Review of Behavioral Economics, 2018, 5: 221–241

Toward A Behavioral Foundation of Normative Economics

Malte F. Dold[1] and Christian Schubert[2]

[1] *New York University, 19 West 4th Street, New York, NY 10012; malte.dold@nyu.edu*
[2] *German University in Cairo, 11835 New Cairo, Cairo, Egypt; christian.schubert@guc.edu.eg*

ABSTRACT

While behavioral economics has had a major impact on positive theorizing in economics, it remains unclear what exactly those new insights about deviations from rational choice mean in terms of policy implications. Given the ever-rising interest in the new, psychologically informed economics, this paper outlines the way in which reasoning about the normative implications of behavioral economics has developed in the last decade. We argue that behavioral economics has inspired new thinking about the prospect of 'behavioral normative economics' (BNE). The paper sketches important approaches in the field, discusses their theoretical shortcomings, and outlines some initial ideas on how to conceptualize individuals' identity as a key task in BNE. We suggest that the dualistic concepts of the individual should be abandoned in favor of a notion of a unified self that is constituted by its capacity to learn and reflect upon new preferences on a continuous basis.

Keywords: Behavioral, Identity, Nudging, Normative Economics, Welfare

JEL Codes: B41, D04, P46

1 Introduction

Given that economics has rediscovered the often intricate and surprising ways in which real-world agents think and choose, it is time to systematically assess the meaning of these behavioral insights for policy prescription. We will argue that normative economics (i.e., the branch of economics that is concerned with

ISSN 2326-6198; DOI 10.1561/105.00000097

welfare and policy advice), can only benefit from a more realistic underlying model of human conduct. We will try to clarify the ways in which this is the case and critically review existing approaches in, what we will refer to as, 'behavioral normative economics' (henceforth BNE).

In general, BNE can be conceptualized as the set of all attempts to modify standard normative economics so that it is better aligned – or 'reconciled' (McQuillin and Sugden, 2012) – with insights from behavioral economics and can be coherently applied in a *behavioral world*, i.e., a world where individuals have limited computational capacities, attention, and willpower. The project begins from the basic insight that economists' standard notion of welfare (which defines welfare, essentially, as the degree of satisfaction of given consistent preferences) cannot easily apply to a behavioral setting where preferences are often incomplete and unstable (Hausman and McPherson, 2009).

We will argue that the key query shaping future developments of BNE (and behavioral policy advice derived from it) concerns the underlying conception of *identity*. Do agents whose choices are marred by cognitive biases consist of multiple selves among which one – the cool 'planner' or the hot, spontaneous 'doer' – needs to be privileged? Or should we try to overcome this modeling approach–presumed in most contemporary contributions to the literature on BNE – in favor of a procedural understanding of identity as a person's idiosyncratic understanding and valuation of her preference formation process? Clearly, questions like this beget the need to rekindle the dialogue between (behavioral) economists and ethicists.

We begin with an exploration of recent attempts to elaborate upon a purportedly 'pragmatic' way to incorporate behavioral findings into policy advice (Section 2). Given the shortcomings of such previous attempts, we will proceed to discuss the branch of BNE that is currently most popular within economics, 'behavioral welfare economics' (Section 3). We then discuss one particularly controversial offshoot of that approach known as 'libertarian paternalism'. Section 4 presents a radical alternative approach to BNE: Robert Sugden's 'opportunity criterion'. While his approach seems promising, his criterion ultimately retains the static character of the mainstream approach. Our penultimate Section 5 sketches a more dynamic proposal of 'preference learning' whose elements may serve as vital building blocks for future theories in the field of BNE. Our final section identifies policy implications and promising avenues for further research.

2 Behavioral Policy Advice: The Pragmatic Way

From the viewpoint of a practicing economist, the simplest way to incorporate behavioral findings into public policy is what we refer to as the 'pragmatic' way. Such a position takes policy objectives as given and answers questions about

how best to satisfy them by incorporating behavioral insights (e.g., about the way in which real humans react to economic incentives in an experimental setting) without questioning the underlying normative premises. Following Keynes' (1917) famous distinction, the 'pragmatists' favor an instrumental understanding of economics (in Keynes' terminology: "the art of political economy") as opposed to normative economic theorizing (in Keynes' terminology: "applied ethics"). While the latter examines alternative policy-goals and the applicability of welfare criteria, the former only assesses policies in light of given ends. Prominent advocates of this approach are Bolton and Ockenfels (2012), among others, who refer to it as 'behavioral economic engineering'.[1]

For example, a kindergarten director wishing to reduce the number of parents who are late in picking up their children would be wise not to introduce a monetary fee as this has been shown to crowd out certain social norms, thus generating counterproductive behavioral effects (Gneezy and Rustichini, 2000). In this and many other contexts, standard economic analysis is not sufficient to identify the sources of impediments to market efficiency since psychological and cognitive biases of consumers, and not their rational behavior, are assumed to be the main driver for 'behavioral market inefficiencies' (Madrian, 2014, p. 681).

The engineering approach views standard regulation policies with economic incentives (e.g., taxes or subsidies) and behaviorally inspired interventions (e.g., default rules or cooling-off-periods) as complementary entities. Chetty states that "the decision to include behavioral factors in economic models should be viewed as a pragmatic rather than philosophical choice" (2015, p. 28). Therefore, theoretical analysis in behavioral economic engineering adjusts standard economic models only *at the margin* (e.g., modelling boundedly rational agents equipped with social preferences).

This method of incorporating behavioral findings into policy advice aims at making economic institutions more robust with regard to the presupposed normative ideal of 'rationality,' in the standard economic sense of consistency. Behaviorally enriched models can inform policy-makers about the effectiveness of standard economic measures to reach given policy goals and contribute to the development of a new set of policy tools that are more cost-effective compared to traditional ones (Madrian, 2014, p. 663). Importantly, this approach also puts emphasis on gathering experimental evidence. Bolton and Ockenfels (2012, p. 667) state, "the objective behind behavioral economic engineering is to catch problems in the robustness of new economic institutions and mechanisms prior to their engagement on a large social scale, where failure can be far costlier."

