREY (1995) showed that people are able to anticipate some psychological effects of preference modification, but do not recognize other mechanisms which are also widely agreed upon in the psychology literature.

The question whether people can rationally deal with their habits has also given rise to experimental investigations. In a laboratory environment, preferences with habit formation effects can be induced via the payoff function, which entails the advantages that those preferences are consistent and fully known to both the subjects and the experimenters. Experimental evidence about the rationality of habit formation is also ambiguous. FEHR and ZYCH (1998) found that their subjects in several successive experimental life cycles systematically overconsumed in early periods without any indication of learning although there was no uncertainty about their income stream and they had been provided with extensive training before the experiment. By contrast, BROWN, CHUA, and CAMERER (2009) reported of subjects who quickly learned in a more complex setting with income uncertainty, such that their behavior did not significantly deviate from the optimal path after a few life cycles. Moreover, when their subjects were provided with an example of a successful intertemporal consumption and saving pattern, the observed behavior was close to optimal right from the beginning. Results from experiments of intertemporal consumption and saving decisions without habit formation effects point at heterogeneous behavior, as well (see e. g.

BALLINGER ET AL., 2011). Only a part of the subjects was able to deal with that task successfully.

Assuming myopic habit formation, a series of early studies has been concerned with the long-run behavior of agents (see EL-SAFTY, 1976a, 1976b; GORMAN, 1967; HAMMOND, 1976; MCCARTHY, 1974; POLLAK, 1970, 1976b). Under certain conditions, intertemporal consumption patterns lead to an equilibrium in which demanded quantities remain stable. Even if such an equilibrium exists and is reached, it is based upon myopic behavior and does not constitute a case of intertemporal utility maximization. Only under restrictive conditions can the behavior be rationalized with a consistent long-run utility function. A further point of discussion has been welfare analysis under myopic habit formation. VON WEIZSÄCKER (1971) suggested that the long-run equilibrium preferences be taken as basis for welfare considerations. He justified this by proving that if a long-run utility function exists and this function assigns a higher value to consumption bundle A than to B, it is possible for the individual to move from B to A in a finite series of steps and with every step perceive an improvement according to her respective current preferences. A more restrictive welfare criterion was put forward by WEISBROD (1977). This is a kind of Pareto criterion. According to it, A can only be regarded superior to B if A is preferred under both preference structures that prevail if the individual initially consumes either A or B. POLLAK, in contrast, criticized all approaches to judge individual behavior from a long-run or intertemporal welfare perspective, though the individual herself is unable to make such a judgment due to her myopia, as “schizophrenic” (1976b, p. 296). In a similar vein, MARSCHAK (1978) pointed out that if one treats the individual with different preferences following from different consumption patterns or state policies as different persons and requires unanimity, truly taste changing actions are rendered impossible.

7 Interdependent preferences

Economists’ considerations of interdependent preferences¹² are as old as the science of economics itself. In *The Theory of Moral Sentiments* SMITH wrote:

[W]hat are the advantages which we propose by that great purpose of human life which we call bettering our condition? To be observed, to be attended to, to be taken notice of with

¹² While we stick to the term “interdependent preferences” in this survey, the synonymous terms “social preferences” and “other-regarding preferences” are frequently used in the literature.

sympathy, complacency, and approbation, are all the advantages which we can propose to derive from it. It is the vanity, not the ease, or the pleasure, which interests us. (1976, p. 50)

MARX unambiguously stated that “wants and pleasures have their origin in society; we therefore measure them in relation to society; we do not measure them in relation to the objects which serve for their gratification. Since they are of a social nature, they are of a relative nature.” (1902, p. 42). Around 1900, VEBLEN (1934) described the behavior of a well-to-do “leisure class” that demonstrates its wealth by conspicuous leisure and conspicuous consumption. At about the same time, PIGOU wrote that “people do, in fact, desire many things, not merely for their own sake, but, in the main, on account of the reputation or distinction which the possession of them confers” (1913, p. 20).¹³

Very generally, interdependent preferences can be described by a utility function U of individual i that not only contains her own consumption x^i , but also the consumptions of other people x^j as arguments, i. e.

$$U = U(x^i, x^j), i \neq j, \quad (8)$$

where x^j is to be interpreted as a vector in case there are several relevant others who influence i 's utility.¹⁴ The sign of $\partial U / \partial x^k$ is in general indeterminate *a priori* and depends on whether i looks on k with sympathy and altruistic feelings or with hatred and envy instead.

Without further structure, the utility function (8) does not seem parsimonious enough for practical applications as almost any behavior can follow from it. This might explain the negligence with which interdependent preferences have been treated by standard economics. Yet, special cases of the utility function (8) have been proposed that promise to be viable. One approach parallels the habit formation model (see GAERTNER, 1974; POLLAK, 1976a). Like in equation (6), utility additionally depends on a state variable which is this time made up of discounted past consumptions by other people. Another advance is based on the idea that each individual wants to receive a fair share of some total either due to envious feelings or following from a desire to maintain self-esteem or, in sociological terminology, to avoid a feeling of “relative deprivation”

¹³ A short overview over the history of interdependent preferences in economics can be found in CULLISON (1978).

¹⁴ Even more generally, one might consider the case of interrelated utility functions with A's utility level entering into B's utility function and *vice versa*. This might lead to an infinite regress problem. Besides their better traceability, utility functions defined over allocations of goods appear to be more useful for economic applications. BERGSTROM (1999) analyzed under which conditions a system of interrelated utility functions induces unique utility functions over commodity allocations as described in equation (8).

(see e. g. DAVIS, 1959; RUNCIMAN, 1966). Thus, the share of i 's consumption is incorporated as an argument in i 's utility function,

$$U = U\left(x^i, \frac{x^i}{x^i + \sum_{j \neq i} x^j}\right). \quad (9)$$

The underlying psychological rationale requires that the partial derivative of equation (9) with respect to its second argument be negative. A utility function of this type was proposed by DUESENBERY (1949) and can be found in the social preference model by BOLTON and OCKENFELS (2000). A similar psychological reasoning leads to a formulation suggested e. g. by FRANK (1985b) in which utility is rank-dependent, i. e.

$$U = U(x^i, r^i), \quad (10)$$

where r^i denotes the rank of i in a distribution of income or of consumption of certain goods, and a higher rank confers additional utility to i . SCOTT (1972) reckoned that i 's feelings towards k (or, formally speaking, the sign and magnitude of the partial derivative of equation (8) with respect to x^k) depend on k 's consumption level compared to i 's. Namely, i feels altruism if k 's consumption is low compared to i 's, but avarice if it is high. Both feelings express i 's discontent with the respective situation; thus, for a given consumption level x^i , i 's utility attains a maximum for some intermediate x^k that represents to i an equitable outcome. FEHR and SCHMIDT (1999) developed a social preference model with a special form of inequity aversion in which the absolute value of consumption differences affects i 's utility:

$$U = U(x^i, |x^j - x^i|). \quad (11)$$

Empirical and experimental studies have been concerned with a pile of phenomena that may be explained by interdependent preferences. DUESENBERY (1949) long ago recognized that savings data were inconsistent with standard theory. While higher incomes yielded the expected effect of a higher savings rate in cross-sectional data, the effect was absent in time series data. According to DUESENBERY's explanation, when average income in a society increases, consumers more often observe others consuming superior products, which instills in them the desire for those products. Hence, an additional effect increases consumption and decreases savings when average income rises. KAPTEYN, WANSBEEK, and BUYZE (1978) and KAPTEYN ET AL. (1997) found that allowing for a habit stock formed by other people's past consumption considerably improves the fit of Dutch consumption data. An experiment about life cycle consumption and saving choices by CARBONE and DUFFY (2014) found that overconsumption

compared to the optimal path was larger in a treatment where subjects were provided with information about other subjects' past consumption choices.

Another finding from empirical studies is that others' behavior is of different importance to consumption in different commodity groups. This can be explained by varying conspicuousness of consumption. For example, it is to be expected that expenditure for housing is influenced by status concerns more strongly than expenditure for food since one's type of accommodation is easily observable while food is mostly consumed in private. This idea is corroborated by CHAO and SCHOR (1998), who had cosmetics ordered by their conspicuousness. For more conspicuous products, the correlation between price and quality as assessed by consumer reports was weaker, hinting at status consumption.

Apart from consumption behavior, there are other economic decisions that have been argued to be influenced by interdependent preferences. FRANK (1984a, 1984b, 1985a) claimed that wages diverge less than marginal products of labor because in enterprises there is an implicit market for status in which employees with high status are "bribing the lowest-ranked members of their groups to provide ego support through their continued presence" (1985a, p.53). BOWLES and PARK (2005) explained longer work hours in societies with more unequal income distributions as a result of the desire of the poor to emulate the consumption patterns of the rich.

Further evidence for interdependent preferences has been provided by happiness research. EASTERLIN (1974) was the first to note that happiness levels stated in surveys did not increase with income over time or across different societies, though a positive relationship between income and happiness was evident in cross-sectional data collected in any particular society at any particular point in time. This *prima facie* paradox can easily be explained if individual utility is decreased by increasing incomes of other people. By now, a pile of empirical studies has confirmed this pattern (for a survey see CLARK, FRIJTERS, and SHIELDS, 2008). When subjects in experiments are asked to choose between hypothetical states of the world with differing own and average income in the society, they are willing to forego purchasing power in order to avoid earning below average income (see ALPIZAR, CARLSSON, and JOHANSSON-STENMAN, 2005; SOLNICK and HEMENWAY, 1998). In an experiment by ZIZZO and OSWALD (2001), a substantial share of the subjects was ready to pay to reduce other subjects' payoffs. The results of a hypothetical choice experiment by LOEWENSTEIN, THOMPSON, and BAZERMAN (1989), in contrast, point to individual preferences that incorporate inequity aversion. Popular critiques of consumerism like SCITOVSKY (1992) and HIRSCH (1995) also rely in their argumentation on preference interdependence. Further results that challenge standard theory have been produced in

experimental tests of game theory such as the ultimatum and dictator game (see ENGEL, 2011; GÜTH, 1995; ROTH, 1995 for surveys). The social preference models by FEHR and SCHMIDT (1999) and BOLTON and OCKENFELS (2000) are reactions to this experimental evidence.

Notwithstanding all this evidence, it can be argued that individuals do not really hold social preferences, but use status as an instrument to attain greater monetary success in imperfect markets (see COLE, MAILATH, and POSTLEWAITE, 1992; POSTLEWAITE, 1998). Some confirmatory evidence for this hypothesis was generated by BALL ET AL. (2001), who found in an experiment that individuals with high status were able to influence market outcomes to their advantage. Nevertheless, even if the hypothesis is true, it might still be convenient to integrate preference interdependence into utility functions to describe actual behavior, instead of treating the instrumentality explicitly. BÉNABOU and TIROLE (2006) assumed that prosocial behavior can be motivated by both altruism and status concerns. Rewarding such behavior may then entail a crowding out effect since the true motive of the agent becomes doubtful. Similarly, HOLLÄNDER (1990) found that state intervention for the provision of a public good can induce less provision by agents who voluntarily provide since they derive utility from the accompanying social approval.

If individuals' preferences exhibit negative interdependence, i. e. envy or malice, conspicuous consumption of status goods imposes negative externalities on others. Society is then caught in a prisoners' dilemma since it would be Pareto-improving, though individually irrational, to collectively reduce consumption of status goods (see e. g. FRANK, 1985b, 1989, 2005). These considerations raise the question whether state intervention can augment welfare. A strand of literature is concerned with optimal taxation given externalities from interdependent preferences (see e. g. ABEL, 2005; FRANK, 2008; IRELAND, 2001).

Just as with the effects of habit formation, one can ask the question where the process of mutual adaptation to others' consumption behavior will take consumers in the long run. Explorative studies on this issue were provided by KRELLE (1972) and GAERTNER (1974); GRANOVETTER and SOONG (1986) and RAUSCHER (1992, 1993) showed that a stable equilibrium consumption vector may exist, but, depending on parameter values, also complex dynamic patterns may be reasonable which are indistinguishable from chaos for an outside observer. IANNACCONI (1989), on the other hand, criticized the information assumptions of these models. If one assumes consumers to have adaptive expectations, the likelihood of instability diminishes.

While the evidence presented in this section impressively shows that people's utility often depends on feelings such as altruism, envy, or a sense of

justice, it is unclear to what degree this means that preferences are unstable. Of course, preferences will appear unstable from the viewpoint of the narrow standard model which ignores the interdependence. However, interdependent preferences may also entail inherent preference instability if the described feelings are intrinsically changeable. A strand of literature concerned with this question yields ambiguous results.

ANDREONI and MILLER (2002) in an experiment varied the price of giving to others. They were able to show that individual preferences for altruism largely obey the axioms following from revealed preference theory and can thus be rationalized by a stable utility function that incorporates as an argument the payoffs of others. FISMAN, KARIV, and MARKOVITS (2007) reported similar results in an extended investigation. BLANCO, ENGELMANN, and NORMANN (2011) and DE OLIVIERA, ECKEL, and CROSON (2012) conducted series of different experiments with their subjects to investigate the consistency of interdependent preferences across different contexts. Both studies found largely consistent preferences, at least at the aggregate level. COX, FRIEDMAN, and SADIRAJ (2008) incorporated interdependent preferences into revealed preference theory. For this purpose, they defined orderings over preferences and over opportunity sets in a two-person setting. First consider the ordering of two preference sets a and b that may be held by individual i . If, for every possible income distribution between i and j , i , when holding preferences a , is willing to give up at least as much of her own income for one additional unit of j 's income as under preferences b , then i 's preferences a are said to be more altruistic than her preferences b . As for opportunity sets, i perceives opportunity set F as more generous than set G if i 's maximum attainable income is higher in F than in G and the increase in i 's maximum attainable income, compared to G , is at least as high as the increase in j 's maximum attainable income. The central axiom by COX, FRIEDMAN and SADIRAJ describes reciprocity. It states that, in a sequential game where the first mover through her choice determines the second mover's opportunity set, i 's choice of a more generous opportunity set for j induces more altruistic preferences in j . Like in standard revealed preference theory, if one assumes preferences to be stable, a choice in one opportunity set reveals information about preferences and thus predicts ranges of choices compatible with these preferences in other opportunity sets.

Whereas the last paragraph illustrates that consistency of an individual's interdependent preferences in a given point in time is supported by substantial evidence, the issue of stability across time is less clear. BROSIG, RIECHMANN, and WEIMANN (2007) conducted three identical waves of dictator and prisoners' dilemma games with the same subjects and time intervals of one month between each two waves. They observed that subjects became more selfish over time.

