



HAL
open science

The Homer economicus narrative: from cognitive psychology to individual public policies

Guilhem Lecouteux

► **To cite this version:**

Guilhem Lecouteux. The Homer economicus narrative: from cognitive psychology to individual public policies. *Journal of Economic Methodology*, 2023, *The Soul of Economics*, 30 (2), pp.176-187. 10.1080/1350178X.2023.2192222 . hal-03791951v2

HAL Id: hal-03791951

<https://hal.science/hal-03791951v2>

Submitted on 12 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The *Homer economicus* narrative: from cognitive psychology to individual public policies

Accepted version for the special issue ‘The Soul of Economics’, *Journal of Economic Methodology* (2023), 30(2): 176-187.

Publication link: <https://www.tandfonline.com/doi/full/10.1080/1350178X.2023.2192222>

Guilhem Lecouteux

Affiliation: Université Côte d’Azur, CNRS, Gredeg, France.

Postal address: CNRS – GREDEG, Campus Azur; 250 rue Albert Einstein, CS 10269, 06905 Sophia Antipolis Cedex, France.

Telephone number: +33 (0)4 93 95 43 74

Email: guilhem.lecouteux@univ-cotedazur.fr

ORCID: 0000-0001-6602-7247

The *Homer economicus* narrative: from cognitive psychology to individual public policies

Abstract (102 words): A common narrative among some behavioural economists and policy makers is that experimental psychology highlights that individuals are more like Homer Simpson than the Mr Spock imagined by neoclassical economics, and that this justifies policies aiming to ‘correct’ individual behaviours. This narrative is central to nudging policies and suggests that a better understanding of individual cognition will lead to better policy prescriptions. I argue that this *Homer economicus* narrative is methodologically flawed, and that its emphasis on cognition advances a distorted view of public policies consisting in fixing malfunctioning individuals, while ignoring the characteristics of the socio-economic environment that influence individuals’ behaviours.

Keywords: Homer Simpson and Mr Spock; *homo economicus*; rational choice; replication crisis; behaviourally informed policy.

Subject classification codes: A12, B41, C91, D04, D91.

Word count: 7530 (4840 without footnotes and references)

Acknowledgements: I am grateful to seminar audiences in Nice, Nancy, Paris, as well as to the participants to the conference The Soul of Economics (2019) and the 15th Biennial Conference of the International Network for Economic Method (2021) for their comments and suggestions on earlier versions of this work. I also thank the editors and two anonymous referees for their comments that substantially improved this paper.

“One of our major goals in this book is to see how the world might be made easier, or safer, for the Homers among us (and the Homer lurking somewhere in each of us). If people can rely on their Automatic Systems without getting into terrible trouble, their lives should be easier, better, and longer.” (Thaler & Sunstein, 2008a, *Nudge*, p.22)

“Well, excuse me for having enormous flaws I don't work on.” (Homer, *The Simpsons*, Season 18, episode 07)

A popular approach to introduce behavioural economics [BE] is to imagine the characteristics of the fictive ‘*Econs*’ and of the supposedly real ‘*Humans*’. *Econs* are described as ‘highly intelligent beings that are capable of making the most complex of calculations but are totally lacking in emotions’ (Thaler 2015b) and are commonly depicted as Mr Spock from *Star Trek*. *Humans*, on the other hand, would love to be as rational as an *Econ*, though they are probably more like Homer Simpson. They are affected by many emotions and biases, and therefore regularly make errors and choose poorly. This fable of Mr Spock and Homer Simpson is meant to highlight the contrast between the *homo economicus* (*Econ*) populating economic models and the real *homo sapiens* (*Human*). It can be found in numerous publications (e.g. Thaler & Sunstein 2008a, p.42, Ariely 2011, Sunstein 2015, p.12), in introductory courses on BE, and has also reached a much wider audience through its diffusion in the press (e.g. Thaler 2015b). The story we are told is that neoclassical economists have stubbornly studied *Econs*, despite the obvious unrealism of this model, and that BE aims to increase ‘the explanatory power of economics by providing it with more realistic psychological foundations’ (Camerer and Loewenstein, 2004, p.3). An explicit objective of this project is that a ‘reunification [of economics and psychology] should lead to better predictions about economic behavior and better policy prescriptions’ (Camerer, 1999, p.10575).