While this pragmatic approach appears attractive at first sight, it is important to realize its limits. Essentially, it leaves out two key questions: *Where do standard economic policy goals come from? Is it the economist's job to*

[1]For related 'pragmatist' approaches, see Gul and Pesendorfer (2007), Madrian (2014), Bhargava and Loewenstein (2015), and Chetty (2015).

critically discuss those goals as well? We argue that a majority of economists would likely follow Gul and Pesendorfer (2007) and avoid the intricate terrain of questioning the underlying welfare assumptions of policy goals. Depending on one's own answer to the second question, above, there are essentially three ways to move ahead.

First, one may use the 'behavioral economic engineering' toolbox merely to investigate the hypothetical effectiveness of alternative instruments to achieve any exogenously given goal. Typically, that goal may be 'imported' from the policy-makers' realm and make more or less economic sense. However, even this approach relies on some conception of 'welfare' in order to compare alternative instruments as to their 'welfare effects.' Thus, proponents of this approach will typically resort to the notion of welfare developed by standard welfare economics (i.e., welfare as the degree of satisfaction of given and perfectly consistent preferences) or to a measure of subjective well-being (as, e.g., Chetty, 2015). As we discuss below, this limits the range of issues to which this approach can be coherently applied. Second, one might find that a critical discussion of policy goals is part of the economist's job description and argue that policy should focus upon goals that maximize 'welfare' – or install market efficiency – in the sense of standard welfare economics. In this scenario, the same caveat regarding the range of applications applies. Third, one may be skeptical about standard welfare goals (i.e., preference consistency at the individual level and market efficiency at the aggregated level) and argue that valuable properties of markets are grounded in non-standard welfare considerations. Then, morally valuable properties of markets might require much less in terms of individual rationality and allow for incoherent preferences on behalf of consumers. Sugden (2004) argues that the distinctive prerequisite for markets to produce the classical liberal values of *individual opportunities* and *responsibility* – which, for Sugden, are the basis for a morally integrated individual – is not the rationality of consumers, but just their price sensitivity.[2]

The following sections elaborate on the third possibility of non-standard welfare considerations. We begin with a discussion of the approach that is closest to standard accounts in economics, i.e., 'behavioral welfare economics' with its policy paradigm of 'libertarian paternalism'. Then, we contrast it with Robert Sugden's alternative account that deviates radically from the mainstream in that it argues for giving up rationality assumptions in economic welfare analyses. In spite of its many promising features, we argue that Sugden's accounts still suffers from a questionable notion of individuals as 'responsible agents'. Lastly, we sketch our own proposal of the non-standard account to welfare economics that focuses on procedural aspects of preference learning and identity formation.

[2]Becker (1962) makes a similar point and argues that the welfare-producing property of markets is much more compatible with 'irrational behavior' than is commonly assumed in economic theory. Becker distinguishes between market outcomes and individual behavior and places rationality at the market level, not at the level of individual preferences.

3 Behavioral Welfare Economics: Laundering Preferences

If we presuppose, in theory, that it is the economist's legitimate job to critically discuss alternative policy goals, then the question of BNE enters the discussion. We then face the challenge of reconciling normative economics with behavioral insights. In general, one can again distinguish three different ways to approach this query. In the first approach, we can attempt to salvage standard welfare economics with its foundation in the revealed preference approach. Alternatively, we can abandon the traditional focus of normative economics on welfare and redirect our attention to some other valuable properties of markets (e.g., autonomy or opportunities). Lastly, we may attempt to take a step back and focus on the way in which real-world individuals understand and formulate their preferences (i.e., their identity or character) in the first place. In this and the following sections, we will review these various strategies. We begin by examining the first strategy that attempts to salvage traditional welfare economics.

3.1 The Basic Idea

The approach of 'behavioral welfare economics', most prominently advanced by (Bernheim and Rangel, 2007; 2008; 2009), is so widespread in the literature that it can safely be viewed as the dominant approach to BNE.[3] Bernheim and Rangel acknowledge the fact that the individual preferences of real-world individuals are often incongruent with the 'well-behaved' coherence assumed in the textbook. Thus, they suggest identifying, among the incoherent set of an agent's preferences, the subset whose satisfaction indicates a gain in well-being for that agent. Within the confines of that subset, observed choices are respected, which, according to the authors, establishes sound classical liberal credentials. In essence, "choices provide appropriate guidance because they are choices, not because they reflect something else" (Bernheim and Rangel, 2008, p. 52).[4]

The challenge here is that the behavioral welfare economist must first identify the subset of preferences whose satisfaction is taken to reliably increase an agent's well-being. In other words, the messy set of actual 'manifest' preferences must be 'laundered' first in order to reconstruct what the agent 'really' prefers (or what she should prefer), were she free from the cognitive

[3]See Infante *et al.* (2016) for an overview. They also find this approach, with minor variations, in Bleichrodt *et al.* (2001), Kőszegi and Rabin (2008), and Salant and Rubinstein (2008). Manzini and Mariotti (2014) provide a thorough critique.

[4]Bernheim and Rangel (2008) give two arguments in support of this choice-centered approach: One, taking choice as the normative welfare criterion is a commitment to individual autonomy as a good in itself. The other is that (in spite of all the behavioral anomalies) revealed choices are still the best among all the deficient welfare measures policy-makers might take as a compass to detect individual well-being.

biases that make her preferences inconsistent in the first place. Technically, Bernheim and Rangel resort to a *multiple-selves methodology* that allows them to model the observed choices as maximization under constraints, but with a twist. The constraints include all sorts of contextual factors, or framing effects, that are deemed normatively irrelevant in the economics textbook. Hence, a given agent activates different utility functions (i.e., different 'selves'), depending upon the specific situational context that she finds herself in. This may result in different choices even when the option set, relative prices, and budget constraints remain unchanged. Importantly, a good 'A' is only deemed superior to an alternative good 'B' if 'A' is preferred over 'B' by *all* of the agent's 'selves.' In other words, revealed preferences are being respected as normatively relevant if and only if they are robust under all conceivable contextual constraints and have the standard well-defined structure. Consequently, Bernheim and Rangel's approach rests upon the assumption that "each person has a neoclassical agent deep within himself which is struggling to surface" (Whitman and Rizzo, 2015, p. 423). Individual decisions are deemed to be erroneous if they fail to produce choices that fulfil the rationality axioms of transitivity and completeness. Put differently, a decision error is not independently defined in terms of well-being or any other sensible normative criterion. Rather, an 'error' results from any choice that does not make standard choice theory descriptively accurate. Hence, rather than providing a constructive answer to the questions raised by the reconciliation problem, behavioral welfare economics sidesteps it.