The only subjects behaving consistently across all three waves were those who acted completely selfishly as the standard model predicts. By contrast, VOLK, THÖNI, and RUIGROK (2012) found largely stable preferences for conditional cooperation in their subjects when conducting three waves of public good games over five months. SASS and WEIMANN (2012) argued that this result was confounded by a reputation effect. In a public good experiment with four waves in one week intervals and a design that excluded reputation effects, they again observed subjects becoming more selfish. While conditional cooperation was the most frequent behavior in the first wave, about one third of the cooperators turned into selfish free riders until the last wave. BRUHIN, FEHR, and SCHUNK (2016) used data from dictator and reciprocity games to elicit interdependent preferences and then categorized their subjects into three types which emerge endogenously by employing a finite mixture model. Since they conducted another session with the same subjects five months later, their data are also suited to test preference stability. BRUHIN, FEHR and SCHUNK found remarkable preference stability on the type level. Most of their subjects were either moderately or strongly altruistic. Their model re-assigns a subject to the same type as in the first session in 56–84 percent of the instances, depending on the assigned type in the first session. CARLSSON, JOHANSSON-STENMAN, and NAM (2014) observed substantial stability of cooperation preferences over several years in villagers from rural Vietnam.

Thus, stable interdependent preferences have been found only in part of the studies. In a series of further economic experiments where unstable preferences and increasing selfishness were observed, a moral self-licensing effect was put forward as an explanation (see SASS and WEIMANN, 2015a, 2015b; SASS, TIMME, and WEIMANN, 2015; SCHMITZ, 2015). The phenomenon of moral self-licensing is well-known in the psychology literature (see MERRITT, EFFRON, and MONIN, 2010). It describes that a person, having committed a good deed, thereby feels entitled to behave immorally, e.g. to act more selfishly, next time. Some studies also investigated the boundaries of the effect. SASS and WEIMANN (2015a) detected that moral self-licensing was context-specific. Being unselfish in one situation thus did not give subjects in that study the feeling to be entitled to be selfish in a different context. SCHMITZ (2015) found that the moral self-licensing effect was weaker when the time interval between the waves of his experiment was longer. He set up a model reminiscent of habit formation in which an individual derives additional utility from a stock of committed good deeds that depreciates over time. If an individual has committed a good deed recently, her stock is still high and she has no need to be unselfish. By contrast, a good deed will increase utility if the last good deed lies sufficiently far in the past and the stock thus has sufficiently depreciated. This approach can potentially explain why

interdependent preferences have been observed as stable in some studies but as unstable in others.

8 Character and emotions

To find an answer to the still unresolved question whether interdependent preferences are inherently unstable, it may be helpful to investigate the underlying factors of such preferences. This section goes one step further as it considers what determines such feelings as altruism and envy and how steady, in turn, those determining factors are.

While the stability of interdependent preferences is still subject to discussion, the considerable heterogeneity of such preferences is well-documented. ANDREONI and MILLER (2002) found that significant fractions of their subjects behaved completely selfishly, treated payoffs to self and others as perfect substitutes, or, respectively, treated them as perfect complements, implying equal payoffs. FISCHBACHER, GÄCHTER, and FEHR (2001), conducting a public good experiment, categorized their subjects according to three behavioral patterns: free riding, conditional cooperation, and hump-shaped contributions that first increase and then decrease again with rising contributions of others. As mentioned already in the last section, BRUHIN, FEHR, and SCHUNK (2016) determined three preference types endogenously: moderate altruists, strong altruists, and negatively inequity averse. CAPLAN (2003) argued that individual preference stability and heterogeneity can be explained with diverse characters of different people. In psychology, the five-factor model of personality may now be considered canonical. According to it, human personality varies on five dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (for an introduction to the model, see e.g. DIGMAN, 1990; MCCRAE and JOHN, 1992). A series of longitudinal studies revealed that personality traits in adults are remarkably stable considering periods of several years or even decades (see e.g. CONLEY, 1985; COSTA and MCCRAE, 1988; SOLDZ and VAILLANT, 1999). (BOONE, DE BRABANDER, and VAN WITTELOOSTUIJN (1999), LÖNNQVIST, VERKASALO, and WALKOWITZ (2011), and VOLK, THÖNI, and RUIGROK (2012) applied these insights to economics. The results of the prisoners' dilemma and public good experiments they conducted support their hypothesis that an individual's inclination to cooperation varies with her personality. Specifically, high openness to experience, high agreeableness, and low neuroticism were related to more cooperative behavior.

BELL (1982) and LOOMES and SUGDEN (1982) proposed regret as another feeling that may influence behavior. According to regret theory, in a choice situation under uncertainty, an individual feels regret when, having made her choice, consequences are revealed and she observes that she would have done better had she decided differently. Note that regret occurs even if the choice maximized standard expected utility *ex ante*. Conversely, if a chosen option comes out to have consequences superior to those of its alternatives, the individual experiences a pleasurable feeling called rejoicing. Regret theory assumes that the individual anticipates her possible regret and rejoicing and accounts for them when making her choice in the first place. That is, she maximizes a modified version of her expected utility that considers the impacts of these feelings. The incorporation of regret makes utility choice set-dependent since the occurrence of regret depends upon the existence and the outcomes of further options. From the viewpoint of standard theory, the utility function becomes intransitive. Therefore, it depends on the definition of alternatives whether a sequence of choices can be interpreted to reveal preference instability. Suppose an individual chooses A from the set {A, B}, but chooses B from the set {A, B, C}. While for standard theory this means a preference reversal, choice set-dependent regret theory does not define A as an alternative. Rather, there are two distinct alternatives in the two situations: “choosing A and rejecting B” in the first and “choosing A and rejecting both B and C” in the second. No inference can be made from the choice in the one situation on the choice in the other. Yet, such a narrow definition of alternatives renders the occurrence of preference reversals in the real world extremely unlikely as the choice set has to be exactly held constant. Unstable preferences may thus simply be defined out of existence. However, regret theory is able to explain several behavioral anomalies such as the Allais paradox and the standard pattern of preference reversal (see also LOOMES and SUGDEN, 1983).

While character traits and the inclination to feel regret are relatively constant factors, emotions can arise and fade rapidly. LOEWENSTEIN (1996, 2000) discussed the effects of negative emotions and drive states, such as anger, fear, hunger, sexual desire, and pain, under the term “visceral factors”. Visceral factors can be incorporated into standard theory by defining utility as state-dependent. When an individual is in a certain drive state, her utility function U switches. More precisely, the marginal utility of commodities connected to her drive greatly increases, so that the desire for those commodities that can satisfy her craving becomes dominant. For instance, a person may be willing to forego considerably more units of another commodity for one unit of food when she is hungry. Formally,

$$U = \begin{cases} U_1(x_a, x) & \text{if } a = 1 \\ U_0(x_a, x) & \text{if } a = 0 \end{cases}, \quad (12)$$

with $\partial U_1 / \partial x_a \gg \partial U_0 / \partial x_a$, where x_a denotes commodities connected to the drive and a is an indicator variable that equals one when the individual is in a drive state.¹⁵ ARIELY and LOEWENSTEIN (2006) showed in an experiment that sexual arousal influences hypothetical decisions about behaviors related to sex. Moreover, LERNER, SMALL, and LOEWENSTEIN (2004) found that emotional effects can also carry over on unrelated and payoff-relevant decisions. Note that visceral states, while they can appear and disappear rapidly, do not occur arbitrarily. Rather, their occurrence depends on the presence and the intensity of certain cues that are in general predictable. For instance, the occurrence of a hunger feeling depends on factors such as the elapsed time since the last meal and the proximity of palatable food.

An axiomatic approach is presented by DIETRICH and LIST (2013). They set up a model in which different motivational states bring about different preference orderings. In their model, only a subset of the relevant properties of the choice alternatives is salient in a certain motivational state and only salient properties affect preferences. Thus, while there is an unique underlying preferences relation, called “weighing relation” by the authors, which ranks different property combinations, that weighing relation defines different preferences in different motivational states with different sets of salient properties. The underlying weighing relation need not be transitive and complete in order to define transitive and complete preferences in each motivational state. DIETRICH and LIST remain silent on the question what causes certain motivational states with certain salient properties to occur.

LAIBSON (2001) deployed a model similar to LOEWENSTEIN’s to explain addictive behavior. In his model, when a commodity is regularly consumed in a certain situation or environment, a complementarity between the commodity and the properties of the environment may develop, so that the marginal utility of consuming the commodity is increased in that environment. In this case, the environment or its properties become a cue for consumption. In the long run, several different equilibria are possible, in one of which the individual consumes the commodity only when the cue is present. For example, there may be an equilibrium in which an individual drinks alcohol only on weekends or only when in companionship. In contrast to LOEWENSTEIN’s approach, the commodity need not draw on the satisfaction of a basic physiological want. Neither need the

¹⁵ LOEWENSTEIN’s (1996) more complex model also allows for multiple drives and continuous drive levels.

reaction to the cue be innate, but can be acquired through habituation. One may thus interpret LAIBSON's approach as a generalization of habit formation theory, where other factors than past consumption can enter the habit stock and bring about a complementarity to present consumption.

LOEWENSTEIN (1996) doubted that visceral factors should be treated as changing preferences, the reason being their connection to irrational behavior. People would often regret *ex post* the decisions they made when they were in drive states, or were aware of the sub-optimality of their actions even in the very moment they act. LOEWENSTEIN furnished this view with examples of people suffering from a phobia who are well aware that there is no reason to be afraid, yet are unable to act according to this insight, or of addicts who realize that taking drugs is against their self-interest, but cannot escape the compulsion to do so. Another instance of irrationality is that people generally fail to fully anticipate how future visceral influences will affect their desires and actions even if they have many times undergone such states in the past (see e. g. READ and VAN LEEUWEN, 1998). Thus, an alternative explanation arises: emotions may not change utility, but inhibit information processing. Emotions may then be an instance of bounded rationality rather than changing preferences. KAUFMAN (1999), referring to the Yerkes-Dodson law in psychology which postulates an optimal level of arousal, related excessive arousal to bounded rationality. LEITH and BAUMEISTER (1996) found in several experiments that subjects who were aroused and in a bad mood were more prone to take foolish risks. Since this tendency vanished when the subjects were advised to thoroughly think through their decision, the authors considered deliberation impaired by emotion to be the most likely explanation. The results of some economic experiments in which subjects had the opportunity to diminish the payoff of other players at their own expense also corroborate this view. Although such behavior can in principle also be explained as a rational reaction to negatively interdependent preferences as discussed in the last section, one would then expect careful trade-offs between the utility derived from harming others and the accompanying costs. Yet, observations point at impulsive, emotion-driven actions. Thus, ZIZZO and OSWALD (2001) found that the demand for destroying other players' payoffs was very price-inelastic and BOSMAN and VAN WINDEN (2002) observed that subjects often chose a corner solution and forewent their whole own payoff in order to hurt the other player as much as possible.

Reciprocal behavior may be regarded as a hybrid of interdependent and emotion-driven preferences. Positive reciprocity, i. e. rewarding behavior, fits well into the interdependent preferences framework. However, negative reciprocity, i. e. punishing behavior, is closer connected to emotions as it is often impulsive. In contrast to the models discussed in Section 7, reciprocity entails

that one's utility U does not only depend on outcomes for oneself x^i and for others x^j , but also on others' intentions a^j ,

$$U = U(x^i, x^j, a^j). \quad (13)$$

When j by her actions determines a vector (x^i, x^j) or a set of such vectors as an opportunity set, i can, depending on the context, perceive the same material outcomes as resulting from either kind or unkind behavior by j . Some experimental evidence for the role of intentions stems from modified ultimatum games in which the proposer could choose between only two allocations. When a proposer chose to propose an unequal division leaving the larger part of the pie to herself, higher responder rejection rates were observed if the proposer's alternative option was an equal division than if the alternative option entailed even more inequality (see BRANDTS and SOLA, 2001; FALK, FEHR, and FISCHBACHER, 2003). Presumably, depending on the feasible alternative, the responders regarded the same division to reveal the proposer's unkindness in the first case, but her kindness in the second. Revealed intentions thus bring about preferences which appear as choice set-dependent if the influence of intentions is neglected. In a similar vein, rejection rates differed, depending on whether a division was intentionally chosen by the proposer or it was determined randomly and the responders knew about that (see BLOUNT, 1995). The role of intentions in related settings was investigated by COX (2004) and FALK, FEHR, and FISCHBACHER (2008).¹⁶ Further experimental evidence for positive reciprocity was produced by BERG, DICKHAUT, and MCKABE (1995) and CROSON (2007) whereas behavior in accordance with negative reciprocity was observed by FEHR and GÄCHTER (2000) and KAGEL and WOLFE (2001).

Since intentions cannot be directly observed, beliefs about intentions based on observable behavior have to be formed. That is, the intentions variable a^j in expression (13), in turn, is a function of beliefs. GEANAKOPOLOS, PEARCE, and STACCHETTI (1989) analyzed which consequences the inclusion of beliefs into the utility function brings about for game theory. While backward induction can no longer be applied, they proved existence theorems for normal-form Nash equilibrium and subgame-perfect equilibrium. However, there can be multiple subgame-perfect equilibria even if there are no payoff ties. Since GEANAKOPOLOS, PEARCE and STACCHETTI assume only initial beliefs to be relevant and neglect rational belief updating, the application of their concept to multi-stage games may yield

¹⁶ Cox (2004) conducted a trust game, where first-mover transfers to second movers are either predetermined or intentionally chosen. FALK, FEHR, and FISCHBACHER (2008) dealt with the so-called moonlighting game which adds to the trust game the opportunity of costly punishment for the second movers.

illogical results. For this reason, RABIN (1993), who was the first to suggest a concrete reciprocity model based on the new framework, restricted his analysis to normal-form games.