My aim in this paper is to highlight that this specific interpretation of BE – emphasising the cognitive limitations of ‘real’ individuals compared to their neoclassical *alter ego*, and which lies at the core of nudging policies – offers a distorted view of behavioural public policies by shifting the focus from social institutions and the functioning of markets (which is the primary interest of economists) to consumers’ cognition. The objective of public policies becomes ‘to Simpsonize consumer law’ (Frerichs 2011), making the market a safer place for the Homer lurking within each of us, while largely ignoring the supply side as well as the individual’s socio-economic environment. In what follows, I will use the generic term ‘theorist’

[she] to designate the actual behavioural economist, experimenter, welfare economist, philosopher, or expert, who intends to model the behaviour of an ‘individual’ [he], and who may be tempted to adopt the *Homer economicus* narrative when reporting her findings.

I start by characterising the *Homer economicus* narrative and the articulation between its specific interpretation of BE findings and the policies it justifies (section 1). I will then highlight some of its methodological weaknesses, in particular its mischaracterisation both of *homo economicus* as a hyper-rational Mr Spock (section 2), of *homo sapiens* as Homer Simpson (section 3), and its limited approach to public policy (section 4). I conclude by arguing that an alternative interpretation of the empirical findings of BE, in line with what Hands (2013) labels ‘contemporary revealed preference theory’, offers a radically different perspective by focusing on the processes by which individual preferences are shaped (section 5).

1. The *Homer economicus* narrative

The *homo economicus* is a convenient straw man for BE to justify its investigation into more realistic psychological foundations. Both in academic publications and in more mainstream media, we are told the following story:

If you look at economic textbooks, you will learn that *homo economicus* can think like Albert Einstein, store as much memory as IBM’s Big Blue, and exercise the willpower of Mahatma Gandhi. Really. But the folks we know are not like that. (Thaler & Sunstein 2008a, p.6)

Knowing that most individuals ‘predictably err’ (Thaler & Sunstein 2008a, p.7), it is advanced that ‘the antipaternalist position is [...] a literal nonstarter’ (Sunstein & Thaler 2003, p.1165), which *requires* the theorist to act and design policies aimed at helping the individual. This interpretation of deviations from rational choice in the lab in terms of ‘biases’ and ‘mistakes’ has significantly shaped the field of BE since the 2000s and the publications of Sunstein & Thaler (2003) and Camerer *et al* (2003), respectively, call for a ‘libertarian’ and ‘asymmetric’ paternalism. It also greatly influenced policy discourses around the world with the popularisation of nudges, with which most policy makers are now accustomed.

The *Homer economicus* narrative (HEN) designates the discourse of bounded rationality drawing a direct implication from the accumulation of empirical findings in the lab that human subjects do not generally behave in line with the predictions of expected utility theory, to the justification of new forms of public policies (see Loewenstein and Chater 2017 for a typology).¹ I define a ‘narrative’ as a ‘story [...] the ways in which we construct disparate facts in our own worlds and weave them together cognitively in order to make sense of our reality’ (Patterson & Monroe 1998, 315). My contribution is not primarily about the use of narratives in economics in the construction of models or for argumentative purposes (see McCloskey 1998, Morgan & Wise 2017; and Truc 2021 on the use of narratives in BE), nor about the narrativity of economic agents (see Juille & Jullien 2017), but rather about their political use in the justification of public policies. Discourses are not neutral in the sense of merely describing our world: they can also influence what we do and how we act (Knights and Morgan 1991). From the perspective of social constructivism,² ‘discourse can be analysed to reveal the hegemonic structures and linguistic practices that determine meaning, value, and identity in its field of operation’ (Mehta 2013, 1251). The recurrent use of terms such as ‘myopic’, ‘naïve’, ‘impulsive’, or even ‘idiotic’ to characterize individual behaviours, and the systematic references to expressions such as ‘deviation’ or ‘mistake’, implying that individuals ‘suffer’ from their ‘biases’ clearly tend to pathologise the individual, by comparison to his ‘far-sighted’ and ‘sophisticated’ neoclassical *alter ego* (see Mehta 2013, pp.1253-1255 for a more exhaustive semantic analysis). By concentrating the discourse on the individual’s deficient psychology, his cognitive faculties implicitly become the dominant explanation of his behaviour from the theorist’s perspective, while possibly neglecting other factors. As an illustration, obesity is usually analysed in BE as a self-control problem, revealing the inability for the individual to control his impulses for highly calorific food, rather than the result of poverty (Drewnowski and Specter 2004) or misaligned economic incentives (Galizzi 2012, 19). Designing policies to limit the obesity pandemic could for instance consist in fighting poverty more globally or regulating the price of calories in the junk food industry, while the HEN will point to the responsibility of the consumer, who ought to be rescued from his biases by benevolent experts.