Moreover, this approach ranks states only when the 'intrapersonal Pareto condition', described above, is satisfied. However, it is not immediately clear why, from a normative point of view, a person should not have different preference rankings at different points in time. One might change her lifestyle or advance her moral values, which could produce incoherence between the different preference rankings when judged from the viewpoint of a 'unitary self'. Why should we assume such a unitary meta-perspective rather than an evolutionary view of the self (Hargreaves Heap, 2013, p. 989)?

3.2 An Application: Libertarian Paternalism

One prominent practical application of 'laundering' manifest preferences in order to reconstruct allegedly 'true' preferences has been suggested by Thaler and Sunstein (2003) and Thaler and Sunstein (2008). Their approach of libertarian paternalism, a normative public policy program, aims at improving individuals' welfare without interfering with their freedom of choice.[5] It does

[5]The term 'libertarian paternalism' is somewhat misleading in that there are good reasons to question its 'libertarian' character (e.g. Grüne-Yanoff, 2012); beyond that, only a subset of all the applications suggested in books such as Thaler and Sunstein (2008) and Sunstein (2014) are actually paternalistic in nature. Some of them aim at internalizing

so through the use of nudges (i.e., subtle modifications of situational contexts, one's 'choice architecture') that work by exploiting cognitive biases or by responding to them.[6] The 'nudge agenda' has become immensely influential among policy-makers, particularly in the Anglo-Saxon world (Oliver, 2015; OECD, 2017).

The ubiquity of libertarian paternalism does not warrant a detailed description and critique in this paper.[7] The point here is to emphasize that its advocates follow the general strategy pursued by behavioral welfare economics: to 'launder' people's actual preferences before letting them enter the individual (or social) welfare calculus. Sunstein and Thaler (2003, p. 1162) are explicit in identifying those choices as 'inferior' and in need of correction that "[people] would change if they had complete information, unlimited cognitive abilities, and no lack of self-control." The underlying idea is, similar to the starting point of Bernheim and Rangel, to take homo economicus as the ideal decision maker that guides the public policy agenda. Thus, the libertarian paternalists presume that choosing a specific outcome (e.g., healthy food, non-smoking, or regular contributions to one's retirement savings account) is welfare-increasing if it corresponds to what homo economicus, 'the Econ' within, would have done.[8]

Without entering into a debate about the appropriate use of economic 'rationality' conceptions, we merely wish to point out the potential pitfalls of taking homo economicus as the normative benchmark that real-world individuals should follow. In the parlance of dual-process theory, which is popular among behavioral economists, this is tantamount to privileging an agent's purportedly long-term or System 2 self with her specific preferences, as opposed to the short-term preferences of the spontaneous and affect-driven System 1 self.[9] Unless one is given adequate normative reasons for this specific starting point, it is not clear why one should accept this approach, given that individuals' allegedly long-term preferences ("I wish I didn't want to smoke!") are often expressive in nature (Schnellenbach, 2012) in addition to the controversial nature of dual-process theorizing in general (Buturovic and

externalities. On the latter point, see, e.g., Guala and Mittone (2015).

[6] In his more recent work, Sunstein (2017, p. 10) discusses educative, non-exploitative nudges that address deliberate System 2 thinking. However, he defends exploitative nudges to a large extent and concedes that sometimes "System 1 nudges have far higher net benefits than System 2 nudges, which can be a waste of time and effort."

[7] A critical overview of ethical objections is given by Rizzo and Whitman (2009), Hausman and Welch (2010), Schnellenbach (2012), and Whitman and Rizzo (2015).

[8] On this point, see, e.g., Thaler (2015, pp. 25, 29, 57). See also the critique advanced by Gigerenzer (2015), drawing inspiration from the 'old' behavioral economics school of Herbert Simon *et al.*

[9] Kahneman (2011) popularized this terminology of System 1/2 thinking. He assumes that real-world individuals are, at any given time, influenced by both a fast, intuitive, spontaneous System 1 and a slow, reflective, rational System 2. Typically, so the argument goes, both systems are in conflict with each other, but eventually the former prevails, which predictably leads to choices the individual regrets ex post.

Tasic, 2015; Rizzo, 2016).

Moreover, libertarian paternalism rests upon a "dualistic model of the human being, in which an inner rational agent is trapped inside a psychological shell" (Infante *et al.*, 2016, p. 1). This 'dualistic approach' has at least three problematic implications. First, by naming homo economicus as a normative role model, this method transforms the entire substance of behavioral economics along with its psychological content into a 'problem' to be overcome rather than something perfectly normal to be constructively used in positive (causal) explanations. In doing so, the approach effectively denies the possibility that deviations from rational choice might be something to be acknowledged as reasonable (Rizzo, 2017; Berg and Gigerenzer, 2007). Second, the dualistic model lacks any psychological explanation of the rational behavior that it cherishes. It presumes that the inner rational agent's 'latent' preferences are perfectly consistent and independent of context. Ultimately, the psychological mechanism meant to give rise to those rational preferences is left as a black box, which introduces a methodological asymmetry (Infante *et al.*, 2016). Third, the dualistic approach fails to establish a coherent conception of *personal identity* in the way that it models human agents. Agents are split into two 'selves' that are represented by different and typically conflicting utility functions. The utility function that maps the System 2 self is then taken to represent the agent's 'true' preferences or her 'true' identity. In the remainder of this paper, we will focus on this third implication, the *identity problem*, and develop a way to overcome the notion of a 'dualistic self'.

4 A Contractarian Approach: The Opportunity Criterion

Through a series of papers, Sugden (2004; 2006; 2008) has developed a criterion that aims at overcoming the dualistic model described above. Sugden elaborates upon a normative criterion explicitly couched in a contractarian framework meant to show that behavioral economics (the positive findings of which he, by and large, accepts) is not only incompatible with paternalistic policy implications in the sense of libertarian paternalism, but calls for an anti-paternalistic stance on public policy design.[10] According to Sugden, normative economists should take a contractarian approach and directly address the citizens affected by their advice rather than viewing themselves as policy advisors that assess the value of social states from a purportedly neutral perspective. Therefore, the task would be to help negotiating "a fair agreement between individuals, each of whom is looking at the world from his own viewpoint" (Sugden, 2008, pp. 229f). Again, we do not want to present his

[10] A behavioral economic case against government paternalism can of course also be made on non-normative grounds, e.g., from a public choice perspective, see Schnellenbach and Schubert (2015) and Schubert (2017).

approach in detail, but will focus on what we see as its key contribution to the fundamental debate shaping BNE: the issue of identity.