DUFWENBERG and KIRCHSTEIGER (2004) introduced belief updating and were thus able to tackle also sequential games. The utility function proposed by DUFWENBERG and KIRCHSTEIGER is

$$U = x^i(a_i, b_{ij}) + \sum_{j \neq i} (\theta_{ij} \kappa_{ij}(a_i, b_{ij}) \lambda_{iji}(b_{ij}, c_{ijk})). \quad (14)$$

It nicely illustrates how first- and second-order beliefs co-determine utilities. a_i denotes i 's strategy, b_{ij} i 's belief about j 's strategy, and c_{ijk} i 's second-order belief about j 's belief about k 's strategy. Strategies and beliefs about them are subject to updating, i. e. they are functions of the history of the game. The first term on the right hand side of equation (14) represents i 's material payoff if she chooses strategy a_i and the other players choose their strategies as she believes them to do. The second term represents i 's utility derived from reciprocity. It consists of the sum of reciprocity terms with respect to each other player $j \neq i$. i 's reciprocity utility with respect to any other player j is the product of three factors: (i) a reciprocity sensitivity parameter $\theta_{ij} \geq 0$, (ii) the kindness of i 's strategy to j , κ_{ij} , given i 's beliefs about j 's strategy, and (iii) i 's belief about j 's kindness to i , λ_{iji} , which depends on i 's beliefs about j 's strategy as well as on her second-order beliefs about j 's beliefs about i 's and others' strategies. If beliefs are assumed to be correct, the last factor becomes equivalent to the kindness of j 's strategy to i , i. e. $\lambda_{iji} = \kappa_{ji}$. The value of θ_{ij} indicates how much i cares for reciprocity to j ; it is exogenously given and does not depend on past behavior. Normalize the kindness value of a strategy that is perceived as neither kind nor unkind to zero. i 's utility then increases if κ_{ij} and λ_{iji} have the same sign, i. e. if i treats j (un)kindly and believes to be treated (un)kindly by j . By contrast, i 's utility decreases if i treats j kindly but believes to be treated unkindly by j , or *vice versa*. DUFWENBERG and KIRCHSTEIGER specified the kindness of i 's strategy a_i to j as the difference between j 's material payoff that results from i 's strategy, given that the other players' behave as i believes them to do, and a reference-state payoff for j . The reference-state payoff, in turn, is defined as the arithmetic mean of the maximum and the minimum payoff for j that i believes to be able to bring about by choosing any strategy from her strategy set. As for the determination of the minimum payoff, an additional constraint ensures that only those possible strategies are considered that induce a Pareto-efficient allocation of material payoffs.

FALK and FISCHBACHER (2006) presented a related reciprocity model which works without belief updating. Instead, reciprocity utility components are

defined for each node of the game. For a given history of the game, the components of all nodes that can still be reached are summed up while those that are no longer feasible do no longer influence utilities. LEVINE (1998) chose a somewhat different approach. In his model, kindness is not defined by actions, but by a parameter value that indicates how kind or altruistic a person is in general. That degree of altruism is drawn from an exogenous distribution and remains fixed over time. However, as the degrees of altruism are private knowledge, actions can reveal information and belief updating plays a role.

Another model by COX, FRIEDMAN, and GJERSTAD (2007) can be considered as an extension of the social preference approach by ANDREONI and MILLER (2002). In a two-person framework, COX, FRIEDMAN and GJERSTAD set up a general CES utility function with own and other's payoff as goods. The share parameter for other's payoff is called the "emotional state". It is a function of a reciprocity and a status variable. The determination of the reciprocity variable can be thought of as similar to the process of calculating kindness in the above model by DUFWENBERG and KIRCHSTEIGER (2004). The status variable accounts for factors that establish generally recognized deviations in payoff claims from equal shares, such as effort. COX, FRIEDMAN and GJERSTAD used experimental data from various games to estimate the parameters of their model.

While the reciprocity models mentioned so far are more in the spirit of the interdependent preference models from Section 7, the approach by CHARNES and RABIN (2002) takes into account the impulsivity that is at least sometimes connected to reciprocal behavior. CHARNES and RABIN added reciprocity to the inequity aversion model by FEHR and SCHMIDT (1999). In contrast to other models where reciprocity is incorporated by a continuous variable, CHARNES and RABIN employed a utility function with indicator variables. If another person has misbehaved, the respective indicator variable switches and the weight of that person's payoff in the utility function is discretely diminished. This approach is reminiscent of the visceral influence model by LOEWENSTEIN (1996) in the form presented above.

9 Culture and institutions

According to standard economic theory, individual preferences are unaffected by the type of society and its institutions an individual grows up and lives in. This view stands in contrast to experimental evidence which showed that members of different industrial (see ROTH ET AL., 1991) as well as native small-scale societies (see HENRICH ET AL., 2001) behave differently in the ultimatum

game, presumably due to divergent perceptions and norms of fairness. The theory of cultural transmission (see BISIN and VERDIER, 2000; CAVALLI-SFORZA ET AL., 1982) provides an explanation how societal influences co-determine preferences: many aspects of individual preferences are not inborn, but formed during childhood and adolescence following a process of social learning and imitation of role models. HOLBROOK and SCHINDLER (1989; SCHINDLER and HOLBROOK, 1993) found that tastes for music and personal appearance each develop within a critical age period. Evidence from neurobiology and psychology points at more fundamental influences. Studies with mono- and dizygotic twins revealed impacts of non-genetic, environmental factors on emotions and even on the formation of neural connections within the brain (see ZIZZO, 2003). This evidence corroborates the hypothesis by MARKUS and KITAYAMA (1991) that cultural socialization may induce in an individual a certain construal, i. e. perception of the world and herself, leading to differences among cultures in such basic functions as cognition, experience, and emotion. Cultural transmission can also explain stratification within a single society when the experiences of its young members, who are in their preference-forming period, are, either unconsciously or intentionally to maintain class identity, restricted to a certain milieu (see AKERLOF, 1983). AKERLOF and KRANTON (2000) incorporated social identities into agents' utility functions and then analyzed the economic effect in different situations, thus elucidating e. g. discrimination and social exclusion.

Note that cultural environments are usually widely static and mostly change, if at all, only over generations. Cultural factors therefore do not explain true preference instability on the individual level, but rather preference heterogeneity across cultures and generations. However, this is less true for the further institutional factors discussed in this section, i. e. the legal, economic, and political system. While these factors usually are quite stable, as well, they sometimes change abruptly and can thus give rise to individual preference change.

Extending the economic approach to the legal system, crime appears as a negative external effect. An optimal policy thus requires that marginal enforcement costs be equated with the marginal external costs caused by a crime. The socially optimal level of crime then arises endogenously as individuals, when maximizing their expected utility, decide whether committing a certain crime pays off for them or not (see BECKER, 1968). SUNSTEIN (1986) and DAU-SCHMIDT (1990) criticized this view, arguing that shaping people's preferences also is an important objective of law. DAU-SCHMIDT interpreted several features of legal systems all over the world as empirical evidence for his position. For instance, intent often matters tremendously for the fierceness of a punishment – also attempted crimes are punished – while the actual harm caused seems to be only of secondary

importance. Furthermore, individuals are also punished for crimes they committed on behalf of artificial persons. Viewing the legal system as means to internalize external effects, these regulations appear as superfluous and even welfare-reducing whereas they make much sense when a preference-shaping function of law is admitted. COOTER (1998) applied the adaptive preference approach to the legal system. If laws make opportunities contingent on preferences, it may be a rational reaction for an individual to adapt her preferences to avoid that her opportunities are curtailed. For example, since being dishonest can, after a corresponding sentence, result in being excluded from certain offices, one may decide to acquire a preference for being honest in the first place.¹⁷

Economic experiments have produced some evidence that also the economic system itself may influence preferences. While the results of bargaining experiments such as the ultimatum game stand in sharp contrast to the assumption of narrow self-interest and point to interdependent, social preferences, the behavior of subjects in most cases quickly converges to the equilibrium implied by the standard model assuming pure egoism when the experimental design resembles a competitive market setting. This holds true even in those cases where the market equilibrium entails an extreme distribution with one market side earning almost the whole surplus (see CASON and WILLIAMS, 1990; GROSSKOPF, 2003; PRASNIKAR and ROTH, 1992; ROTH ET AL., 1991).¹⁸ So, do markets change preferences? More precisely, do competitive markets make people selfish? This conjecture is all but new; MARSHALL already reported that “[t]he term ‘competition’ has gathered about it evil savour, and has come to imply a certain selfishness and indifference to the wellbeing of others” (1930, p. 6). The standard economic approach explains this phenomenon with a change in opportunities that is brought about by the different institutional designs. This change in opportunities triggers a change in behavior, but leaves preferences unaffected. Recognizing that in a large, competitive market a single individual is unable to affect the market outcome, also an altruist has to accept that behaving altruistically will not help anyone and that he can therefore simply go for maximizing his own payoff. Hence, egoistic behavior does not imply egoistic

17 Further consequences of changes in the design of the legal system that are unexpected from the viewpoint of standard economics may result from motivation crowding effects. FREY (1997) argued that distrusting laws may crowd out the civic virtues of citizens and therefore yield adverse effects.

18 But note as a counter example the study of FEHR, KIRCHLER, and WEICHBOLD (1998) in which a competitive market setting was unable to undermine reciprocal behavior.

preferences (see BUCHANAN, 1978; SOBEL, 2009). FEHR and SCHMIDT (1999) and BOLTON and OCKENFELS (2000) pointed out that competitive market outcomes are compatible with their models of interdependent preferences if there is a small minority of agents who hold egoistic preferences since this minority can induce the competitive equilibrium.¹⁹

Nevertheless, the story does not end here. BUCHANAN (1978) argued that if the ruling economic system of a large-scale market economy deprives the individuals of the opportunity to express their true preferences and let these true preferences take effect, the state should change the system to a more locally segmented structure which were more in accordance with the moral-ethical principles of its citizens. Moreover, BOWLES (1998) mentioned a channel through which markets indirectly could nonetheless affect preferences. In the long run, the as-if selfish acting on competitive markets could rub off on preferences. Combining this view with the theory of cultural transmission yields worrying implications. Even if competitive markets do not change our preferences, they may cause our children to become egoists as they copy our as-if selfish behavior in those markets.

It is now quite a small step to the Marxist critics, who build on a model with malleable preferences when they argue that capitalists are indoctrinating people to maintain power. Note that a model of endogenous preferences perfectly fits the Marxist theory of historical materialism according to which “all history is nothing but a continuous transformation of human nature” (MARX, 1992, p.107). Leaving aside the influence of markets, there is some empirical evidence which suggests that the political system affects preferences, as well. GERBER and JACKSON (1993) set up a model in which voters modify their attitudes towards political issues after the party they feel affiliated to has changed its position. They considered the evolution of opinions following the debates on the civil rights movement and the Vietnam war in the US to be consistent with their model. ALESINA and FUCHS-SCHÜNDELN (2007) used the reunification of Germany as a natural experiment. They were able to show that Germans raised in socialist East Germany are more likely to approve of state provision for social insurance when income is held constant and a wealth of

¹⁹ BESTER and GÜTH (1998) showed that a minority of egoists (and even malevolents, as noted by BOLLE, 2000) can spread if the actions of the agents are strategic substitutes. In contrast, when actions are strategic complements, i. e. decent behavior will be reciprocated and nasty behavior will be punished, a minority of altruists can spread. CAMERER and FEHR (2006) provide an overview over the conditions under which selfish or, respectively, social behavior is likely to dominate.

demographic variables is controlled for. This corroborates another theoretical channel of preference modification suggested by BOWLES (1998): as the extent to which certain goods such as social insurance are provided by the state varies among political systems, citizens used to an extensive welfare state may come to value its benefits more due to an endowment effect. OCKENFELS and WEIMANN (1999) also compared the behavior of West and East Germans in an experimental study. Yet, they found that East Germans showed less solidarity, which, the authors speculate, might be a consequence of the experienced shortages of goods in a socialist system. A subsequent study on the same matter by BROSIG-KOCH ET AL. (2011) found that the differences in solidarity between West and East Germans had not diminished 20 years after the German reunification. This may be explained by the fact that the extent of solidarity is based on local social norms which can be expected to be rather stable as they are, in contrast to political attitudes, frequently reinforced in social interactions.

Strong criticism of capitalism was raised by GINTIS (1974), who claimed, referring to empirical studies which illustrate with US data that individual income is far stronger correlated with schooling than with cognitive abilities, that the US educational system would be more concerned with teaching conformity to the economic system, i. e. changing preferences, than transferring knowledge. If, however, the economic system changes preferences, then total welfare which, in turn, depends on these preferences can no longer provide a valid evaluation of the economic system, for otherwise a circular argument ensues. In a similar vein, GALBRAITH (1958) argued that the increase of production in modern industrial societies has become an end in itself as the want for new products was artificially created through preference manipulation. MARCUSE claimed:

We may distinguish both true and false needs. "False" are those which are superimposed upon the individual by particular social interests in his repression [...]. No matter how much such needs may have become the individual's own [...]; no matter how much he identifies himself with them and finds himself in their satisfaction, they continue to be what they were from the beginning – products of a society whose dominant interest demands repression. (1964, pp. 4 f.)

He saw the individual in capitalist society as a mere instrument serving economic interests and demanded people's, if necessary coercive, liberation from their existence in servitude.

At this point, the theory of indoctrination obviously runs into problems. Who has the ability and authority to decide which preferences of others are true and which are false? Who is allowed, and under which conditions, to

manipulate others' preferences against those others' wills?²⁰ Literature from diverse social sciences may at least provide an orientation. GOODIN (1993), while considering respect towards people's preferences the essence of democracy, allowed for cases of "permissible paternalism". Such cases could arise when an individual was uninformed or unable to correctly process information or when her preferences were inconsistent or only superficial. Whether one of these conditions is fulfilled for a particular case, however, appears to be often debatable. A similar catalogue of cases in which individual preferences may be legitimately overturned was put up by SUNSTEIN (1986). According to DAU-SCHMIDT'S (1990) view, the electorate of a country, moderated by the structure of the political system, decides through its voting behavior which and whose preferences are valued how much by the society. SUNSTEIN (1986) emphasized the safeguarding function of political institutions. In his view, the constitution of a state embodies its citizens' second-order preference that certain rights and liberties be protected, possibly also against their more changeable and superficial first-order preferences.

SUNSTEIN'S and THALER'S (2003; THALER and SUNSTEIN, 2009) concept of "libertarian paternalism" or "nudging" has received much interest. They postulated the inevitability of preference manipulation. This inevitability arises since human behavior is contingent on a multitude of context effects (which are discussed in Section 10 of this survey). Yet, a choice situation has to be presented to the choosers with some design or framing. Given that influencing people's preferences is to some extent unavoidable, the framing should, while retaining freedom of choice, promote a decision that is in the individual's true interest. As a prime example, THALER and SUNSTEIN mentioned the order of different foods in a cafeteria which has a tremendous impact on which items are consumed. A clever order in this case could prevent individuals from eating unhealthy food in a sudden moment of weak will. BOVENS (2009) looked at the nudging approach with ambivalent feelings. Though acknowledging its merits, he warned that nudging could in the long run lead to the infantilization of the chooser. WRIGHT and GINSBURG (2012) criticized that a superiority of regulation was mostly assumed *a priori* by proponents of liberal paternalism although a thorough cost-benefit analysis was needed. Moreover, the value of autonomy of choice *per se* and the chance to learn from one's errors and develop one's personality were ignored. Further criticism was advanced by GRÜNE-YANOFF (2012), who insisted that it does make a difference for people's liberty whether their preferences are manipulated only accidentally or intentionally. SUGDEN (2008)

²⁰ MARCUSE (1964) himself also conceded these problems, see pp. 40 f.

questioned the inevitability of preference manipulation, proposing as an alternative to leave it to the market how to design and present choice situations.