¹ Note that my argument is not targeted to behavioural economists in general, but only to the theorists – whether they be behavioural economists, psychologists, policy advisors – who endorse this type of discourse. In many academic publications, behavioural economists take extra precautions to avoid bold claims about the possible irrationality of individuals, though this is less the case with public intellectuals such as Thaler or Sunstein (who are however much more influential regarding the design of public policies).

² See Burr & Penny (2017) for an introduction. A central figure in this approach is Foucault (1977), who studied how discourses – in penal practice notably – define categories and objects of study, and how those categories shape our behaviours (due to the desire of individuals to conform to what is categorised as ‘normal’ in society).

The HEN is built around three claims:

- (i) Neoclassical economics is based on a descriptively inadequate model of the economic agent, the *homo economicus*, that supposes that individuals reason unflinchingly well.
- (ii) Deviations from expected utility theory [EUT] in the lab are widely documented, they reveal that individuals exhibit various biases, and therefore that they are predictably irrational.
- (iii) A better understanding of individual cognition will allow better policy prescriptions.

The first claim is that the picture of *homo economicus* as a hyper-rational Mr Spock is descriptively wrong. Furthermore, by presenting individual behaviours as ‘deviations’ from the standard (neoclassical) model of decision making, it is implicitly argued that not deviating would be preferable: this is an implicit normative claim of (ii), i.e. that real individuals, if given the choice, *would prefer* to behave like Mr Spock. The second claim emphasises that most people, most of the time, make mistakes. Deviations from the standard model are not epiphenomena, they are constitutive of individual behaviour: from the perspective of the theorist, the norm is that individuals deviate from the standard model (they *predictably* err). As a direct implication of (i) and (ii), the theorist must investigate the individual’s cognition to help him make better choices and improve the design of public policies.

2. *Homo economicus* and Mr Spock

Regarding the identification of the *homo economicus* as Mr Spock, a major worry is that there are probably only very few (if any) economists who genuinely believe that this model is psychologically realistic. Even though it is likely that many economists do not precisely know how to offer an adequate methodological justification of this model (see e.g. Heidl 2016 for a review of possible justifications), none of them will invoke ‘psychological realism’ as an explanation. The practice of identifying the *homo economicus* as a hyper-rational being can be seen as a convenient pedagogical tool to justify (to undergraduate students for instance) the use of mathematical methods of optimisation (e.g. Leamer 2009). This does not mean however that

it can be rejected on the sole basis that it does not provide a realistic description of individual cognition. Furthermore, this element of discourse in the HEN seems to be in contradiction with the practice of behavioural economists, as can be seen with e.g. the formal representation of loss aversion. Given the recurrent celebrations of cumulative prospect theory [CPT] as psychologically more realistic than EUT, we could expect that CPT offers a realistic description of how individuals choose when facing prospects involving gains and losses. However, Wakker and Tversky (1993) proposes an axiomatisation of CPT in which the only difference from EUT is the domain of validity of the principle of tradeoff consistency, since CPT can elicit orderings of value difference from choice data only if all prospects are comonotonic and the outcomes are all nonnegative or all nonpositive (see also Wakker 2010). An individual whose behaviour is consistent with sign-comonotonic tradeoff consistency can exhibit loss aversion, though he is as ‘cold-blooded [a] maximiser’ (Thaler 2015a, p.7) as a *homo economicus*: he indeed behaves as if he was maximising a function characterised by a utility function over outcomes and a weighting probability function, which are possibly different for gains and losses.³

An alternative interpretation of EUT – i.e. that individuals can be represented as choosing the alternatives that maximise their (subjective) expected utility – is that its behavioural axioms provide general characterisations of behaviour (Friedman and Savage 1948, 1952).⁴ This is in line with the approach labelled by Hands (2013) as ‘contemporary revealed preference theory’, offering a renewed approach to revealed preference theory based on the interpretation of preferences and utility functions as summaries of the individual’s choice behaviour, which cannot provide causal explanations of behaviour (as in Bernheim and Rangel 2008, Binmore 2009, Gul and Pustendorfer 2008).⁵ Ross (2005, 2014, 2021) offers the most elaborate version of this approach, which he labels ‘Neo-Samuelson Philosophy of Economics’, by interpreting revealed preference theory as an application of Dennett’s (1987) intentional

³ More generally, the interpretation of deviations from EUT in terms of biases implies that we model individual behaviour as a departure from the benchmark of EUT. The incremental addition of new biases while being committed to solving optimisation problems (as explicitly advocated by Rabin 2013) necessarily increases the computational complexity of the problem to be solved. While this is a reasonable methodology – an alternative to the use of optimisation methods would be to develop models of bounded rationality (Harstad and Selten 2013) – it incidentally implies that Homer might eventually solve an even more complex mathematical problem than Mr Spock!