Instead of specific preference relations (e.g., of transitivity and completeness), Sugden suggests that 'opportunities' should be the ultimate carriers of value in normative economics. In this context, a given agent's opportunities are defined as "something that he has the power to bring about, if he so chooses" (Sugden, 2010, p. 49). The normative role model is not homo economicus, but rather the 'responsible agent' who treats her past, present, and future preferences as fully her own, even when facing uncertainty or regret (Sugden, 2004, p. 1018). Sugden's 'opportunity criterion' states that from a contractarian viewpoint, behind a hypothetical veil of uncertainty, individuals exhibit an overarching interest in maximizing the chance to satisfy their potential preferences – understood as 'passions' that ultimately lack the need to be justified to any external authority. Sugden's underlying normative intuition is that every agent should be free to try to satisfy *whatever* incoherent and possibly eccentric preferences she finds herself with at any given point in time, provided that she does not harm other people in the process. While individual preferences may very well be 'unconsidered', individuals' "valuing the *opportunity to satisfy them* is considered" (Sugden, 2006, p. 217, italics added).

Sugden claims that his approach overcomes the multiple-selves model of the individual in that he conceptualizes agents as continuing 'loci of responsibility' who endorse any preference their own selves once had or will have. In light of this role model, it is always unconditionally good for each acting agent to maximize her own opportunity set. By pointing to opportunities instead of laundered or nudged preference sets, Sugden admits the possibility of incoherent preferences while retaining the normative substance of liberal welfare economics (i.e., the principle of consumer sovereignty). His 'opportunity criterion' says that institutional arrangements such as markets are good to the extent that they help individuals maximize their opportunities.

4.1 Problems of Sugden's Proposal

At first glance, this approach appears more attractive than its main counterpart, behavioral welfare economics – with its prominent offspring, libertarian paternalism – as it seems to avoid the futile search for people's 'true' preferences (Fumagalli, 2013). However, upon closer inspection, it becomes apparent that although Sugden avoids splitting of the individual into multiple selves, he nonetheless fails to provide a convincing answer to the *identity problem* illustrated in the preceding section. Instead of privileging long-term preferences that a homo economicus would indulge, Sugden advocates the opposite: agents should be free to pursue any preference *at the moment of choice* independent of any concern as to the consequences in terms of well-being. In doing so, Sugden implicitly sides with the dualistic approach. Sugden (2004, p. 243) further

asserts that only by giving authority to the 'impulsive self', rather than the self as "maker of plans or source of reflective judgement about the well-being of the continuing person," might one be able to appreciate the value of the market in providing for opportunity. This view assumes that the absence of interferences (i.e., negative liberty) is a sufficient empirical condition for the quality of agency necessary for the normative role model of the responsible agent. This, however, runs contrary to basic insights from empirical research, which point to the fact that unconstrained consumption tends to produce negative welfare effects for myopic consumers (Mueller *et al.*, 2010). In addition, it also underestimates the necessary cognitive and material resources required for agency capabilities (Sen, 2009). A person who either lacks the necessary 'inner prerequisites' (e.g., in the form of imagination and valuation) or the 'external means' (in the form of time and information) to make reasoned choices is likely to end up with a substantial set of welfare-decreasing decisions since she is not able to correctly reflect upon the consequences of various choice options (Dold, 2018).

In our view, the critical point concerns a key implication that Sugden himself derives from his normative role model of the 'responsible agent'. While such an agent may wish to steer the future development of her own preferences by summoning willpower (which Sugden refers to as 'self-command'), Sugden claims that she *never* wishes to constrain her opportunity set by external means ('self-constraint').[11] If this truly follows from the responsible agent conception, then the model of identity underlying Sugden's approach remains questionable. Many real-world agents are not keen to maximize their own personal opportunity sets independent of any concern for its consequences in terms of personal well-being. Rather, they wish to engage in both self-command *and* self-constraint, depending upon the circumstances (Schubert, 2015). Sugden seems to confuse legitimate self-commitment – comprising both self-command and self-constraint – with illegitimate paternalism (Vanberg, 2014, pp. 338–341).

In summary, Sugden's claim that most real-world individuals would embrace *any* preference, however inconsistent or eccentric, is assumed rather than supported by empirical evidence. It appears implausible to argue that most individuals adopt the particular relationship toward their own preferences that is implicit in Sugden's account. One may argue with Frankfurt (1971) that the capacity to critically reflect upon one's own tastes, inclinations, and dispositions (i.e., factors that ultimately constitute one's preferences) sets human beings apart from other animals. Neglecting that capacity is tantamount to dissolving the individual agent into a faceless being that stumbles through life, driven around by whichever tastes she happens to acquire at that moment of choice. There is no conceptual basis for any notion of agents' idiosyncratic identity

[11] An example of the latter would be an alcoholic's strategy to lock away the spirits in a cupboard and throw away the only key. On the issue of self-command, see Sugden (2004, p. 1018). See Schubert (2015) for a more detailed critique.

in Sugden's account. He merely replaces the arbitrary privileging of people's 'System 2' preferences – predominant in behavioral welfare economics – with the equally arbitrary privileging of the preferences of the acting self. The static character of both approaches does not do justice to the evolving nature of human preferences, or more generally, identity formation.

4.2 A Constructive Critique

We suggest a generalization of this point by arguing that dualistic approaches lack a systematic place for a nuanced critical distance toward one's preferences. One might object that allowing agents to assume a critical position towards their own preferences would presuppose some specific conception of the good life. This, in turn, would narrow the relevance of such an alternative approach to those agents who happen to endorse exactly that particular conception. For instance, allowing preferences to be criticized according to how well they promote one's own happiness makes the whole approach irrelevant to all non-hedonists.