10 Reference points

Although there are earlier studies on reference dependence in the form of adaptation in psychology and psychophysics (see BRICKMAN, COATES, and JANOFF-BULMAN, 1978; HELSON, 1948; for an overview see FREDERICK and LOEWENSTEIN, 1999) as well as in economics (see MARKOWITZ, 1952), the approach of generally reference-dependent preferences gained impact in the economics literature first for decisions under risk. Phenomena such as the Allais and Ellsberg paradoxes led TVERSKY and KAHNEMAN (1974) to argue that, due to constrained cognitive capacities, men often do not optimize, but come to a decision by employing certain heuristics. Acting according to those heuristics, while appropriate in most cases, would lead to seriously suboptimal decisions in some instances. In a seminal paper, they (KAHNEMAN and TVERSKY, 1979) introduced “prospect theory” as an empirical model of decision-making under risk.

According to prospect theory, the process of decision-making consists of two phases. The first phase is the editing phase during which the individual redefines the outcomes of her choice alternatives as prospects, i. e. she codes them as gains or losses relative to her reference point. The question how this reference point, in turn, is determined is postponed until later in this section. Furthermore, the individual simplifies the choice situation. Such simplifications may include the elimination of dominated alternatives from the choice set, disregarding common consequences of all alternatives,²¹ and rounding probabilities. In the following evaluation phase, the individual assesses the simplified alternatives and chooses the alternative that conveys the highest value V . For assessment, she deploys a value function

$$V = \sum_s \pi_s v_s. \quad (15)$$

While the structure of this function resembles the calculation of expected utility, the entering factors differ. π_s denotes the subjective probability of a state s . KAHNEMAN and TVERSKY claimed that men are prone to overestimate small and underestimate middle and high probabilities. Highly probable

²¹ KŐSZEGI and SZEIDL (2013) deal with the tendency to neglect common consequences by suggesting a focus-weighted utility function. Attributes for which the chooser's options provide a larger range of different realizations receive higher decision weights.

consequences might be perceived as certain. Moreover, the subjective value of a state v_s does not directly depend on the outcome x_s associated with it, but rather on the change this outcome constitutes assessed from the individual's reference point r , i. e.

$$v_s = v_s(x_s - r). \quad (16)$$

Thus, if the argument of v_s is positive or negative, the outcome is perceived as a gain or loss, respectively. The value of an outcome identical to the reference point is normalized to zero: $v_s(0) = 0$. The subjective value function is assumed strictly increasing. Moreover, in order to meet the empirical evidence for risk aversion with respect to gains and risk-seeking with respect to losses, the function is assumed concave in its gain domain and convex in its loss domain, i. e. $d^2v_s/d(x_s - r)^2 \begin{cases} < \\ > \end{cases} 0$ for $x_s \begin{cases} > \\ < \end{cases} r$. The empirical feature of loss aversion, i. e. losses psychologically weigh heavier than equal-sized gains, is incorporated by the demand that the function be steeper in its loss than in its gain domain. TVERSKY and KAHNEMAN (1991) extended prospect theory to decisions in deterministic environments.²² Basically, expression (15) then becomes obsolete and the individual simply, after editing them in the first phase, evaluates her alternatives according to the function given in equation (16).

TVERSKY and SIMONSON (1993) set up a model that rationalizes choice set-dependent preferences in decisions under certainty in that the chooser uses other options of the choice set as reference points when she assesses the value of an option. The value of an option A given choice set S is described by the value function

$$V(A, S) = \sum_{i=1}^n v_i(a_i) + \beta \sum_{B \in S, B \neq A} RADV(A, B). \quad (17)$$

The first summand of the right hand side of equation (17) represents the reference-independent value of option A . It is assumed that there are n value-relevant dimensions and that the option can be described as a vector of characteristics a_i , one for each relevant dimension. The reference-independent value of A is made up of the sum of A 's values v_i in the single dimensions. The second summand represents the choice set-dependent value of A , where the coefficient β determines the relative importance of choice set effects. $RADV(A, B)$ denotes the relative advantage of option A over option B . In the second summand, the

²² An alternative model of reference-dependent preferences for decisions under certainty is presented by MUNRO and SUGDEN (2003). This model is an extension of neoclassical consumer theory rather than prospect theory.

relative advantages of option A over all other options in the choice set are summed up. The relative advantage of A over B, in turn, is defined as

$$RADV(A, B) = \frac{ADV(A, B)}{ADV(A, B) + DISADV(A, B)}, \quad (18)$$

where $ADV(A, B) = \sum_{i|v_i(a_i) > v_i(b_i)} (v_i(a_i) - v_i(b_i))$ and $DISADV(A, B) = f(ADV(B, A))$ with $DISADV(A, B) \geq ADV(B, A) \forall A, B$. That is, the relative advantage of A over B is the absolute advantage of A over B divided by the sum of the absolute advantage and the absolute disadvantage of A over B. The absolute advantage of A over B is determined by A's and B's characteristics a_i and b_i in the value-relevant dimensions. $(v_i(a_i) - v_i(b_i))$ describes the value difference between options A and B in dimension i . The absolute advantage of A over B is the sum of these differences over all dimensions in which the difference is positive, i. e. A is better than B. The absolute disadvantage of A over B is not simply the same as the absolute advantage of B over A. Analogously to prospect theory, TVERSKY and SIMONSON suppose that disadvantages loom larger than advantages of equal size. The absolute disadvantage of A over B is therefore always at least as large as the accompanying absolute advantage of B over A. Moreover, the absolute disadvantage of A over B is assumed to be an increasing and convex function of the absolute disadvantage of B over A.²³

TVERSKY's and SIMONSON's (1993) model can explain how the addition of a further option to the choice set can change the preference order between two options already contained in the set. Let there be two options A and B and two value-relevant dimensions. Let A be better than B in dimension 1 and B be better than A in dimension 2; overall let the chooser be indifferent between options A and B in a binary choice situation. Now a third option C is added to the choice set. Let C be still better than A in dimension 1 and still worse than A in dimension 2. Following the introduction of option C, the evaluation of options A and B changes in that now their relative advantages over C have to be considered. If these relative advantages are different, the indifference between options A and B will turn into a preference for one option. If $RADV(A, C) > RADV(B, C)$, the chooser will now prefer A to B. It is reasonable to assume that the stated inequality will hold due to the convexity of the disadvantage function. Option C adds to the consideration of the value of option

²³ The modified version of TVERSKY's and SIMONSON's (1993) model presented here abstracts from a further reference point effect, the background effect, and assumes that relative advantages are evaluated also in binary choice situations. These simplifications are innocuous with respect to the choice set effect we are interested in.

A (B) a small (large) disadvantage in dimension 1 and a small (large) advantage in dimension 2. Intuitively, if the new large disadvantage of option B dominates, option B is now regarded inferior to option A.

There is a vast literature on empirical and experimental phenomena that may be explained by prospect theory and reference point-dependent preferences. Due to the multitude of phenomena, not all can be dealt with in the following. CAMERER's (2000) summary gives an impression of the variety of them.

Inquiries into the causes of preference reversal suggest that the elicitation procedure may influence the individual's reference point as it draws the attention of the subject to that dimension of the available gambles which corresponds to the dimension in which the answer is required. That is to say, when subjects are asked for a selling price of a gamble, they concentrate on the amount that can be won whereas they focus more on winning probabilities when choosing between gambles (see TVERSKY and THALER, 1990; TVERSKY, SLOVIC, and KAHNEMAN, 1990). THALER (1980) discussed the empirically observable neglect of opportunity costs and consideration of sunk costs. Since opportunity costs are forgone gains, the principle of loss aversion predicts that they will receive less psychological weight than costs which lead to disbursements and are thus perceived as losses. The second phenomenon can be explained by assuming that the individual's reference point has not yet adjusted to a sure loss. BENARTZI and THALER (1995) showed that loss aversion combined with sufficiently frequent portfolio evaluations by investors can account for the equity premium puzzle. A considerable strand of literature has been concerned with the gap between the willingness to pay for a good and the amount demanded for giving up the very same good and the related phenomena of status quo bias and endowment effect. The well-documented effect that an individual values a good more highly if she is in possession of it is explained by proponents of prospect theory as a reference point shift. While the individual perceives a good she does not own as a gain, giving up the same good once she is endowed with it means a loss. Alternative accounts of the phenomenon that deviate more from standard economics draw on incomplete preferences (see MANDLER, 2004) or different procedures of memory retrieval being triggered for preference construction (see JOHNSON, HÄUBL, and KEINAN, 2007). Note that while the endowment effect and related asymmetries have been reproduced in many cases (see BATEMAN ET AL., 1997; KAHNEMAN, KNETSCH, and THALER, 1990, 1991; KNETSCH, 1989; KNETSCH and SINDEN, 1984), there is evidence of and an ongoing discussion about conditions under which the effect does not or no longer occur (see ISONI, LOOMES, and SUGDEN, 2011; KNETSCH and WONG, 2009; LIST, 2003, 2004, 2006; PLOTT and ZEILER, 2005, 2007, 2011). Especially, the endowment effect vanishes if subjects gain sufficient practice and experience.

Let us now return to the question what determines an individual's reference point. The answer to this question is not straightforward. TVERSKY and KAHNEMAN stated: "A diversity of factors determine the reference outcome in everyday life. The reference outcome is usually a state to which one has adapted; it is sometimes set by social norms and expectations; it sometimes corresponds to a level of aspiration, which may or may not be realistic" (1981, p. 456). Thus, the approaches of habit formation (aspirations raised by past consumption), interdependent preferences (aspirations raised by others' consumption), and cultural and institutional influences (social norms) may be interpreted as sub-cases of a general reference dependence of preferences. Following HOCH and LOEWENSTEIN (1991), also emotional influences on preferences can be subsumed under a reference-dependent model by assuming that cues such as physical proximity can bring about a temporary shift of the reference point. If the individual has psychologically adapted to performing an action triggered by the cue, she may perceive a failure to do so as a loss. Accordingly, the desire for acting can increase. Furthermore, to explain the effects of the framing of the choice environment on preferences, which gave rise to prospect theory in the first place, KAHNEMAN and TVERSKY approved that the reference point "can be affected by the formulation of the offered prospects" (1979, p. 274). Examples of framing effects can be found in TVERSKY and KAHNEMAN (1981). For instance, given the hypothetical situation that 600 people of a population suffer from a life-threatening disease, a majority of subjects favored a therapy that saves 200 lives with certainty over one that involves a chance of one third that all 600 are saved and a chance of two thirds that no one is saved. However, when the problem was reformulated in a way such that the first therapy implies the certain death of 400 people while the second involves a two-third chance of 600 deaths and a one-third chance of no deaths, the second option was mostly preferred. Although both formulations describe the same consequences, it is argued that the first formulation implies a reference point in which no one is saved while the second implies a reference point in which all 600 people are saved.

At this point, the model may appear arbitrary. Since the reference point can be influenced by almost anything, any preference change may be explained *ex post* by an appropriate reference point shift. Such an assessment is not quite fair as there are laboratory experiments which found that preferences changed as expected when subjects' reference points were manipulated through framing (see e. g. BATEMAN ET AL., 1997). Yet, as described in the above paragraph, prospect theory does not even contain a precise explanation of how a reference point is determined in a given situation. Experimental tests of the theory hence have to rely on identifying the reference point beforehand by common sense in

situations that are designed such that common sense yields an unambiguous prediction. However, it appears very intricate to control for all factors that can potentially affect an individual's reference point in the field. A model that determines reference points endogenously seems preferable to identifying the probable reference point through psychological reasoning and then testing the accompanying hypotheses on preferences. SCHMIDT and ZANK (2012) presented an approach to derive the reference point from preferences, exploiting the supposition that within both the gain and the loss domain the value function becomes steeper when the reference point is approached. While the reference point can thus be determined endogenously for a given situation, it still does not allow for predictions how the reference point changes with the context. KŐSZEGI and RABIN (2006) took a further step into this direction. They set up a utility model with an additively separable reference-dependent term in which the individual's reference point equals her rational expectations about outcomes. Reference points based on rational expectations have found experimental support in studies by MARZALLI ERICSON and FUSTER (2011) and ABELER ET AL. (2011). Yet, when expected and realized outcomes coincide, the utility function in KŐSZEGI's and RABIN's model boils down to the standard case. Hence, the model does not account for reference point effects under certainty.

Several empirical reference point effects cast doubts on the rationality of such preferences. Preferences also change in reaction to variations of the choice environment that, from a normative point of view, should be irrelevant. The choice between therapy programs mentioned above constitutes such an example. Moreover, TVERSKY and KAHNEMAN (1981) reported that a majority of subjects, when confronted with two binary choices at a time, could be induced by an appropriate framing to choose a pair of options the combined outcome of which was strictly dominated by the combined outcome of the other two options. TVERSKY and KAHNEMAN agreed that the susceptibility of preference ratings to framing effects does not need to imply unstable underlying informed preferences. Rather, the individual may believe that some of her perceptions are erroneous and strive for preference consistency. BERNHEIM and RANGEL (2009) proposed a generalized version of revealed preference theory in which the choice situation is described by the choice objects with their relevant properties and by an ancillary condition. The ancillary condition collects features of the choice environment such as the framing which affect the chooser's preferences, but are considered irrelevant from the viewpoint of a social planner. The model thus makes the irrationality of certain preference changes explicit.

It is not always clear how to resolve a preference inconsistency. While choosing a dominating bundle of options is logically compelling, there is no obvious framing for describing the consequences of a therapy program. More

recent research may even aggravate concerns about irrationality. ARIELY, LOEWENSTEIN, and PRELEC (2003) were able to induce in their subjects a completely arbitrary reference point which obviously has no logical connection to the decision problem. When subjects were first asked whether they were ready to purchase a good for an amount of money equal to the last digits of their social security number and after that their exact willingness to pay for that good was elicited, a strong and significant positive correlation between individuals' willingnesses to pay and the last digits of their social security numbers was observed. Since the supposition that the subjects believed that their social security number would contain some helpful information to determine their willingness to pay appears absurd, it seems most likely that an unconscious psychological priming process was at work.²⁴ Indeed, there is a large literature on mere exposure effects in psychology which acknowledges that sheer repeated exposure of an individual to a stimulus can enhance the individual's preference for that stimulus (see BORNSTEIN, 1989; ZAJONC, 1968, 2001). Such an effect even occurs if the stimulus is presented subliminally. Further studies showed that subliminal stimuli can also affect real consumption behavior in a laboratory setting (see STRAHAN, SPENCER, and ZANNA, 2002; WINKIELMAN, BERRIDGE, and WILBARGER, 2005).

11 Conclusion

As has already been pointed out in the introduction, the question whether individual preferences are stable or unstable depends on the employed model and is thus a matter of definition. However, it is clear that preferences are unstable from the viewpoint of the standard model in microeconomics since behavior is influenced by more factors than just prices and income.