⁴ Adequately engaging with the different positions regarding the interpretations of the axioms would lead me astray from the main argument of the paper, so I refer the interested reader to Heukelom (2014, chapter 2) and Moscati (2019, part 3).

⁵ The precise characterisation of contemporary revealed preference theory also includes the application of the Generalised Axiom of Revealed Preferences and the possibility to apply the theory to non-human entities, as long as they are responsive to incentives.

stance. From an empirical perspective, what matters is not that actual human beings have the mental capacities to solve complex optimisation problems, but that they behave *as if* they indeed solved such a complex problem. The various formulations of expected utility theory by Von Neumann & Morgenstern (1947), Savage (1954), Anscombe & Aumann (1963), among others, are indeed all ‘as if’ theorems: they take choice data as the primitive of analysis, and state that, if choice data respect certain axioms (such as transitivity and independence), it is as if the individual had chosen by maximising his expected utility – the utility function being defined *a posteriori* to represent his behaviour. This confusion about the interpretation of ‘preferences’ and ‘utility’ in BE – as causing or representing behaviour (see Guala 2019, Thoma 2021b) – is unfortunately very common and probably fuels the idea that economists are committed to an implausible model of behaviour.⁶

Just as an atom does not need to know the law of physics (and solve a complex mathematical problem) to determine how to behave in a given environment, an individual does not have to bother with the mathematical details of the model that the theorist uses to describe his behaviour. From this perspective, EUT is neither *descriptive* (it does not explain how the individual chooses) nor *prescriptive* (it does not suggest how the individual should choose), it is *analytic* – it merely ‘says’ what the individual chooses. It is a formal framework, the *language* that economists use when talking about preferences, choice, and beliefs.⁷ If this interpretation of revealed preference theory is correct,⁸ the experimental test of EUT is not whether the individual indeed maximised a well-defined utility function given his beliefs, but whether it correctly represents the data (and therefore, whether there exists a specification of the utility function and beliefs such that the axioms are respected). If ever experimental data cannot be expressed in terms of EUT, then the theorist needs to enrich her vocabulary (with e.g. probability weighting functions in rank-dependent expected utility or reference-dependence in cumulative prospect theory) to represent more accurately the behaviour of the individual. She should not – as in the HEN – blame the individual for making ‘mistakes’ and behaving in a way that is unintelligible from the perspective of her model.

⁶ Heukelom (2014) suggests that this was already a point of misunderstanding between Savage (1954) and Edwards (1954) and the developing behavioural decision research approach, the latter equating utility with Lewinian valence.

⁷ This is the explicit formulation of Aumann & Brandenburger (1995, 1174-1175) in the context of Bayesian game theory (see Lecouteux 2018 for a similar analysis of EUT in game theory)

⁸ The behaviouristic interpretation of RPT – and its rejection of any interpretation of preferences or beliefs as mental states – is widely disputed in the philosophical literature, such as with Hausman (2000, 2012). See Thoma (2021a,b) for a discussion of the literature and a defence of the behaviouristic interpretation.

3. Interpreting experimental results from the lab

After having portrayed the *homo economicus* as an absurdly rational model of behaviour, the HEN invokes some ‘robust’ findings in psychology to reject its empirical validity. Best-sellers like *Nudge* (Thaler and Sunstein 2008a), *Thinking, Fast and Slow* (Kahneman 2011) or *Predictably Irrational* (Ariely 2008) are full of thought experiments and references to results in cognitive psychology highlighting the many biases that affect our daily lives. Many of those results are surprising and amusing, and this has certainly contributed to the books’ popularity. The promise of combining insights from psychology (‘behavioural insights’) with the elaborated statistical techniques used in economics fuelled the optimism that BE would gain improved scientific credibility and realism, leading *in fine* to better policies. Good statistical practices, however, reveal that many of those findings are *not* robust.

As an illustration, Kahneman dedicates the fourth chapter of *Thinking Fast and Slow* to various forms of priming, and writes:

When I describe priming studies to audiences, the reaction is often disbelief. [...] The idea you should focus on, however, is that disbelief is not an option. The results are not made up, nor are they statistical flukes. You have no choice but to accept that the major conclusions of these studies are true. (pp.56-57)

Schimmack *et al* (2017) however report the impossibility to replicate the great majority of the results mentioned in this chapter. Psychology and social sciences currently face a ‘replication crisis’ (Pashler and Wagenmakers 2012; Gelman & Loken 2014) and systematic replication studies (e.g. the Experimental Economics Replication Project; Replication-Index; Data Replicada) highlight the difficulty of replicating many experimental results (see Ortmann 2021). This suggests that the robustness of the empirical findings on which the HEN is built has been largely overstated, ironically resulting in a possible overconfidence bias among its

proponents,⁹ as well as a confirmation bias that makes them ignore some of the studies that contradict the narrative (see e.g. Gigerenzer 2015 for references).