In our view, this problem can be avoided by a dynamic, agent-relative perspective that focuses on the underlying process *giving rise to* the very preferences that are the evaluandum in behavioral welfare economics. Sugden's 'opportunity criterion' does not evaluate preferences per se, but focuses on the maximization of choice sets while neglecting the process that brings the preferences for these choices about. In contrast, focusing on the process of preference formation frees the economist from the need to critically assess single preferences according to whether they satisfy certain criteria (e.g., Do they satisfy rationality axioms? Are they 'authentic'?) or unconditionally endorse any preference an agent happens to hold (as is the case in Sugden's approach). Instead, it asks whether the agent concerned is willing to accept and potentially embrace the process itself.

What is then missing in the dualistic models more generally and in Sugden's approach, in particular? Regarding the former, their static character is readily apparent: at least two sets of conflicting preference orderings are exogenously given and the critical aspect consists of rejecting one of them as 'erroneous' or 'inauthentic' without much justification. In terms of the latter approach, the lack of a proper systematic place for a dynamic account of preference formation becomes apparent when we look at the few instances where Sugden explicitly discusses preference change. On the one hand, he recommends his criterion as well-suited to situations where individuals only form their preferences in the act of choice itself. Here, the 'opportunity criterion' addresses not only actual, but predominantly potential preferences (Sugden, 2008, p. 230). On the other hand, conceptualizing preferences as 'passions' implies that nothing of substance can be said about their origin or about the learning processes behind them. The only instances where Sugden talks about preference change

in the series of papers under discussion is where he models an agent's risk preferences as being mood-dependent. In this way, they switch from one state (risk-friendly) to another (risk-averse) in a non-cumulative way, which merely reflects the instability of individuals' preferences (Sugden, 2006). Importantly, there is no cumulative, irreversible learning of preferences involved.[12] Hence, there is no process that could assume the role of the *evaluandum*.

5 Back to the Unified Self: Preference Learning and Active Choice

The key problem facing both the dualistic approach of behavioral welfare economics and Sugden's alternative has to do with the fact that their implicit conceptions of identity are incomplete. In economics, a person's 'identity' is typically equated with her utility function (i.e., her preference ordering), that is, in turn, exogenously given and stable. Behavioral welfare economics, by separating the agent into (at least) two separate 'selves' with conflicting utility functions, gives up the notion of a unified identity. The one property that clearly sets behavioral welfare economics apart from its neoclassical cousin is the fact that the former introduces a 'faulty' self that, due to its psychological baggage, precludes its rational counterpart from acting upon her 'true' preferences.

In his rejection of behavioral welfare economics, Sugden ultimately retains the focus upon preferences and assigns value to the *opportunity* to satisfy potential preferences, rather than the satisfaction of actual preferences. In his case, identifiable identity is located, not in a particular utility function, but in a specific *relationship* toward one's own potential future preferences. However, that relationship is simply postulated and rather extreme, as it involves total and unconditional endorsement of the preferences of the (future) acting self. His approach conspicuously lacks a theory that relates the degree of endorsement of one's own potential preferences to the nature of the corresponding processes of preference formation. In this section, we will outline the contours of a normative approach that accounts for the possibility that the process of preference formation not only determines the degree of endorsement of potential future preferences, but is itself the key *evaluandum*.

The dualistic approach splits human identity into two selves: one with a long-term view that is slowly deliberating and consistent versus the other, short-term, impulsive and spontaneous. An alternative starting point would be to avoid any normative pre-commitment in favor of either the former (as in behavioral welfare economics) or the latter (as in Sugden's approach). Rather,

[12] In our context, 'cumulative' preference formation means to exclude phenomena such as random preference reversal and mood-dependent switches of preferences; it therefore is a minimal criterion for generic 'preference learning'. On this point, see Witt (2010) and Schubert (2015).

one may choose to endorse the hypothesis that striking a reflective balance between both 'selves' that is neither overly strict nor overly permissive is what makes a human life a successful one (Cowen, 1991). We follow Davis (2016) in postulating that individuals establish and continually re-establish their *identity* through the reflection upon the conflicting present and future demands. According to Davis (2016), a plausible conception of identity requires a notion of *individuation*, the individuals' capacity to critically reflect upon their evolving preferences, that avoids the circularity of the neoclassical model of man, whereby identity is equated with a utility function that is assumed to be given.

Considering behavioral economic insights about the nature of human preferences, it seems that most of them tend to be highly contingent, sometimes even critically dependent on changes in contextual factors that can only be deemed arbitrary (e.g. Bowles, 1998; Lichtenstein and Slovic, 2006). Moreover, processes of preference formation are often not subject to control by the agent herself. Nevertheless, real-world human beings value the fact that they are able to form their own preferences in a way that 'feels' self-determined (Deci and Ryan, 2000).

The highly contingent nature of individual preferences leads us to suggest that economists should take a step back and focus on the quality of the process of preference formation (or 'construction'), rather than on single preferences at any particular moment in time. In doing so, the role of preferences for welfare analysis is qualified: they are only seen as provisional inputs for decision-making processes and are not the crucial unit for normative assessments.[13] What seems to be valuable in a 'behavioral world' is making the experience, on an ongoing basis, of trying out new preferences, discarding some of them, and keeping others. This may be referred to as a process of preference learning (Schubert, 2015). It resonates with the idea, advanced by James Buchanan in his 1979 article *Natural and Artifactual Man*, that human beings face the existential challenge to creatively construct their own identity over the course of their lifetime. By interacting with their environment, agents reconstitute themselves again and again. As Buchanan puts it, "the 'individual', as described by a snapshot at any given moment, is an artifactual product of choices that have been made in prior periods, both by himself or herself and others" (Buchanan, 1979, p. 287).

[13]Long before the current debate on the significance of behavioral evidence for economic welfare analysis, Rothenberg (1962, p. 282) argued for the qualification of the role of preferences (or has he calls it: 'tastes') in the context of a discussion of the notion of consumer sovereignty in the American Economic Review: "Consumers' choices may not reflect their true tastes; ... maybe these tastes cannot accurately be known; or ... are not really 'owned' but only 'loaned' tastes anyway, passed on from one person to another. What really can belong to the self and be accurately known is the experience of making and taking responsibility for choices, whether right or wrong, and seeking to know by this continuing dialogue across the permeable boundary of the self what if anything is worth preserving. It is possible that this quest, given any reasonable degree of responsiveness in the outside world, is what consumers want more than being given what they are told they really want."