Some approaches circumvent treating changing behavior as changing preferences by ascribing the changes to altered side conditions rather than an altered utility function. Information is one of those side conditions. As far as altered behavior can reliably be traced back to changes of the individual's information about the environment, it seems legitimate not to speak about changing tastes. Yet, one should keep in mind that the individual, due to

²⁴ ARIELY, LOEWENSTEIN, and PRELEC (2006) showed in another experiment that, while subjects' relative valuations for different goods may be consistent, the absolute valuation level can be influenced by arbitrary anchors.

psychological processes, may not always be able to correctly distinguish preferences from information. When information about an individual's own preferences is concerned, the formal treatment becomes ambiguous. Another side condition refers to the individual's endowment with specialized human capital. The question then becomes by which factors and through which channels human capital rather than preferences is affected. A model in which human capital adjusts arbitrarily in order to keep the assumption of stable preferences does not constitute any progress.

Various approaches include additional arguments into the individual's utility function. The adaptive preference approach proposes that the individual strategically, albeit unconsciously, manipulates her preferences in order to increase her utility. While there is considerable evidence for dissonance reduction phenomena in the psychology literature, applications to economics have so far been scarce. Pursuing this avenue further may help to better explain behavior also in some economic environments. Nevertheless, from a normative viewpoint, adaptive preferences appear indefensible. Preference adaptation can be seen as self-deception that endangers the individual's freedom of choice. Besides, the rationality of preference adaptation is questionable. The claim that one rationally deceives oneself in order to be better off clearly stretches the meaning of rationality.

In the cases of habit formation and interdependent preferences, broad empirical evidence that favors these approaches can be produced. A bulk of phenomena that are inconsistent with standard theory can be explained. Models that impose particular forms on the relationship between the additional exogenous factors and utility have been developed and, successfully, put to the test. Thus, the preference changes caused by one's own past or others' consumption are systematized and become predictable. From a moral standpoint, one may decline addiction as an extreme form of habit formation or negative preference interdependence in the form of envy as a basis for a normative model. Yet, such an argument does neither apply to non-pathological effects of habituation nor to preference interdependence in the form of altruism. As for rationality, it is an old but still unresolved discussion whether habit formation effects are anticipated by the individual or she reacts to them myopically, leading to potentially inconsistent dynamic patterns. However, most habit formation models employed in the literature assume rational habit formation and are successful in explaining empirical data under this assumption. Apart from adding exogenous factors to the standard model, interdependent preferences may also cause inherent intertemporal instability. When and why social preferences are inherently unstable is the object of current experimental investigations.

The approach that considers emotional influences is less developed, but promising. While one may intuitively ascribe some human behavior patterns to emotional influences on preferences, such anecdotic evidence remains speculative. However, further supporting evidence stems from controlled laboratory experiments. Due to the connection of emotions to the intensity of measurable cues and physiological factors, the approach is not arbitrary. First models that formalize the effect of these factors on preferences have been developed. More progress has been made in formalizing the influences of reciprocal concerns. Whether reciprocity should rather be considered a sub-case of interdependent or of emotional preferences depends on whether impulsivity is assumed in reciprocal behavior and thus on the concrete model formulation. Normatively, emotional influences appear questionable with respect to both morality and rationality. Behavior triggered by intense emotions is often deemed indecent or at least foolish, and switching preferences lead individuals to behave inconsistently such that they frequently regret *ex post* the actions they committed when driven by their feelings.

The evidence for preference change caused by changes in various institutional factors is scattered. An overarching model does not exist. Anyway, institutional factors mostly remain constant; hence, including them into the analysis is worthwhile only in some special cases. However, even when changes in both institutions and economic outcomes are observed, the effects of institutional factors on preferences and opportunities have to be carefully disentangled. Preference manipulations by the state via its institutions appear highly dubious in general. Nevertheless, the argument that people should, without coercion, be “nudged” to certain behaviors is quite popular.

A theory of reference-dependent preferences in its general form assembles the previous approaches in that the reference point may be influenced by various of the above-mentioned factors. Ample evidence supports prospect theory and its extension to deterministic environments. Nonetheless, further research is needed to systematize the effects of the choice environment on the reference point. The empirical changeability of behavior in response to framing manipulations may also lead economists to reconsider the theory that preferences, to some degree, are constructed only when needed for coming to a decision. Yet, also an approach of constructed preferences, lest it be arbitrary, needs to specify which aspects of the choice environment tend to trigger which psychological processes or preference formation techniques which, in turn, lead to what kind of preferences. The developing field of neuroeconomics promises to provide further insights in this respect. As far as sheer framing effects on preferences are concerned, the reference dependence seems indefensible from the point of rationality.

As this survey has shown, diverse approaches exist that extend standard theory to account for changing tastes. However, there are considerable differences in the elaboration and, hence, the usefulness of these theories. Although some of the approaches – notably habit formation and interdependent preferences and, to some degree, prospect theory and emotional influences – possess a sound theoretical foundation, describe influences on preferences that are relevant for many choice situations economists deal with, and are supported by ample empirical and experimental evidence, no theory of unstable preferences has made its way into standard microeconomics textbooks yet.

A last remark shall be made on further topics that might fit into this survey, but were not dealt with. Firstly, the literature on time inconsistency and hyperbolic discounting (see e.g. LAIBSON, 1997; STROTZ, 1956) has been neglected. Formalizing the original theory of time inconsistency, time itself needs to be included as an additional argument in the utility function. Thus, an empirically untestable model results since time can never be held constant. Newer advances suggest a “quasi-hyperbolic” discounting scheme which attributes a special salience to the respective present, but otherwise allows for time consistency. One may therefore refer to the approaches presented in Sections 8 and 10 and regard the salience of a present reward as a cue or treat the present as a reference point (see LOEWENSTEIN, 1988). Secondly, evolutionary influences on preferences have recently been discussed extensively in the literature (see e.g. SAMUELSON, 2001). However, genetic evolution can lead to altered preferences only over successive generations whereas the preferences of any particular individual are fixed, because of which this strand of literature has been omitted.²⁵

Acknowledgements: I thank Joachim Weimann, Till Requate and three anonymous referees for helpful suggestions on earlier versions of this survey which led to considerable improvements.

References

- ABEL, A. B. (1990): Asset Prices Under Habit Formation and Catching up with the Joneses, *American Economic Review, Papers and Proceedings* 80, 38–42.

²⁵ The term “evolution”, in a broader, non-genetic sense, is also used for changes on the individual level. In this sense, however, “evolution” is synonymous to “learning” whose effects are discussed in Section 3.

- ABEL, A. B. (2005): Optimal Taxation When Consumers Have Endogenous Benchmark Levels of Consumption, *Review of Economic Studies* 72, 21–42.
- ABELER, J., A. FALK, L. GOETTE and D. HUFFMAN (2011): Reference Points and Effort Provision, *American Economic Review* 101, 470–492.
- AKERLOF, G. A. (1983): Loyalty Filters, *American Economic Review* 73, 54–63.
- AKERLOF, G. A. (1991): Procrastination and Obedience, *American Economic Review, Papers and Proceedings* 81, 1–19.
- AKERLOF, G. A. and W. T. DICKENS (1982): The Economic Consequences of Cognitive Dissonance, *American Economic Review* 72, 307–319.
- AKERLOF, G. A. and R. E. KRANTON (2000): Economics and Identity, *Quarterly Journal of Economics* 115, 715–753.
- ALCHIAN, A. A. (1950): Uncertainty, Evolution, and Economic Theory, *Journal of Political Economy* 58, 211–221.
- ALESINA, A. and N. FUCHS-SCHÜNDELN (2007): Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences, *American Economic Review* 97, 1507–1528.
- ALLAIS, M. (1953): Le Comportement De l'Homme Rationnel Devant Le Risque: Critique Des Postulats Et Axiomes De l'École Américaine, *Econometrica* 21, 503–546.
- ALPIZAR, F., F. CARLSSON and O. JOHANSSON-STENMAN (2005): How Much Do We Care About Absolute Versus Relative Income and Consumption?, *Journal of Economic Behavior and Organization* 56, 405–421.
- ANAND, P. (1987): Are the Preference Axioms Really Rational?, *Theory and Decision* 23, 189–214.
- ANAND, P. (1993): The Philosophy of Intransitive Preference, *Economic Journal* 103, 337–346.
- ANDERSEN, S., G. W. HARRISON, M. I. LAU and E. E. RUTSTRÖM (2008): Lost in State Space: Are Preferences Stable?, *International Economic Review* 49, 1091–1112.
- ANDERSON, L. R. and C. A. HOLT (1997): Information Cascades in the Laboratory, *American Economic Review* 87, 847–862.
- ANDREONI, J. and J. MILLER (2002): Giving According To GARP: An Experimental Analysis of the Consistency of Preferences for Altruism, *Econometrica* 70, 737–753.
- ARIELY, D. and G. F. LOEWENSTEIN (2006): The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making, *Journal of Behavioral Decision Making* 19, 87–98.
- ARIELY, D., G. F. LOEWENSTEIN and D. PRELEC (2003): “Coherent Arbitrariness”: Stable Demand Curves Without Stable Preferences, *Quarterly Journal of Economics* 118, 73–105.
- ARIELY, D., G. F. LOEWENSTEIN and D. PRELEC (2006): Tom Sawyer and the Construction of Value, *Journal of Economic Behavior and Organization* 60, 1–10.
- ARMSTRONG, W. E. (1939): The Determinateness of the Utility Function, *Economic Journal* 49, 453–467.
- ARONSON, E. (1997): Back to the Future: Retrospective Review of Leon Festinger's “A Theory of Cognitive Dissonance”, *American Journal of Psychology* 110, 127–137.
- AUMANN, R. J. (2008): Rule-Rationality versus Act-Rationality. Hebrew University of Jerusalem, Discussion Paper No. 497.
- BALL, S., C. ECKEL, P. J. GROSSMAN and W. ZAME (2001): Status in Markets, *Quarterly Journal of Economics* 116, 161–188.
- BALLINGER, T. P., E. HUDSON, L. KARKOVIATA and N. T. WILCOX (2011): Saving Behavior and Cognitive Abilities, *Experimental Economics* 14, 349–374.

- BANERJEE, A. V. (1992): A Simple Model of Herd Behavior, *Quarterly Journal of Economics* 107, 797–817.
- BATEMAN, I., A. MUNRO, B. RHODES, C. STARMER and R. SUGDEN (1997): A Test of the Theory of Reference-Dependent Preferences, *Quarterly Journal of Economics* 112, 479–505.
- BECKER, G. S. (1965): A Theory of the Allocation of Time, *Economic Journal* 75, 493–517.
- BECKER, G. S. (1968): Crime and Punishment: An Economic Approach, *Journal of Political Economy* 76, 169–217.
- BECKER, G. S. (1974): A Theory of Social Interactions, *Journal of Political Economy* 82, 1063–1093.
- BECKER, G. S. (1992): Habits, Addictions, and Traditions, *Kyklos* 45, 327–346.
- BECKER, G. S. and K. M. MURPHY (1988): A Theory of Rational Addiction, *Journal of Political Economy* 96, 675–700.
- BELL, D. E. (1982): Regret in Decision Making Under Uncertainty, *Operations Research* 30, 961–981.
- BÉNABOU, R. and J. TIROLE (2006): Incentives and Prosocial Behavior, *American Economic Review* 96, 1652–1678.
- BEN-ARAB, M., E. BRIYS and H. SCHLESINGER (1996): Habit Formation and the Demand for Insurance, *Journal of Risk and Insurance* 63, 111–119.
- BENARTZI, S. and R. H. THALER (1995): Myopic Loss Aversion and the Equity Premium Puzzle, *Quarterly Journal of Economics* 110, 73–92.
- BERG, J., J. DICKHAUT and K. MCKABE (1995): Trust, Reciprocity, and Social History, *Games and Economic Behavior* 10, 122–142.
- BERGSTROM, T. C. (1999): Systems of Benevolent Utility Functions, *Journal of Public Economic Theory* 1, 71–100.
- BERNHEIM, B. D. and A. RANGEL (2009): Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics, *Quarterly Journal of Economics* 124, 51–104.
- BESTER, H. and W. GÜTH (1998): Is Altruism Evolutionary Stable?, *Journal of Economic Behavior and Organization* 34, 193–209.
- BETTMAN, J. R., M. F. LUCE and J. W. PAYNE (1998): Constructive Consumer Choice Processes, *Journal of Consumer Research* 25, 187–217.
- BIKHCHANDANI, S., D. HIRSHLEIFER and I. WELCH (1992): A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, *Journal of Political Economy* 100, 992–1026.
- BISIN, A. and T. VERDIER (2000): “Beyond the Melting Pot”: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits, *Quarterly Journal of Economics* 115, 955–988.
- BLANCO, M., D. ENGELMANN and H.-T. NORMANN (2011): A Within-Subject Analyses of Other-Regarding Preferences, *Games and Economic Behavior* 72, 321–338.
- BLOUNT, S. (1995): When Social Outcomes Aren’t Fair: The Effect of Causal Attributions on Preferences, *Organizational Behavior and Human Decision Processes* 63, 131–144.
- BOLLE, F. (2000): Is Altruism Evolutionary Stable? and Envy and Malevolence?, *Journal of Economic Behavior and Organization* 42, 131–133.
- BOLTON, G. E. and A. OCKENFELS (2000): ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review* 90, 166–193.
- BOONE, C., B. DE BRABANDER and A. VAN WITTELOOSTUIJN (1999): The Impact of Personality on Behavior in Five Prisoner’s Dilemma Games, *Journal of Economic Psychology* 20, 343–377.