Consider again loss aversion. We find recurrent claims in the literature that loss aversion is well-documented both in the lab and in the field, with casual evaluations of λ around 2, such as Thaler's (2015a, p.34) claim that 'losses hurt about twice as much as gains make you feel good' – in line with Tversky and Kahneman (1992) initial estimate of $\lambda = 2.25$. A careful reading of the papers making such claims reveals, however, much more nuanced results, because of non-incentivised tasks, restrictions to the gain domain, insufficient theoretical specification, or inadequate statistical reporting¹⁰ (see Harrison and Swarthout 2023 for an extensive review). Probability weighing and reference-dependent preferences (notwithstanding the major issue of the determination of the reference point) have some empirical support, though the existence of a genuine aversion to losses captured by $\lambda > 1$ seems more disputed.

An additional difficulty is that, even though EUT is explicitly rejected as an empirical model, the HEN maintains it offers a prescriptive model about how to behave. That is, if Homer Simpson had the possibility to behave like Mr Spock, he should welcome this opportunity to become more rational, which will increase his welfare (since he can now more efficiently choose the best means to reach his own ends).¹¹ This normative-descriptive distinction in human decision-making, which recognises the possibility of human error, was central in the development of behavioural decision research (see Heukelom 2014) and is found both in experimental psychology and Savage (1954). An important caveat however is that Savage explicitly acknowledges that EUT is a normative theory when choosing in *small worlds*, which is true of the controlled environment of the lab, but not of 'real' decision problems with which individuals are usually confronted, which raises the question of the external validity of a

⁹ Writing on the 'law of small numbers', Tversky and Kahneman (1971) propose that 'psychologists have an exaggerated belief in the likelihood of successfully replicating an obtained finding' (p.105) and may as a result have 'an exaggerated confidence in the validity of conclusions based on small samples' (p.106)

¹⁰ Tversky and Kahneman's (1992) experiment has for instance non-salient monetary rewards, and their results report *median* values for the different parameters (very probably from different subjects). This means that we have no information about the values of the other parameters (curvature of the utility function and probability weighing function) for the median individual whose λ equals 2.25 (and no information to test whether this is statistically different from 1).

¹¹ Anecdotally, in one *Simpsons* episode occasionally used by BE promoters (at least Camerer *et al* 2003 and Thaler and Sunstein 2008, the latter thanking Matthew Rabin for the reference) in which a surgeon is hammering a crayon back into Homer's nose to lower his IQ (he is then saying more and more stupid things and 'the surgeon knows the operation is complete when Homer finally exclaims: "Extended warranty! How can I lose?"', Thaler & Sunstein 2008, p.87), Homer had such an opportunity (his intelligence was significantly higher with the crayon removed) and *willingly chose* to stay stupid, as a means to preserve his social relations. If being Homer is a problem from the theorist's perspective, it may not be the case from Homer's own perspective.

possible bias found in the lab.¹² Furthermore, the normative arguments in favour of choosing in line with EUT are considerably weaker than usually considered (Lecouteux 2021a),¹³ and there is little evidence that individuals who deviate from EUT see a reduction in their welfare (Berg & Gigerenzer 2003, pp.148-149, Arkes *et al* 2016).

The characterisation in the HEN of the individual as a *Human* akin to Homer Simpson – who is predictably irrational (if not stupid) – lacks proper empirical and normative support. First, considerable doubt has been cast over many experimental ‘findings’ in psychology, and deviations from EUT might be less common than usually thought. Second, deviations from EUT in the lab do not necessarily reveal irrationality outside the lab. Those two caveats do not mean that individuals have vast cognitive capacities (cf previous section), that EUT is descriptively valid (there are also successful replications), nor that experiments in economics are useless if we intend to investigate choices in large worlds. The preferences of the individual inferred in the lab (which might be consistent with EUT or not) might for instance still be used as the welfare metric when discussing choices in large worlds (Harrison and Ross 2018). What we should simply avoid is interpreting a deviation from EUT in the lab as the necessary outcome of an underlying cognitive deficiency that will likely affect the choices of the individual outside the lab.