Accepting this observation, every individual partakes in idiosyncratic processes of 'self-constitution' throughout their lives – processes whose success critically depends upon reasoning and active choice (Korsgaard, 2009). Individuals are, then, modeled as causally effective agents that, through their own idiosyncratic learning and conditioning history, acquire what is commonly called a 'character'. In this way, individual agents are conceptualized as 'loci of learning'. Importantly, no 'utility functions' are assumed to be given at the outset. In orthodox economics parlance, utility functions are only constructed – a process that involves trial-and-error-based adjustments – in the course of an agent's lifetime.

Such a process-oriented view of welfare and the individual avoids the notion that normative economics needs to take a view on people's 'true interests'. Rather, it should be directed "at the conditions under which people acquire the sense of interest on which they act" (Hargreaves Heap, 2013, p. 995). In interpreting 'sense of interest' as 'preferences', we support a negative specification of Hargreaves Heap's proposal. Here, the normative desideratum is to make sure that the process of preference learning continues in a way that *is not unacceptable* to the individual agents concerned. As such, public policy – and society in general – must make sure that formal institutions *do not hinder* individuals from reconstituting themselves as 'loci of learning' on a continuous basis.

How might formal institutions interfere in this ongoing learning process? In our view, the recently popular policy tool of nudging provides notable examples. Government-issued nudges may impair agents' 'individuation' if they systematically discourage active choice – understood as deliberate and effortful decision-making and the exercise of self-control. The design of defaults is considered one of the most prominent and most effective types of nudging. For example, if defaults are set to maximize retirement saving, then people no longer need to weigh costs and benefits of their relevant choices. Thus, they no longer need to take care of their retirement savings, which remains a fundamental financial question faced by many individuals in their lifetimes. Conversely, one could offer a counterargument that favors this kind of nudge in that it frees scarce mental resources that can be used for other important decisions instead. However, individuals who constantly rely on some external authority to set defaults might lose the capacity for their own active choice, i.e., a continuous lack of exercise of self-control leads to a decrease of that capacity over time (Baumeister *et al.*, 2006, pp. 1779–1786). Interestingly, the decline in the capacity of self-control in one area, e.g. financial affairs, can decrease the exercise of self-control in another, seemingly unrelated area of choice, e.g., diet or alcohol consumption (Oaten and Cheng, 2007). Given that the capacity of self-control seems to be a key prerequisite for successful processes of identity formation, crowding-out effects such as this one must be carefully studied.

In another example, consider the use of psychological tools to influence individuals' behavior (e.g., in the form of corporate commercials or product placements in supermarkets). At the most basic level, the deliberate design of the 'choice architecture' aims to shape individuals' beliefs and steer their choices in a particular manner. Individuals who are partly aware of the fact that they are subject to psychological influences may respond by a display of revolting behavior. Over time, however, there willpower might be depleted and they may succumb to the power of these nudges whose behavioral effects may actually be hedonically pleasurable.[14] In the end, individuals might start questioning the degree to which their choices reflect what one might plausibly call their 'own preferences' or whether they are the product of a myriad of subtle influences. As Waldron (2014) puts it, "what becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interests?" Self-respect, we might add, should be viewed as a crucial product of a 'successful' process of self-constitution; in other words, lack of self-respect may be taken to be one indicator of a deficient process of preference learning.

Thus, not only the capacity for active choice might be jeopardized by nudges that exploit decision biases (System 1 nudging), but also individuals' appreciation of their own decision-making capacity may be negatively affected when their choice situations are carefully engineered with the aim of bringing about certain choice patterns. In the case of System 1 nudging, people's trust in their own preferences is at stake. This is due to the fact that the underlying process of preference formation seems to be, to a problematic degree, under the control of some external authority.

How does the proposal outlined here differ from both behavioral welfare economics and Sugden's 'opportunity criterion'? It differs from the former in its inherently dynamic and procedural nature. In addition, it differs from the latter in allowing for the process of preference learning that has no place in Sugden's account, namely, preference change that is both non-temporary and cumulative.

The three approaches (behavioral welfare economics, Sugden's 'opportunity criterion', and our own preliminary proposal) differ markedly with respect to the value that they assign to single preferences. While behavioral welfare economics makes that value a function of the preferences' formal properties, Sugden shifts the focus away from the satisfaction of actual preferences towards

[14] An example of revolting behavior is the reaction of people to a policy change of the default rule for organ donation from opt-in to opt-out. When the Dutch government passed a bill that shifted the default designation to one of presumed consent, the number of Dutch citizens who registered as nondonors rose to roughly 40 times the number observed in months before the bill was passed. Yet, this spike in active rejections was only temporary (Krijnen *et al.*, 2017).

the satisfaction of potential ones. However, this makes him assign a maximum value to satisfying any preference that an individual may hold now or in the future. Thus, Sugden takes a first step towards focusing on processes rather than outcomes by shifting our attention to the value inherent in the ability to satisfy yet unformed preferences. In contrast, our proposal favors a truly procedural outlook. However, our conceptualization of individuals as *loci of learning* implies an important qualification of the value assigned to the satisfaction of particular preferences as outcomes. We take preferences and choices to be valuable and important in that they serve as constructive tools within an agent's ongoing process of individuation. They do so to the extent that the agent has good reason to accept them as being the outcome of a process of preference learning that *she can reasonably accept as 'her own'*. In line with Hargreaves Heap (2013, p. 985), our proposal acknowledges that people refer to reasons for choice that can make their actions intelligible even if they do not always hold well-defined preferences. In addition, our proposal suggests that social and institutional prerequisites for reasoning about one's preferences – the process of identity formation – constitute the key *evaluandum* that economists should study when they conduct welfare analyses.

6 Conclusion

In light of an ongoing public interest in behavioral policies, one must systematically consider the normative implications of behavioral economics. In this paper, we first discussed the shortcomings of the pragmatists' approach to behavioral policy advice, which focuses on (allegedly) value-free, means-ends calculi. Yet, we argued that any approach in economics that informs public policy must rely on some conception of 'welfare' in order to be able to critically compare alternative instruments. Accepting this premise, we outlined the now dominant approach in BNE, 'behavioral welfare economics,' with its policy offspring 'libertarian paternalism.' We then scrutinized the most prominent alternative account: Robert Sugden's 'opportunity criterion'. We argued that both approaches lack a convincing account of personal identity. If one accepts that incoherent preferences challenge standard notions of welfare in economics and that the quest for 'true' preferences turns out to be as problematic as the claim that real-world individuals happily embrace any preference they might potentially encounter, we argued that it makes sense to shift the conceptual focus on the underlying processes of preference formation and individuation (or 'self-constitution') they imply. We proposed the notion of identity as a useful heuristic for constructing a behaviorally informed normative theory in economics.