- BORNSTEIN, R. F. (1989): Exposure and Affect: Overview and Meta-Analysis of Research, 1968–1987, *Psychological Bulletin* 106, 265–289.
- BOSMAN, R. and F. VAN WINDEN (2002): Emotional Hazard in a Power-to-Take Experiment, *Economic Journal* 112, 147–169.
- BOVENS, L. (2009): The Ethics of Nudge, in: T. GRÜNE-YANOFF and S. O. HANSSON (eds.) *Preference Change: Approaches From Philosophy, Economics and Psychology*. Springer, Dordrecht et al., 207–219.
- BOWLES, S. (1998): Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions, *Journal of Economic Literature* 36, 75–111.
- BOWLES, S. and Y. PARK (2005): Emulation, Inequality, and Work Hours: Was Thorsten Veblen Right?, *Economic Journal* 115, F397–F412.
- BOYER, M. (1978): A Habit Forming Optimal Growth Model, *International Economic Review* 19, 585–609.
- BOYER, M. (1983): Rational Demand and Expenditures Patterns Under Habit Formation, *Journal of Economic Theory* 31, 27–53.
- BRANDTS, J. and C. SOLÅ (2001): Reference Points and Negative Reciprocity in Simple Sequential Games, *Games and Economic Behavior* 36, 138–157.
- BRAUN, P. A., G. M. CONSTANTINIDES and W. E. FERNON (1993): Time Nonseparability in Aggregate Consumption: International Evidence, *European Economic Review* 37, 897–920.
- BREKKE, K. A., S. KVERNDOKK and K. NYBORG (2003): An Economic Model of Moral Motivation, *Journal of Public Economics* 87, 1967–1983.
- BRICKMAN, P., D. COATES and R. JANOFF-BULMAN (1978): Lottery Winners and Accident Victims: Is Happiness Relative?, *Journal of Personality and Social Psychology* 36, 917–927.
- BROSIG, J., T. RIECHMANN and J. WEIMANN (2007): Selfish in the End? An Investigation of Consistency and Stability of Individual Behavior. University of Magdeburg, FEMM Working Paper No. 05/2007.
- BROSIG-KOCH, J., C. HELBACH, A. OCKENFELS and J. WEIMANN (2011): Still Different After All These Years: Solidarity Behavior in East and West Germany, *Journal of Public Economics* 95, 1373–1376.
- BROWN, A. L., Z. E. CHUA and C. F. CAMERER (2009): Learning and Visceral Temptation in Dynamic Saving Experiments, *Quarterly Journal of Economics* 124, 197–231.
- BROWNSTEIN, A. L. (2003): Biased Predecision Processing, *Psychological Bulletin* 129, 545–568.
- BRUHIN, A., E. FEHR and D. SCHUNK (2016): The Many Faces of Sociality: Uncovering the Distribution and Stability of Social Preferences. University of Mainz, Discussion Paper No. 1603.
- BUCHANAN, J. M. (1978): Markets, States, and the Extent of Morals, *American Economic Review, Papers and Proceedings* 68, 364–368.
- CAMERER, C. F. (2000): Prospect Theory in the Wild: Evidence From the Field, in: D. Kahneman and A. Tversky (eds.) *Choices, Values, and Frames*. Russell Sage Foundation and Cambridge University Press, New York et al., 288–300.
- CAMERER, C. F. and E. FEHR (2006): When Does “Economic Man” Dominate Social Behavior?, *Science* 311, 47–52.
- CAMPBELL, J. Y. and J. H. COCHRANE (1999): By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* 107, 205–251.
- CAMPBELL, J. Y. and A. DEATON (1989): Why Is Consumption so Smooth?, *Review of Economic Studies* 56, 357–373.

- CAPLAN, B. (2003): Stigler-Becker Versus Myers-Briggs: Why Preference-Based Explanations Are Scientifically Meaningful and Empirically Important, *Journal of Economic Behavior and Organization* 50, 391–405.
- CARBONE, E. and J. DUFFY (2014): Lifecycle Consumption Plans, Social Learning and External Habits: Experimental Evidence, *Journal of Economic Behavior and Organization* 106, 413–427.
- CARLSSON, F., O. JOHANSSON-STENMAN and P. K. NAM (2014): Social Preferences Are Stable Over Long Periods of Time, *Journal of Public Economics* 117, 104–114.
- CARRASCO, R., J. M. LABEAGA and J. D. LÓPEZ-SALIDO (2005): Consumption and Habits: Evidence From Panel Data, *Economic Journal* 115, 144–165.
- CARROLL, C. D., J. OVERLAND and D. N. WEIL (2000): Saving and Growth with Habit Formation, *American Economic Review* 90, 341–355.
- CASON, T. N. and A. W. WILLIAMS (1990): Competitive Equilibrium Convergence in a Posted-Offer Market with Extreme Earnings Inequities, *Journal of Economic Behavior and Organization* 14, 331–352.
- CAVALLI-SFORZA, L. L., M. W. FELDMAN, K. H. CHEN and S. M. DORNBUSCH (1982): Theory and Observation in Cultural Transmission, *Science* 218, 19–27.
- CHAO, A. and J. B. SCHOR (1998): Empirical Tests of Status Consumption: Evidence From Women's Cosmetics, *Journal of Economic Psychology* 19, 107–131.
- CHARNES, G. and M. RABIN (2002): Understanding Social Preferences with Simple Tests, *Quarterly Journal of Economics* 117, 817–869.
- CHU, Y.-P. and R.-L. CHU (1990): The Subsidence of Preference Reversals in Simplified and Marketlike Experimental Settings: A Note, *American Economic Review* 80, 902–911.
- CLARK, A. E., P. FRIJTERS and M. A. SHIELDS (2008): Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles, *Journal of Economic Literature* 46, 95–144.
- COHEN, M. D. and R. AXELROD (1984): Coping with Complexity: The Adaptive Value of Changing Utility, *American Economic Review* 74, 30–42.
- COLE, H. L., G. J. MAILATH and A. POSTLEWAITE (1992): Social Norms, Savings Behavior, and Growth, *Journal of Political Economy* 100, 1092–1125.
- CONLEY, J. J. (1985): Longitudinal Stability of Personality Traits: A Multitrait-Multimethod-Multioccasion Analysis, *Journal of Personality and Social Psychology* 49, 1266–1282.
- CONLISK, J. (1980): Costly Optimizers Versus Cheap Imitators, *Journal of Economic Behavior and Organization* 1, 275–293.
- CONLISK, J. (1996): Why Bounded Rationality?, *Journal of Economic Literature* 34, 669–700.
- CONSTANTINIDES, G. M. (1990): Habit Formation: A Resolution of the Equity Premium Puzzle, *Journal of Political Economy* 98, 519–543.
- COOTER, R. (1998): Expressive Law and Economics, *Journal of Legal Studies* 27, 585–607.
- COSMIDES, L. and J. TOOBY (1994): Better Than Rational: Evolutionary Psychology and the Invisible Hand, *American Economic Review, Papers and Proceedings* 84, 327–332.
- COSTA, P. T. JR and R. R. MCCRAE (1988): Personality in Adulthood: A Six-Year Longitudinal Study of Self-Reports and Spouse Ratings on the NEO Personality Inventory, *Journal of Personality and Social Psychology* 54, 853–863.
- COWEN, T. (1989): Are All Tastes Constant and Identical? A Critique of Stigler and Becker, *Journal of Economic Behavior and Organization* 11, 127–135.
- COX, J. (2004): How to Identify Trust and Reciprocity, *Games and Economic Behavior* 46, 260–281.

- COX, J. C., D. FRIEDMAN and S. GJERSTAD (2007): A Tractable Model of Reciprocity and Fairness, *Games and Economic Behavior* 59, 17–45.
- COX, J. C., D. FRIEDMAN and V. SADIRAJ (2008): Revealed Altruism, *Econometrica* 76, 31–69.
- COX, J. C. and D. M. GREYER (1996): The Preference Reversal Phenomenon: Response Mode, Markets and Incentives, *Economic Theory* 7, 381–405.
- CROSON, R. T. A. (2007): Theories of Commitment, Altruism and Reciprocity: Evidence From Linear Public Good Games, *Economic Inquiry* 45, 199–216.
- CUBITT, R. P., C. STARMER and R. SUGDEN (2001): Discovered Preferences and the Experimental Evidence of Violations of Expected Utility Theory, *Journal of Economic Methodology* 8, 385–414.
- CUBITT, R. P. and R. SUGDEN (2001): On Money Pumps, *Games and Economic Behavior* 37, 121–160.
- CULLISON, W. E. (1978): Examining the Effect of Interdependent Consumer Preferences on Economic Growth, or Re-Discovering Adam Smith and His Eighteenth-Century Contemporaries, *Southern Economic Journal* 44, 937–944.
- DAU-SCHMIDT, K. G. (1990): An Economic Analysis of the Criminal Law as a Preference-Shaping Policy, *Duke Law Journal* 39, 1–38.
- DAVIS, J. A. (1959): A Formal Interpretation of the Theory of Relative Deprivation, *Sociometry* 22, 280–296.
- DE OLIVEIRA, A. C. M., C. ECKEL and R. T. A. CROSON (2012): The Stability of Social Preferences in a Low-Income Neighborhood, *Southern Economic Journal* 79, 15–45.
- DIETRICH, F. and C. LIST (2013): Where Do Preferences Come From?, *International Journal of Game Theory* 42, 613–637.
- DIGMAN, J. M. (1990): Personality Structure: Emergence of the Five-Factor Model, *Annual Review of Psychology* 41, 417–440.
- DUESENBERY, J. S. (1949): *Income, Saving and the Theory of Consumer Behavior*. Harvard University Press, Cambridge, MA.
- DUFWENBERG, M. and G. KIRCHSTEIGER (2004): A Theory of Sequential Reciprocity, *Games and Economic Behavior* 47, 268–298.
- DYNAN, K. E. (2000): Habit Formation in Consumer Preferences: Evidence From Panel Data, *American Economic Review* 90, 391–406.
- EASTERLIN, R. A. (1974): Does Economic Growth Improve the Human Lot? Some Empirical Evidence, in: P. A. DAVID and M. W. REDER (eds.) *Nations and Households in Economic Growth: Essays in Honour of Moses Abramowitz*. Academic Press, New York, 89–125.
- EL-SAFY, A. E. (1976a): Adaptive Behavior, Demand and Preferences, *Journal of Economic Theory* 13, 298–318.
- EL-SAFY, A. E. (1976b): Adaptive Behavior and the Existence of Weizsäcker's Long-Run Indifference Curves, *Journal of Economic Theory* 13, 319–328.
- ELLSBERG, D. (1961): Risk, Ambiguity, and the Savage Axioms, *Quarterly Journal of Economics* 75, 643–669.
- ELSTER, J. (1982): Sour Grapes – Utilitarianism and the Genesis of Wants, in: A. K. SEN and B. A. O. WILLIAMS (eds.) *Utilitarianism and Beyond*. Cambridge University Press, Cambridge, UK et al., 219–238.
- ELSTER, J. (2004): Costs and Constraints in the Economy of the Mind, in: I. BROCAS and J. D. CARRILLO (eds.) *The Psychology of Economic Decisions, Vol. 2: Reasons and Choices*. Oxford University Press, Oxford et al., 3–14.
- ENGEL, C. (2011): Dictator Games: A Meta Study, *Experimental Economics* 14, 583–610.

- FALK, A., E. FEHR and U. FISCHBACHER (2003): On the Nature of Fair Behavior, *Economic Inquiry* 41, 20–26.
- FALK, A., E. FEHR and U. FISCHBACHER (2008): Testing Theories of Fairness – Intentions Matter, *Games and Economic Behavior* 62, 287–303.
- FALK, A. and U. FISCHBACHER (2006): A Theory of Reciprocity, *Games and Economic Behavior* 54, 293–315.
- FEHR, E. and S. GÄCHTER (2000): Cooperation and Punishment in Public Goods Experiments, *American Economic Review* 90, 980–994.
- FEHR, E., E. KIRCHLER and A. WEICHBOLD (1998): When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets, *Journal of Labor Economics* 16, 324–351.
- FEHR, E. and K. SCHMIDT (1999): A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics* 114, 817–868.
- FEHR, E. and P. K. ZYCH (1998): Do Addicts Behave Rationally?, *Scandinavian Journal of Economics* 100, 643–662.
- FERSON, W. E. and G. M. CONSTANTINIDES (1991): Habit Persistence and Durability in Aggregate Consumption: Empirical Tests, *Journal of Financial Economics* 29, 199–240.
- FESTINGER, L. (1957): *A Theory of Cognitive Dissonance*. Row Peterson, Evanston et al.
- FISCHBACHER, U., S. GÄCHTER and E. FEHR (2001): Are People Conditionally Cooperative? Evidence From a Public Goods Experiment, *Economics Letters* 71, 397–404.
- FISMAN, R., S. KARIV and D. MARKOVITS (2007): Individual Preferences for Giving, *American Economic Review* 97, 1858–1876.
- FRANK, R. H. (1984a): Are Workers Paid Their Marginal Products?, *American Economic Review* 74, 549–571.
- FRANK, R. H. (1984b): Interdependent Preferences and the Competitive Wage Structure, *RAND Journal of Economics* 15, 510–520.
- FRANK, R. H. (1985a): *Choosing the Right Pond*. Oxford University Press, New York et al.
- FRANK, R. H. (1985b): The Demand for Unobservable and Other Nonpositional Goods, *American Economic Review* 75, 101–116.
- FRANK, R. H. (1989): Frames of Reference and the Quality of Life, *American Economic Review, Papers and Proceedings* 79, 80–85.
- FRANK, R. H. (2005): Positional Externalities Cause Large and Preventable Welfare Losses, *American Economic Review, Papers and Proceedings* 95, 137–141.
- FRANK, R. H. (2008): Should Public Policy Respond to Positional Externalities?, *Journal of Public Economics* 92, 1777–1786.
- FRANKFURT, H. G. (1971): Freedom of the Will and the Concept of a Person, *Journal of Philosophy* 68, 5–20.
- FREDERICK, S. and G. F. LOEWENSTEIN (1999): Hedonic Adaptation, in: D. KAHNEMAN, E. DIENER and N. SCHWARZ (eds.) *Well-Being: The Foundations of Hedonic Psychology*. Russell Sage Foundation, New York, 302–329.
- FREDERICK, S., G. F. LOEWENSTEIN and T. O'DONOGHUE (2002): Time Discounting and Time Preference: A Critical Review, *Journal of Economic Literature* 40, 351–401.
- FREY, B. S. (1997): A Constitution for Knaves Crowds Out Civic Virtues, *Economic Journal* 107, 1043–1053.
- FREY, B. S. and R. EICHENBERGER (1994): Economic Incentives Transform Psychological Anomalies, *Journal of Economic Behavior and Organization* 23, 215–234.
- FRIEDMAN, M. (1953): *Essays in Positive Economics*. University of Chicago Press, Chicago et al.