4. Fixing individuals or the market?

I have argued so far that the contrast between fictive *Econs* and supposedly real *Humans* in the HEN is based on two fundamental mischaracterisations, first of EUT (for *Econs*) and second of individual behaviours in the lab (for *Humans*). The third claim in the HEN draws a direct implication from the first two claims to the realm of public policy, by suggesting that an improved knowledge of individual cognition will lead to better policy prescriptions (Chetty 2015). Consider for instance Thaler and Sunstein’s (2008b) analysis of the 2008 financial crisis:

¹² See Guala (2005, 141-160) for a general discussion, and Lecouteux (2021b) on the distinction between small and large worlds in behavioural welfare economics.

¹³ This is partly because EUT is a good theory for choosing in small worlds, while there are many problems for which it does not make sense to look for an ‘optimal’ solution – e.g. how much to save for one’s retirement for an individual who gets his first job in 2022 (who knows what the world and systems of social security could possibly look like in 40 or 50 years?). Proposing to increase individual savings might be a reasonable policy objective for e.g. macroeconomic reasons, though justifying it because this is what people *should* rationally do if they were better informed and less biased mostly reveals a disproportionate overconfidence in the theorist’s ability to predict what the future will be.

Why did Mr Greenspan, along with the rest of the world’s regulators, fail to foresee that this could happen? We think their mistake was to neglect the role of *human nature* (Thaler & Sunstein 2008b, ‘Human frailty caused this crisis’, *Financial Times*)

Their diagnosis is that the crisis resulted from the addition of individual irrationalities, from individual consumers to investors, bankers, and regulators. If they are right, then preventing the next crisis will require improving how all those individuals choose: since a large program of re-education may not be conceivable,¹⁴ a solution to regulate their behaviours could be to slightly alter their environment of choice and to nudge them to adopt better-informed behaviours. This is precisely the type of argument developed by Sunstein (2023) in what he calls ‘Hayekian behavioural economics’ – in which regulators should contribute to reduce the knowledge problem by providing cues about the ‘correct’ behaviour (i.e. what people would choose under epistemically favourable conditions).¹⁵

Focusing on individual cognition entails that we attribute to individual traits the causes of phenomena that could result from more global dynamics (e.g. considering that obesity is the result of a lack of self-control, while poverty might be a more relevant explanation). This could have the very unfortunate effect of advancing ‘behaviourally-informed’ policies which could turn out to be inadequate or ineffective. Consider as an illustration the question of retirement savings in the USA. It is often observed that the situation of many households is precarious (with little to no savings), and BE explains this situation by reference to various biases (Benartzi and Thaler 2007). The *obvious* policy objective should therefore be to increase individuals’ retirement savings, for instance by opting them by default into a retirement savings plan (Thaler and Benartzi 2004). Putting aside the question of whether not saving for one’s retirement genuinely reveals an individual mistake (Lecouteux 2015), we might wonder whether this individual behaviour makes sense from a collective perspective. Living in a society that promotes a high level of consumption has many undesirable effects (ecological for instance), though one significant advantage is that maintaining a high level of aggregate demand is likely to stimulate economic activity, which is *in fine* beneficial to all – at least from a macroeconomic

¹⁴ Note that this position is disputed, as with the ‘boost’ program of Hertwig & Grüne-Yanoff (2017), which aims to empower individuals by enhancing their competences.

¹⁵ Sugden (2023), commenting on Sunstein’s paper, notes that this is exactly the whole idea of nudging and libertarian paternalism – and argues that Hayek would very likely have disagreed with Sunstein’s proposal.

perspective. While increasing individual savings would probably benefit each individual household, *ceteris paribus*, there is a risk that increasing the savings of *all* households might lead to an economic slowdown (and possibly to lower savings *in fine*, according to Keynes's Paradox of thrift). What is sensible for the individual may not be true at the social level: the justification of a nudge cannot solely be based on evidence regarding the behaviour of isolated individuals, but must include a richer picture of socially embedded individuals, whose behaviours are influenced by norms which might be justifiable at the collective level, even though they derive from the aggregation of individual 'irrationalities'.

A second difficulty is that the HEN advances a discourse expressed from the perspective of actors in a market, who cannot question the rules of the market itself. Narratives are contextually thick and provide 'a sense of speakers' cognitive maps of themselves, both in relation to others and in the specific contexts of their described behaviors' (Patterson & Monroe 1998). In the case of the HEN, and more particularly of nudging policies, the theorist sees herself as an expert consultant (for the State or the private sector). Consider Thaler and Sunstein's presentation of the cafeteria problem:

A friend of yours, Carolyn, is the director of food services for a large city school system. She is in charge of hundreds of school, and hundreds of thousands of kids eat in her cafeterias every day [...]