In our view, when it comes to normative and policy implications, the key lesson of behavioral economics is the surprising evidence that people's preferences are a function of subtle features of their social and situational context. Real-world individuals may have their preferences 'manipulated' by

both private and governmental agents in a way that economists were unable to discern in traditional models populated by homines economici. Generally, this new body of research gives rise to the question of which channels of influence are acceptable to real-world agents.

We argued with Deci and Ryan (2000) that real-world individuals value the fact that they are able to form their own preferences in a self-determined way. Individuals wish to maintain a certain degree of control over their own processes of preference formation, namely, they wish to accept the processes and their temporary outcomes as phenomena *that they can embrace as their own*. Importantly, being the author of one's own life also requires that one delegates complex choices to external authorities for the simple reason that one's mental resources are scarce. However, it is crucial to acknowledge the trade-off involved: excessive delegation risks jeopardizing one's project of individuation in the form of a reduced ability to make active choices.

Given this refocusing, we advocate for behavioral policy-makers to reconsider their fixation on correcting people's 'decision mistakes'. Instead, behavioral policy should focus on the task of modifying those contextual elements that foster individuals' means of preference learning, viz., *educational and cultural institutions* that provoke individuals' reflection of their evolving preferences and *access to markets* that allow individuals to test different preferences on a continuous basis. These policy implications differ from behavioral welfare economics in that they reject policies that systematically discourage active choice (e.g., in the form of 'sticky' default rules). They also differ from Sugden's 'opportunity criterion' in that they do not advocate mere choice, but also stress the internal and external prerequisites for preference formation.

From our viewpoint, institutions harm individuals when they influence the preference formation process in a way that results in the fragmentation of a person's identity and the feeling of being manipulated. 'Manipulation' is a difficult concept in that it seems to presuppose some kind of superior insight into a distinction between those preferences that are formed in a neutral context (and thus reflect what the individual 'truly' wants) and those preferences that are context-dependent and only 'artificially' induced. Following Hayek (1961), we believe that it is impossible to draw such a sharp distinction. Yet, it seems fair to say that there are contexts that influence choice heavily and others where the manipulative aspect is minimal or reduced. We do not need a reference point of absolute neutrality to relatively assess the degree of manipulation of different preference formation processes.[15] We postulate that preference formation 'works' to the extent that individuals are able to make reasoned choices and are willing to embrace the process itself. In order to secure this result, behavioral policies should meet a basic condition: individuals

[15]In his magnum opus *The Idea of Justice*, Sen (2009) makes this argument regarding the notion of justice: even though we do not know what the ideally just society would look like, we can still rank two states of the world with regard to the relative injustices they comprise.

must not be systematically deprived or discouraged from active choice and learning processes. We deem the theoretical conceptualization of preference learning and the empirical analysis of the way in which real-world individuals understand and develop their preferences to be vital parts of future research in the field of behavioral normative economics.

References

Baumeister, R. F., M. Gailliot, C. N. DeWall, and M. Oaten. 2006. "Self-regulation and personality: How interventions increase regulatory success, and how depletion moderates the effects of traits on behavior". *Journal of Personality*. 74(6): 1773–1802.

Becker, G. S. 1962. "Irrational behavior and economic theory". *Journal of Political Economy*. 70(1): 1–13.

Berg, N. and G. Gigerenzer. 2007. "Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk-taking". *Social Choice and Welfare*. 28(2): 337–359.

Bernheim, B. D. and A. Rangel. 2007. "Toward choice-theoretic foundations for behavioral welfare economics". *American Economic Review*. 97(2): 464–470.

Bernheim, B. D. and A. Rangel. 2008. "Choice-Theoretic Foundations for Behavioral Welfare Economics". In: *The Foundations of Positive and Normative Economics*. Ed. by A. Caplin and A. Schotter. Oxford: Oxford University Press. 7–77.

Bernheim, B. D. and A. Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics". *Quarterly Journal of Economics*. 124(1): 51–104.

Bhargava, S. and G. Loewenstein. 2015. "Behavioral economics and public policy 102: Beyond nudging". *American Economic Review*. 105(5): 396–401.

Bleichrodt, H., J. L. Pinto, and P. P. Wakker. 2001. "Making descriptive use of prospect theory to improve the prescriptive use of expected utility". *Management Science*. 47(11): 1498–1514.

Bolton, G. E. and A. Ockenfels. 2012. "Behavioral economic engineering". *Journal of Economic Psychology*. 33(3): 665–676.

Bowles, S. 1998. "Endogenous preferences: The cultural consequences of markets and other economic institutions". *Journal of Economic Literature*. 36(1): 75–111.

Buchanan, J. M. 1979. "Natural and artifactual man". In: *The Collected Works of James M. Buchanan Volume 1: The Logical Foundations of Constitutional Liberty*. Indianapolis: Liberty Fund, Inc. 246–259.

Buturovic, Z. and S. Tasic. 2015. "Kahneman's failed revolution against economic orthodoxy". *Critical Review*. 27(2): 127–145.