- FRIEDMAN, M. (1971): *Price Theory: A Provisional Text*. Aldine Publishing Company, Chicago, 8th ed.
- FUHRER, J. C. (2000): Habit Formation in Consumption and Its Implications for Monetary-Policy Models, *American Economic Review* 90, 367–390.
- GAERTNER, W. (1974): A Dynamic Model of Interdependent Consumer Behavior, *Zeitschrift Für Nationalökonomie* 34, 327–344.
- GALBRAITH, J. K. (1958): *The Affluent Society*. Riverside Press, Cambridge, MA.
- GAO, L. and U. SCHMIDT (2005): Self Is Never Neutral: Why Economic Agents Behave Irrationally, *Journal of Behavioral Finance* 6, 27–37.
- GEANAKOPOLOS, J., D. PEARCE and E. STACCHETTI (1989): Psychological Games and Sequential Rationality, *Games and Economic Behavior* 1, 60–79.
- GERBER, E. R. and J. E. JACKSON (1993): Endogenous Preferences and the Study of Institutions, *American Political Science Review* 87, 639–656.
- GINTIS, H. (1974): Welfare Criteria with Endogenous Preferences: The Economics of Education, *International Economic Review* 15, 415–430.
- GOODIN, R. E. (1993): Democracy, Preferences and Paternalism, *Policy Sciences* 26, 229–247.
- GORMAN, W. M. (1967): Tastes, Habits and Choices, *International Economic Review* 8, 218–222.
- GRANOVETTER, M. and R. SOONG (1986): Threshold Models of Interpersonal Effects in Consumer Demand, *Journal of Economic Behavior and Organization* 7, 83–99.
- GREYER, D. M. and C. R. PLOTT (1979): Economic Theory of Choice and the Preference Reversal Phenomenon, *American Economic Review* 69, 623–638.
- GROSSKOPF, B. (2003): Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition, *Experimental Economics* 6, 141–158.
- GRÜNE, T. (2004): The Problems of Testing Preference Axioms with Revealed Preference Theory, *Analyse Und Kritik* 26, 382–397.
- GRÜNE-YANOFF, T. (2012): Old Wine in New Casks: Libertarian Paternalism Still Violates Liberal Principles, *Social Choice and Welfare* 38, 635–645.
- GRÜNE-YANOFF, T. and S. O. HANSSON (2009a): Preference Change: An Introduction, in: T. GRÜNE-YANOFF and S. O. HANSSON (eds.) *Preference Change: Approaches From Philosophy, Economics and Psychology*. Springer, Dordrecht et al., 1–26.
- GRÜNE-YANOFF, T. and S. O. HANSSON (2009b): From Belief Revision to Preference Change, in: T. GRÜNE-YANOFF and S. O. HANSSON (eds.) *Preference Change: Approaches From Philosophy, Economics and Psychology*. Springer, Dordrecht et al., 159–184.
- GUARIGLIA, A. and M. ROSSI (2002): Consumption, Habit Formation, and Precautionary Saving: Evidence From the British Household Panel Survey, *Oxford Economic Papers* 54, 1–19.
- GÜTH, W. (1995): On Ultimatum Bargaining Experiments – A Personal Review, *Journal of Economic Behavior and Organization* 27, 329–344.
- HAHNEL, R. and M. ALBERT (1990): *Quiet Revolution in Welfare Economics*. Princeton University Press, Princeton, NJ.
- HAMMOND, P. J. (1976): Endogenous Tastes and Stable Long-Run Choice, *Journal of Economic Theory* 13, 329–340.
- HARSANYI, J. C. (1954): Welfare Economics of Variable Tastes, *Review of Economic Studies* 21, 204–213.
- HARSANYI, J. C. (1982): Morality and the Theory of Rational Behavior, in: A. K. SEN and B. A. O. WILLIAMS (eds.) *Utilitarianism and Beyond*. Cambridge University Press, Cambridge, UK et al., 39–62.

- HAUSMAN, D. M. (2012): *Preference, Value, Choice, and Welfare*. Cambridge University Press, Cambridge, UK et al.
- HEIEN, D. and C. DURHAM (1991): A Test of the Habit Formation Hypothesis Using Household Data, *Review of Economics and Statistics* 73, 189–199.
- HELSON, H. (1948): Adaptation-Level as a Basis for a Quantitative Theory of Frames of Reference, *Psychological Review* 55, 297–313.
- HENRICH, J., R. BOYD, S. BOWLES, C. F. CAMERER, E. FEHR, H. GINTIS and R. MCELREATH (2001): In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies, *American Economic Review, Papers and Proceedings* 91, 73–78.
- HICKS, J. R. (1965): *Capital and Growth*. Oxford University Press, London et al.
- HIRSCH, F. (1995): *Social Limits to Growth*. Routledge, London, revised ed.
- HOCH, S. J. and G. F. LOEWENSTEIN (1991): Time-Inconsistent Preferences and Consumer Self-Control, *Journal of Consumer Research* 17, 492–507.
- HOEFFLER, S. and D. ARIELY (1999): Constructing Stable Preferences: A Look Into Dimensions of Experience and Their Impact on Preference Stability, *Journal of Consumer Psychology* 8, 113–139.
- HOLBROOK, M. B. and R. M. SCHINDLER (1989): Some Exploratory Findings on the Development of Musical Tastes, *Journal of Consumer Research* 16, 119–124.
- HOLLÄNDER, H. (1990): A Social Exchange Approach to Voluntary Cooperation, *American Economic Review* 80, 1157–1167.
- HOUTHAKKER, H. S. (1950): Revealed Preference and the Utility Function, *Economica* 17, 159–174.
- HOUTHAKKER, H. S. and L. D. TAYLOR (1966): *Consumer Demand in the United States, 1929–1970: Analyses and Projections*. Harvard University Press, Cambridge, Mass.
- HSEE, C. K., G. F. LOEWENSTEIN, S. BLOUNT and M. H. BAZERMAN (1999): Preference Reversals Between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis, *Psychological Bulletin* 125, 576–590.
- HUNG, A. A. and C. R. PLOTT (2001): Information Cascades: Replication and Extension to Majority Rule and Conformity-Rewarding Institutions, *American Economic Review* 91, 1508–1520.
- IANNACCONI, L. R. (1986): Addiction and Satiation, *Economics Letters* 21, 95–99.
- IANNACCONI, L. R. (1989): Bandwagons and the Threat of Chaos: Interpersonal Effects Revisited, *Journal of Economic Behavior and Organization* 11, 431–442.
- IRELAND, N. J. (2001): Optimal Income Tax in the Presence of Status Effects, *Journal of Public Economics* 81, 193–212.
- ISONI, A., G. LOOMES and R. SUGDEN (2011): The Willingness to Pay – Willingness to Accept Gap, the “Endowment Effect”, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment, *American Economic Review* 101, 991–1011.
- JEFFREY, R. C. (1974): Preference Among Preferences, *Journal of Philosophy* 71, 377–391.
- JOHNSON, E. J., G. HÄUBL and A. KEINAN (2007): Aspects of Endowment: A Query Theory of Value Construction, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33, 461–474.
- KAGEL, J. H. and K. W. WOLFE (2001): Tests of Fairness Models Based on Equity Considerations in a Three-Person Ultimatum Game, *Experimental Economics* 4, 203–219.
- KAHNEMAN, D., J. L. KNETSCH and R. H. THALER (1990): Experimental Tests of the Endowment Effect and the Coase Theorem, *Journal of Political Economy* 98, 1325–1348.

- KAHNEMAN, D., J. L. KNETSCH and R. H. THALER (1991): Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, *Journal of Economic Perspectives* 5(1), 193–206.
- KAHNEMAN, D. and J. SNELL (1992): Predicting a Changing Taste: Do People Know What They Will Like?, *Journal of Behavioral Decision Making* 5, 187–200.
- KAHNEMAN, D. and A. TVERSKY (1979): Prospect Theory: An Analysis of Decision Under Risk, *Econometrica* 47, 263–291.
- KAHNEMAN, D., P. P. WAKKER and R. SARIN (1997): Back To Bentham? Explorations of Experienced Utility, *Quarterly Journal of Economics* 112, 375–405.
- KAPTEYN, A., T. WANSBEEK and J. BUYZE (1978): The Dynamics of Preference Formation, *Economics Letters* 1, 93–98.
- KAPTEYN, A., S. VAN DE GEER, H. VAN DE STADT and T. WANSBEEK (1997): Interdependent Preferences: An Econometric Analysis, *Journal of Applied Econometrics* 12, 665–686.
- KARNI, E. and D. SCHMEIDLER (1990): Fixed Preferences and Changing Tastes, *American Economic Review, Papers and Proceedings* 80, 262–267.
- KAUFMAN, B. E. (1999): Emotional Arousal as a Source of Bounded Rationality, *Journal of Economic Behavior and Organization* 38, 135–144.
- KNETSCH, J. L. (1989): The Endowment Effect and Evidence of Nonreversible Indifference Curves, *American Economic Review* 79, 1277–1284.
- KNETSCH, J. L. and J. A. SINDEN (1984): Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value, *Quarterly Journal of Economics* 99, 507–521.
- KNETSCH, J. L. and W.-K. WONG (2009): The Endowment Effect and the Reference State: Evidence and Manipulations, *Journal of Economic Behavior and Organization* 71, 407–413.
- KONOW, J. (2000): Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions, *American Economic Review* 90, 1072–1091.
- KOO, A. Y. C. (1963): An Empirical Test of Revealed Preference Theory, *Econometrica* 31, 646–664.
- KRELLE, W. (1972): Dynamisierung Der Nutzenfunktion, *Zeitschrift Für Nationalökonomie* 32, 59–70.
- KÖSZEGI, B. and M. RABIN (2006): A Model of Reference-Dependent Preferences, *Quarterly Journal of Economics* 121, 1133–1165.
- KÖSZEGI, B. and M. RABIN (2008): Choices, Situations, and Happiness, *Journal of Public Economics* 92, 1821–1832.
- KÖSZEGI, B. and A. SZEIDL (2013): A Model of Focusing in Economic Choice, *Quarterly Journal of Economics* 128, 53–104.
- LAIBSON, D. (1997): Golden Eggs and Hyperbolic Discounting, *Quarterly Journal of Economics* 112, 443–477.
- LAIBSON, D. (2001): A Cue-Theory of Consumption, *Quarterly Journal of Economics* 116, 81–119.
- LANCASTER, K. J. (1966): A New Approach to Consumer Theory, *Journal of Political Economy* 74, 132–157.
- LANDSBURG, S. E. (1981): Taste Change in the United Kingdom, 1900–1955, *Journal of Political Economy* 89, 92–104.
- LEITH, K. P. and R. F. BAUMEISTER (1996): Why Do Bad Moods Increase Self-Defeating Behavior? Emotion, Risk Taking, and Self-Regulation, *Journal of Personality and Social Psychology* 71, 1250–1267.
- LENER, J. S., D. A. SMALL and G. F. LOEWENSTEIN (2004): Heart Strings and Purse Strings: Carryover Effects of Emotions on Economic Decisions, *Psychological Science* 15, 337–341.

- LEVINE, D. K. (1998): Modeling Altruism and Spitefulness in Experiments, *Review of Economic Dynamics* 1, 593–622.
- LICHTENSTEIN, S. and P. SLOVIC (1971): Reversals of Preference Between Bids and Choices in Gambling Decisions, *Journal of Experimental Psychology* 89, 46–55.
- LINDMAN, H. R. (1971): Inconsistent Preferences Among Gambles, *Journal of Experimental Psychology* 89, 390–397.
- LIST, J. A. (2003): Does Market Experience Eliminate Market Anomalies?, *Quarterly Journal of Economics* 118, 41–71.
- LIST, J. A. (2004): Neoclassical Theory Versus Prospect Theory: Evidence From the Marketplace, *Econometrica* 72, 615–625.
- LIST, J. A. (2006): Using Hicksian Surplus Measures to Examine Consistency of Individual Preferences: Evidence From a Field Experiment, *Scandinavian Journal of Economics* 108, 115–134.
- LLUCH, C. (1974): Savings and Habit Formation, *International Economic Review* 15, 786–797.
- LOEWENSTEIN, G. F. (1988): Frames of Mind in Intertemporal Choice, *Management Science* 34, 200–214.
- LOEWENSTEIN, G. F. (1996): Out Of Control: Visceral Influences on Behavior, *Organizational Behavior and Human Decision Processes* 65, 272–292.
- LOEWENSTEIN, G. F. (2000): Emotions in Economic Theory and Economic Behavior, *American Economic Review, Papers and Proceedings* 90, 426–432.
- LOEWENSTEIN, G. F. and D. ADLER (1995): A Bias in the Prediction of Tastes, *Economic Journal* 105, 929–937.
- LOEWENSTEIN, G. F., T. O'DONOGHUE and M. RABIN (2003): Projection Bias in Predicting Future Utility, *Quarterly Journal of Economics* 118, 1209–1248.
- LOEWENSTEIN, G. F., L. THOMPSON and M. H. BAZERMAN (1989): Social Utility and Decision Making in Interpersonal Contexts, *Journal of Personality and Social Psychology* 57, 426–441.
- LOOMES, G. and R. SUGDEN (1982): Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty, *Economic Journal* 92, 805–824.
- LOOMES, G. and R. SUGDEN (1983): A Rationale for Preference Reversal, *American Economic Review* 73, 428–432.
- LÖNNQVIST, J.-E., M. VERKASALO and G. WALKOWITZ (2011): It Pays to Pay – Big Five Personality Influences on Co-Operative Behaviour in an Incentivized and Hypothetical Prisoner's Dilemma Game, *Personality and Individual Differences* 50, 300–304.
- MANDLER, M. (2004): Status Quo Maintenance Reconsidered: Changing or Incomplete Preferences?, *Economic Journal* 114, F518–F535.
- MARCH, J. G. (1978): Bounded Rationality, Ambiguity, and the Engineering of Choice, *Bell Journal of Economics* 9, 587–608.
- MARCUSE, H. (1964): *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Routledge and Kegan Paul, London.
- MARKOWITZ, H. (1952): The Utility of Wealth, *Journal of Political Economy* 60, 151–158.
- MARKUS, H. R. and S. KITAYAMA (1991): Culture and the Self: Implications for Cognition, Emotion, and Motivation, *Psychological Review* 98, 224–253.
- MARSHAK, T. A. (1978): On the Study of Taste Changing Policies, *American Economic Review, Papers and Proceedings* 68, 386–391.
- MARSHALL, A. (1930): *Principles of Economics*. Macmillan and Co, London et al., 8th ed.
- MARX, K. (1902): *Wage-Labor and Capital*. New York Labor News Company, New York.