One evening, over a good bottle of wine, she and her friend Adam, a statistically oriented management consultant who has worked with supermarket chains, hatched an interesting idea. Without changing any menus, they would run some experiments in her schools to determine whether the way the food is displayed and arranged might influence the choice kids make. [...]

With hundreds of schools to work with, and a team of graduate student volunteers recruited to collect and analyze the data, Carolyn believes that she now has considerable power to influence what kids eat. Carolyn is pondering what to do with her newfound power. (Thaler & Sunstein, 2008a, 1-2)

The answer to the question of 'what to do with her newfound power' will here depend on her client, i.e. what the person who recruited Carolyn expects from her as an expert in organizing cafeterias. From the perspective of an actor directly involved in the market, there cannot be policy proposals about the actors who are out of reach of her powers. This implies that most

discussions will revolve around the idea of how to ‘fix’ the deficient behaviour of the consumer (e.g. make individuals save more or eat less), while ignoring the behaviour of actors on the supply side. Discussing the cases of ‘nudges that fail’, Sunstein mentions as a possible cause of failure the possibility of ‘counternudges, in the form of compensating behavior on the part of those whose economic interests are at stake’ (Sunstein 2017, 6). The solutions which are put forward for choice architects when nudges are ineffective are:

(1) do nothing, (2) nudge better (or differently), and (3) fortify the effects of the nudge, perhaps through counter-counternudges, or perhaps through incentives, mandates, or bans (Sunstein 2017, 4)

It is striking that, even though it is explicitly recognized that other actors on the market may intervene to limit the efficacy of a nudge, the only solutions advanced are *addressed to the consumer* (the bans and mandates concern indeed the choice of the individual, not a constraint on how the supply side can exploit the bounded rationality of the individual). Public policy is limited to a consulting role with the goal of protecting consumers, without however being able to directly regulate the practices on the supply side. The market within which individuals interact is seen as a confrontation between benevolent choice architects on the demand side, and profit-maximising choice architects on the supply side. The aim of public policy is merely to nudge deviant consumers back to their demand curves, without fundamentally altering the incentives and competition rules in the market.

5. Which alternative to the HEN?

My aim in this paper was to critically review the argumentative strategy of the popular narrative in BE that consists in justifying policies inspired from ‘behavioural insights’ aiming to help cognitively deficient individuals to achieve their own goals (and *in fine* behave just as the standard *homo economicus*). *Humans* are however more like ‘faulty *Econs*’ than *Homo sapiens* (Infante *et al* 2016a, Hands 2020). The HEN is implicitly committed to an atomistic conception of the individual, the model of the inner rational agent (Infante *et al* 2016a): the individual is represented as made up of an inner neoclassically rational agent (our inner Econ), ‘trapped’ within an error-prone psychological shell (our Homerish biases), which distorts how the inner

agent interacts with the real world.¹⁶ By investigating the functioning of this psychological shell, the theorist could then elicit the true preferences of the inner agent – how he would behave if freed from cognitive limitations – as well as determine the policy tools adapted to the individual’s cognition, such that ‘the world might be made easier, or safer, for the Homers among us’ (Thaler and Sunstein 2008a, p.22). The HEN however discards the inherent sociality of the individual, which leads to a narrow view of public policies focused on individual decisions rather than social interactions.

An alternative approach – which would require abandoning BE’s commitment to methodological individualism – is that the behavioural patterns revealed in the lab mostly inform the theorist about the structure of social institutions. Indeed, a methodologically sounder interpretation of EUT – as characterising behaviour – is that preferences elicited in the lab reveal the specificities of the behavioural patterns followed by individuals, while remaining *a priori* agnostic about the mechanisms causing those patterns. The ‘biases’ of the theorist’s model are just convenient labels to designate deviations from EUT, which do not presuppose that such biases genuinely affect an imagined optimisation process in the individual’s mind. Consider for instance the widely claimed result that women are more risk averse than men¹⁷ (e.g. Croson and Gneezy 2009). A first interpretation of this result is that there might be fundamental differences between men and women’s cognition, which could be the outcome of evolutionary pressures (as suggested by Croson and Gneezy 2009). A second interpretation is that stereotypes about gender are internalised by individuals and shape how they behave. Experimental results would not give much information about the individuals’ cognition, but about their social environment. Incidentally, a third interpretation (somewhat related to the discussion in section 3 that theorists may be as biased as the subjects they study) is that the finding that women are more risk averse might also be the consequence of a confirmation bias among economists, who tend to interpret their data in line with socially-held prior beliefs, while the empirical support for this conclusion