Chetty, R. 2015. "Behavioral economics and public policy: A pragmatic perspective". *American Economic Review*. 105(5): 1–33.
Cowen, T. 1991. "Self-constraint versus self-liberation". *Ethics*. 101(2): 360–373.
Davis, J. B. 2016. "Economics, Neuroeconomics, and the Problem of Identity". *Schmollers Jahrbuch*. 136(1): 15–31.
Deci, E. L. and R. M. Ryan. 2000. "The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior". *Psychological Inquiry*. 11(4): 227–268.
Dold, M. F. 2018. "Back to Buchanan? Explorations of welfare and subjectivism in behavioral economics". *Journal of Economic Methodology*. 25(2): 160–178.
Fumagalli, R. 2013. "The futile search for true utility". *Economics & Philosophy*. 29(3): 325–347.
Gigerenzer, G. 2015. "On the supposed evidence for libertarian paternalism". *Review of Philosophy and Psychology*. 6(3): 361–383.
Gneezy, U. and A. Rustichini. 2000. "A fine is a price". *The Journal of Legal Studies*. 29(1): 1–17.
Grüne-Yanoff, T. 2012. "Old wine in new casks: libertarian paternalism still violates liberal principles". *Social Choice and Welfare*. 38(4): 635–645.
Guala, F. and L. Mittone. 2015. "A political justification of nudging". *Review of Philosophy and Psychology*. 6(3): 385–395.
Gul, F. and W. Pesendorfer. 2007. "Welfare without happiness". *American Economic Review*. 97(2): 471–476.
Hargreaves Heap, S. P. 2013. "What is the meaning of behavioural economics?" *Cambridge Journal of Economics*. 37(5): 985–1000.
Hausman, D. M. and M. S. McPherson. 2009. "Preference satisfaction and welfare economics". *Economics & Philosophy*. 25(1): 1–25.
Hausman, D. M. and B. Welch. 2010. "Debate: To nudge or not to nudge". *Journal of Political Philosophy*. 18(1): 123–136.
Hayek, F. A. 1961. "The Non Sequitur of the "Dependence Effect"". *Southern Economic Journal*. 27(4): 346–348.
Infante, G., G. Lecouteux, and R. Sugden. 2016. "Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics". *Journal of Economic Methodology*. 23(1): 1–25.
Keynes, J. N. 1917. *The Scope and Method of Political Economy*. London: Macmillan.
Korsgaard, C. M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.
Kőszegi, B. and M. Rabin. 2008. "Choices, situations, and happiness". *Journal of Public Economics*. 92(8–9): 1821–1832.

Krijnen, J. M., D. Tannenbaum, and C. R. Fox. 2017. "Choice architecture 2.0: Behavioral policy as an implicit social interaction". *Behavioral Science & Policy*. 3(2): 1–18.

Lichtenstein, S. and P. Slovic. 2006. *The Construction of Preference*. Cambridge: Cambridge University Press.

Madrian, B. C. 2014. "Applying insights from behavioral economics to policy design". *Annual Review of Economics*. 6(1): 663–688.

Manzini, P. and M. Mariotti. 2014. "Welfare economics and bounded rationality: the case for model-based approaches". *Journal of Economic Methodology*. 21(4): 343–360.

McQuillin, B. and R. Sugden. 2012. "Reconciling normative and behavioural economics: the problems to be solved". *Social Choice and Welfare*. 38(4): 553–567.

Mueller, A., J. E. Mitchell, R. D. Crosby, O. Gefeller, R. J. Faber, A. Martin, S. Bleich, H. Glaesmer, C. Exner, and M. de Zwaan. 2010. "Estimated prevalence of compulsive buying in Germany and its association with sociodemographic characteristics and depressive symptoms". *Psychiatry Research*. 180(2–3): 137–142.

Oaten, M. and K. Cheng. 2007. "Improvements in self-control from financial monitoring". *Journal of Economic Psychology*. 28(4): 487–501.

OECD. 2017. *Behavioral Insights and Public Policy: Lessons from Around the World*. OECD Publishing: Paris. URL: http://dx.doi.org/10.1787/9789264270480-en.

Oliver, A. 2015. "Nudging, Shoving, And Budging: Behavioural Economic-Informed Policy". *Public Administration*. 93(3): 700–714.

Rizzo, M. J. 2016. "Behavioral economics and deficient willpower: Searching for akrasia". *Georgetown Journal of Law & Public Policy*. 14(1): 789–806.

Rizzo, M. J. 2017. "Rationality—What? Misconceptions of Neoclassical and Behavioral Economics". *Working Paper*. URL: https://ssrn.com/abstract=2927443.

Rizzo, M. J. and D. G. Whitman. 2009. "The knowledge problem of new paternalism". *BYU Law Review*. 2009(4): 905–968.

Salant, Y. and A. Rubinstein. 2008. "(A, f): choice with frames". *The Review of Economic Studies*. 75(4): 1287–1296.

Schnellenbach, J. 2012. "Nudges and norms: On the political economy of soft paternalism". *European Journal of Political Economy*. 28(2): 266–277.

Schnellenbach, J. and C. Schubert. 2015. "Behavioral political economy: A survey". *European Journal of Political Economy*. 40(1): 395–417.

Schubert, C. 2017. "Exploring the (behavioural) political economy of nudging". *Journal of Institutional Economics*. 13(3): 499–522.

Schubert, C. 2015. "Opportunity and preference learning". *Economics & Philosophy*. 31(2): 275–295.

Sen, A. K. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.

Sugden, R. 2004. "The opportunity criterion: consumer sovereignty without the assumption of coherent preferences". *American Economic Review*. 94(4): 1014–1033.

Sugden, R. 2006. "Taking Unconsidered Preferences Seriously". In: *Preferences and Well-Being*. Ed. by S. Olsaretti. Cambridge: Cambridge University Press. 209–232.

Sugden, R. 2008. "Why incoherent preferences do not justify paternalism". *Constitutional Political Economy*. 19(3): 226–248.

Sugden, R. 2010. "Opportunity as mutual advantage". *Economics & Philosophy*. 26(1): 47–68.

Sunstein, C. R. 2014. *Why Nudge?* New Haven: Yale University Press.

Sunstein, C. R. 2017. *Human Agency and Behavioral Economics: Nudging Fast and Slow*. Cham: Palgrave Macmillan.

Sunstein, C. R. and R. H. Thaler. 2003. "Libertarian paternalism is not an oxymoron". *The University of Chicago Law Review*. 70(4): 1159–1202.

Thaler, R. H. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: Norton.

Thaler, R. H. and C. R. Sunstein. 2003. "Libertarian paternalism". *American Economic Review*. 93(2): 175–179.

Thaler, R. H. and C. R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven: Yale University Press.

Vanberg, V. J. 2014. "Evolving preferences and welfare economics: the perspective of constitutional political economy". *Journal of Economics and Statistics*. 234(2–3): 328–349.

Whitman, D. G. and M. J. Rizzo. 2015. "The problematic welfare standards of behavioral paternalism". *Review of Philosophy and Psychology*. 6(3): 409–425.

Witt, U. 2010. "Economic Behavior: Evolutionary vs. Behavioral Perspectives". Papers on Economics and Evolution, No. 1017, Jena, Max Planck Institute of Economics.