- MARX, K. (1992): *The Poverty of Philosophy*. International Publishers, New York.
- MARZALLI ERICSON, K. M. and A. FUSTER (2011): Expectations as Endowments: Evidence on Reference-Dependent Preferences From Exchange and Valuation Experiments, *Quarterly Journal of Economics* 126, 1879–1907.
- MCCARTHY, M. D. (1974): On the Stability of Dynamic Demand Functions, *International Economic Review* 15, 256–259.
- MCCRAE, R. R. and O. P. JOHN (1992): An Introduction to the Five-Factor Model and Its Applications, *Journal of Personality* 60, 175–215.
- MEGHIR, C. and G. WEBER (1996): Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis Using a U.S. Consumption Panel, *Econometrica* 64, 1151–1181.
- MERRITT, A. C., D. A. EFFRON and B. MONIN (2010): Moral Self-Licensing: When Being Good Frees Us To Be Bad, *Social and Personality Psychology Compass* 4/5, 344–357.
- MICHAEL, R. T. and G. S. BECKER (1973): On the New Theory of Consumer Behavior, *Swedish Journal of Economics* 75, 378–396.
- MIRRELES, J. A. (1982): The Economic Uses of Utilitarianism, in: A. K. SEN and B. A. O. WILLIAMS (eds.) *Utilitarianism and Beyond*. Cambridge University Press, Cambridge, UK et al., 63–84.
- MUELLBAUER, J. (1988): Habits, Rationality and Myopia in the Life Cycle Consumption Function, *Annales d'Économie Et De Statistique* 9, 47–70.
- MUNRO, A. and R. SUGDEN (2003): On the Theory of Reference-Dependent Preferences, *Journal of Economic Behavior and Organization* 50, 407–428.
- NAIK, N. Y. and M. J. MOORE (1996): Habit Formation and Intertemporal Substitution in Individual Food Consumption, *Review of Economics and Statistics* 78, 321–328.
- NELSON, P. (1970): Information and Consumer Behavior, *Journal of Political Economy* 78, 311–329.
- NG, Y.-K. and J. WANG (2001): Attitude Choice, Economic Change, and Welfare, *Journal of Economic Behavior and Organization* 45, 279–291.
- OCKENFELS, A. and J. WEIMANN (1999): Types and Patterns: An Experimental East-West-German Comparison of Cooperation and Solidarity, *Journal of Public Economics* 71, 275–287.
- OXOBY, R. J. (2003): Attitudes and Allocations: Status, Cognitive Dissonance, and the Manipulation of Attitudes, *Journal of Economic Behavior and Organization* 52, 365–385.
- OXOBY, R. J. (2004): Cognitive Dissonance, Status and Growth of the Underclass, *Economic Journal* 114, 727–749.
- PASHARDES, P. (1986): Myopic and Forward Looking Behavior in a Dynamic Demand System, *International Economic Review* 27, 387–397.
- PAYNE, J. W., J. R. BETTMAN and E. J. JOHNSON (1992): Behavioral Decision Research: A Constructive Processing Perspective, *Annual Review of Psychology* 43, 87–131.
- PESENDORFER, W. (1995): Design Innovation and Fashion Cycles, *American Economic Review* 85, 771–792.
- PIGOU, A. C. (1913): The Interdependence of Different Sources of Demand and Supply in a Market, *Economic Journal* 23, 19–24.
- PLOTT, C. R. (1996): Rational Individual Behaviour in Markets and Social Choice Processes: The Discovered Preference Hypothesis, in: K. J. ARROW, E. COLOMBATTO, M. PERLMAN and C. SCHMIDT (eds.) *The Rational Foundations of Economic Behavior*. Macmillan Press, Houndmills et al., 225–250.
- PLOTT, C. R. and K. ZEILER (2005): The Willingness to Pay – Willingness to Accept Gap, the “Endowment Effect”, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations, *American Economic Review* 95, 530–545.

- PLOTT, C. R. and K. ZEILER (2007): Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?, *American Economic Review* 97, 1449–1466.
- PLOTT, C. R. and K. ZEILER (2011): The Willingness to Pay – Willingness to Accept Gap, the “Endowment Effect”, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Reply, *American Economic Review* 101, 1012–1028.
- POLKOVNICHENKO, V. (2007): Life-Cycle Portfolio Choice with Additive Habit Formation Preferences and Uninsurable Labor Income Risk, *Review of Financial Studies* 20, 83–124.
- POLLAK, R. A. (1970): Habit Formation and Dynamic Demand Functions, *Journal of Political Economy* 78, 745–763.
- POLLAK, R. A. (1976a): Interdependent Preferences, *American Economic Review* 66, 309–320.
- POLLAK, R. A. (1976b): Habit Formation and Long-Run Utility Functions, *Journal of Economic Theory* 13, 272–297.
- POSTLEWAITE, A. (1998): The Social Basis of Interdependent Preferences, *European Economic Review* 42, 779–800.
- PRASNIKAR, V. and A. E. ROTH (1992): Considerations of Fairness and Strategy: Experimental Data From Sequential Games, *Quarterly Journal of Economics* 107, 865–888.
- RABIN, M. (1993): Incorporating Fairness Into Game Theory and Economics, *American Economic Review* 83, 1281–1302.
- RABIN, M. (1994): Cognitive Dissonance and Social Change, *Journal of Economic Behavior and Organization* 23, 177–194.
- RABIN, M. (1998): Psychology and Economics, *Journal of Economic Literature* 36, 11–46.
- RAUSCHER, M. (1992): Keeping up with the Joneses, *Economics Letters* 40, 287–290.
- RAUSCHER, M. (1993): Demand for Social Status and the Dynamics of Consumer Behavior, *Journal of Socio-Economics* 22, 105–113.
- RAVN, M., S. SCHMITT-GROHÉ and M. URIBE (2006): Deep Habits, *Review of Economic Studies* 73, 195–218.
- READ, D. and B. VAN LEEUWEN (1998): Predicting Hunger. the Effects of Appetite and Delay on Choice, *Organizational Behavior and Human Decision Processes* 76, 189–205.
- ROTH, A. E. (1995): Bargaining Experiments, in: J. H. KAGEL and A. E. ROTH (eds.) *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, 253–348.
- ROTH, A. E., V. PRASNIKAR, M. OKUNO-FUJIWARA and S. ZAMIR (1991): Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An Experimental Study, *American Economic Review* 81, 1068–1095.
- RUNCIMAN, W. G. (1966): *Relative Deprivation and Social Justice*. Routledge, London.
- RYDER, H. E. JR and G. M. HEAL (1973): Optimal Growth with Intertemporally Dependent Preferences, *Review of Economic Studies* 40, 1–31.
- SAMUELSON, P. A. (1937): A Note on Measurement of Utility, *Review of Economic Studies* 4, 155–161.
- SAMUELSON, P. A. (1938): A Note on the Pure Theory of Consumer’s Behaviour, *Economica* 5, 61–71.
- SAMUELSON, L. (2001): Introduction to the Evolution of Preferences, *Journal of Economic Theory* 97, 225–230.
- SASS, M., F. TIMME and J. WEIMANN (2015): The Dynamics of Dictator Behavior. CESifo Discussion Paper No. 5348.
- SASS, M. and J. WEIMANN (2012): The Dynamics of Individual Preferences in Repeated Public Good Experiments. University of Magdeburg, FEMM Working Paper No. 02/2012.

- SASS, M. and J. WEIMANN (2015a): Moral Self-Licensing and the Direct Touch Effect. CESifo Discussion Paper No. 5174.
- SASS, M. and J. WEIMANN (2015b): The Peculiar Power of Pairs. CESifo Discussion Paper No. 5246.
- SCHELLING, T. C. (1984): Self-Command in Practice, in Policy, and in a Theory of Rational Choice, *American Economic Review, Papers and Proceedings* 74, 1–11.
- SCHICK, F. (1984): *Having Reasons*. Princeton University Press, Princeton.
- SCHICK, F. (1986): Dutch Bookies and Money Pumps, *Journal of Philosophy* 83, 112–119.
- SCHINDLER, R. M. and M. B. HOLBROOK (1993): Critical Periods in the Development of Men's and Women's Tastes in Personal Appearance, *Psychology and Marketing* 10, 549–564.
- SCHMIDT, U. and H. ZANK (2012): A Genuine Foundation for Prospect Theory, *Journal of Risk and Uncertainty* 45, 97–113.
- SCHMITZ, J. (2015): One Good Deed a Day? SSRN Discussion Paper No. 2538410.
- SCITOVSKY, T. (1992): *The Joyless Economy*. Oxford University Press, New York et al., revised ed.
- SCOTT, R. H. (1972): Avarice, Altruism, and Second Party Preferences, *Quarterly Journal of Economics* 86, 1–18.
- SEN, A. K. (1973): Behaviour and the Concept of Preference, *Economica* 40, 241–259.
- SEN, A. K. (1977): Rational Fools: A Critique of the Behavioral Foundations of Economic Theory, *Philosophy and Public Affairs* 6, 317–344.
- SEN, A. K. (1993): Internal Consistency of Choice, *Econometrica* 61, 495–521.
- SEN, A. K. (1994): The Formulation of Rational Choice, *American Economic Review, Papers and Proceedings* 84, 385–390.
- SEN, A. K. (1997): Maximization and the Act of Choice, *Econometrica* 65, 745–779.
- SIMON, H. A. (1955): A Behavioral Model of Rational Choice, *Quarterly Journal of Economics* 69, 99–118.
- SIPPEL, R. (1997): An Experiment on the Pure Theory of Consumer's Behaviour, *Economic Journal* 107, 1431–1444.
- SLOVIC, P. (1995): The Construction of Preference, *American Psychologist* 50, 364–371.
- SLOVIC, P. and A. TVERSKY (1974): Who Accepts Savage's Axioms?, *Behavioral Science* 19, 368–373.
- SMITH, A. (1976): *The Theory of Moral Sentiments*. Edited by D. D. RAPHAEL and A. L. MACFIE. Oxford University Press, Oxford et al.
- SNELL, J., B. J. GIBBS and C. VAREY (1995): Intuitive Hedonics: Consumer Beliefs About the Dynamics of Liking, *Journal of Consumer Psychology* 4, 33–60.
- SOBEL, J. (2005): Interdependent Preferences and Reciprocity, *Journal of Economic Literature* 43, 392–436.
- SOBEL, J. (2009): Generous Actors, Selfish Actions: Markets with Other-Regarding Preferences, *International Review of Economics* 56, 3–16.
- SOLDZ, S. and G. E. VAILLANT (1999): The Big Five Personality Traits and the Life Course: A 45-Year Longitudinal Study, *Journal of Research in Personality* 33, 208–232.
- SOLNICK, S. J. and D. HEMENWAY (1998): Is More Always Better? A Survey on Positional Concerns, *Journal of Economic Behavior and Organization* 37, 373–383.
- SPINNEWYN, F. (1981): Rational Habit Formation, *European Economic Review* 15, 91–109.
- STIGLER, G. J. and G. S. BECKER (1977): De Gustibus Non Est Disputandum, *American Economic Review* 67, 76–90.
- STRAHAN, E. J., S. J. SPENCER and M. P. ZANNA (2002): Subliminal Priming and Persuasion: Striking While the Iron Is Hot, *Journal of Experimental Social Psychology* 38, 556–568.

- STROTZ, R. H. (1956): Myopia and Inconsistency in Dynamic Utility Maximization, *Review of Economic Studies* 23, 165–180.
- SUGDEN, R. (1985): Why Be Consistent? A Critical Analysis of Consistency Requirements in Choice Theory, *Economica* 52, 167–183.
- SUGDEN, R. (2004): The Opportunity Criterion: Consumer Sovereignty Without the Assumption of Coherent Preferences, *American Economic Review* 94, 1014–1033.
- SUGDEN, R. (2008): Why Incoherent Preferences Do Not Justify Paternalism, *Constitutional Political Economy* 19, 226–248.
- SUNDARESAN, S. M. (1989): Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth, *Review of Financial Studies* 2, 73–89.
- SUNSTEIN, C. R. (1986): Legal Interference with Private Preferences, *University of Chicago Law Review* 53, 1129–1174.
- SUNSTEIN, C. R. and R. H. THALER (2003): Libertarian Paternalism Is Not an Oxymoron, *University of Chicago Law Review* 70, 1159–1202.
- THALER, R. H. (1980): Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization* 1, 39–60.
- THALER, R. H. and C. R. SUSTEIN (2009): *Nudge*. Penguin Books, London.
- TVERSKY, A. and D. KAHNEMAN (1974): Judgment Under Uncertainty: Heuristics and Biases, *Science* 185, 1124–1131.
- TVERSKY, A. and D. KAHNEMAN (1981): The Framing of Decisions and the Psychology of Choice, *Science* 211, 453–458.
- TVERSKY, A. and D. KAHNEMAN (1991): Loss Aversion in Riskless Choice: A Reference-Dependent Model, *Quarterly Journal of Economics* 106, 1039–1061.
- TVERSKY, A. and I. SIMONSON (1993): Context-Dependent Preferences, *Management Science* 39, 1179–1189.
- TVERSKY, A., P. SLOVIC and D. KAHNEMAN (1990): The Causes of Preference Reversal, *American Economic Review* 80, 204–217.
- TVERSKY, A. and R. H. THALER (1990): Anomalies: Preference Reversals, *Journal of Economic Perspectives* 4(2), 201–211.
- VAN DE KUILEN, G. (2009): Subjective Probability Weighting and the Discovered Preference Hypothesis, *Theory and Decision* 67, 1–22.
- VAN DE KUILEN, G. and P. P. WAKKER (2006): Learning in the Allais Paradox, *Journal of Risk and Uncertainty* 33, 155–164.
- VEBLEN, T. (1934): *The Theory of the Leisure Class: An Economic Study of Institutions*. Modern Library, New York.
- VOLK, S., C. THÖNI and W. RUIGROK (2012): Temporal Stability and Psychological Foundations of Cooperation Preferences, *Journal of Economic Behavior and Organization* 81, 664–676.
- VON WEIZSÄCKER, C. C. (1971): Notes on Endogenous Change of Tastes, *Journal of Economic Theory* 3, 345–372.
- WATHIEU, L. (1997): Habits and the Anomalies in Intertemporal Choice, *Management Science* 43, 1552–1563.
- WEISBROD, B. A. (1977): Comparing Utility Functions in Efficiency Terms or, What Kind of Utility Function Do We Want?, *American Economic Review* 67, 991–995.
- WELSCH, H. (2005): Adaptation of Tastes to Constraints, *Theory and Decision* 57, 379–395.
- WEST, P. M., C. L. BROWN and S. J. HOCH (1996): Consumption Vocabulary and Preference Formation, *Journal of Consumer Research* 23, 120–135.

- WINKIELMAN, P., K. C. BERRIDGE and J. L. WILBARGER (2005): Unconscious Affective Reactions to Masked Happy Versus Angry Faces Influence Consumption Behavior and Judgments of Value, *Personality and Social Psychology Bulletin* 31, 121–135.
- WINSTON, G. C. (1980): Addiction and Backsliding: A Theory of Compulsive Consumption, *Journal of Economic Behavior and Organization* 1, 295–324.
- WRIGHT, J. D. and D. H. GINSBURG (2012): Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty, *Northwestern University Law Review* 106, 1033–1088.
- ZAJONC, R. B. (1968): Attitudinal Effects of Mere Exposure, *Journal of Personality and Social Psychology, Monograph Supplement* 9, 1–27.
- ZAJONC, R. B. (2001): Mere Exposure: A Gateway to the Subliminal, *Current Directions in Psychological Science* 10, 224–228.
- ZIZZO, D. J. (2003): Empirical Evidence on Interdependent Preferences: Nature or Nurture?, *Cambridge Journal of Economics* 27, 867–880.
- ZIZZO, D. J. and A. J. OSWALD (2001): Are People Willing to Pay to Reduce Others' Incomes?, *Annales d'Économie Et De Statistique* 64, 39–65.