¹⁶ As noted by Hausman (2016), the assumption that agents have true preferences for certain alternatives does not commit behavioural welfare economists to the ontological claim that there exists a genuine ‘inner rational agent’ within the individual. The argument of Infante *et al* (2016a,b) is however that behavioural welfare economists believe that there exists some kind of mode of latent reasoning *that would generate neoclassical preferences*. In the absence of arguments supporting that a ‘reasoning free from psychological imperfections’ would necessarily lead to complete, transitive, and context-independent preferences (it is particularly difficult to justify that completeness will be realised), we should question even the *counterfactual* possibility of such an inner rational agent.

¹⁷ This is not technically a ‘bias’ as described in the HEN, though what matters for my argument here is the different ways to interpret the same experimental findings.

is much more nuanced (Nelson 2014; see Nelson 2015 and Filippin and Crosetto 2016 for reviews of the empirical literature).

When formulating normative judgements about ‘biases’ as defined in her model, the theorist therefore forms normative judgements about the social norms that influence how the individuals behave. The fact that retirement savings are low in the USA can for instance be analysed as the direct consequence of living in a society in which individuals consume and save as if they had a present bias. If it is collectively and democratically agreed that this poses a normative problem (based e.g. on different welfare metrics proposed by theorists to fuel the public discussion), then public policies should not merely target individual decision processes (by changing default options), but more globally consider the mechanisms leading to such a preference for immediacy, such as a culture of speed (Tomlinson 2007) or the impact of poverty on impatience (De Bruijn & Antonides 2021). What economists would need to improve policy recommendations would thus be more sociology, rather than more psychology.¹⁸

References

- Ariely, D. (2008). *Predictably irrational*. New York, NY: Harper Audio.
- Ariely, D. (2011) *The upside of irrationality: The unexpected benefits of defying logic at work and at home*. HarperCollins, London.
- Arkes, H. R., Gigerenzer, G., & Hertwig, R. (2016). How bad is incoherence?. *Decision*, 3(1), 20.
- Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *Annals of mathematical statistics*, 34(1), 199-205.
- Aumann, R., & Brandenburger, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica: Journal of the Econometric Society*, 1161-1180.
- Benartzi, S., & Thaler, R. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic perspectives*, 21(3), 81-104.

¹⁸ Ross (2023) argues that this is precisely the evolution of the field, with a progressive convergence between economics and sociology (especially thanks to a common interest in networks). This shift can also be observed in Truc’s (2021) bibliometric analysis of BE, which identifies the recent rise of social preferences and norms as the most studied topic in BE, while observing a decline since the 1980s of publications by behavioural economists in psychology journals.

- Berg, N., and G. Gigerenzer. 2010. As-if behavioral economics: neoclassical economics in disguise? *History of Economic Ideas* 18: 133–165.
- Bernheim, D., & Rangel, A. (2008). Choice-theoretic foundations for behavioral welfare economics. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 155–192). Oxford: Oxford University Press.
- Binmore, K. (2009) *Rational decisions*. Princeton University Press.
- Burr, Vivien and Dick, Penny (2017) Social Constructionism. In: The Palgrave Handbook of Critical Social Psychology. Palgrave Macmillan, London, UK. ISBN 9781137510174
- Camerer, Colin. 1999. Behavioral economics: Reunifying psychology and economics. *Proceedings of the National Academy of Sciences*, 96(19): 10575-10577.
- Camerer, C., Issacharoff, S., Loewenstein, G., O'donoghue, T., & Rabin, M. (2003). Regulation for Conservatives: Behavioral Economics and the Case for " Asymmetric Paternalism". *University of Pennsylvania law review*, 151(3), 1211-1254.
- Camerer, Colin and George Loewenstein. 2004. Behavioral Economics: Past, Present, Future, in Colin Camerer, George Loewenstein and Matthew Rabin (eds), *Advances in Behavioral Economics*, Princeton University Press, pp. 3–51.
- Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review*, 105(5), 1-33.
- Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2), 448-74.
- De Bruijn, E. J., & Antonides, G. (2021). Poverty and economic decision making: a review of scarcity theory. *Theory and Decision*, 1-33.
- Dennett, D. C. (1987). *The intentional stance*. MIT press
- Drewnowski, A., & Specter, S. E. (2004). Poverty and obesity: the role of energy density and energy costs. *The American journal of clinical nutrition*, 79(1), 6-16.
- Edwards, W. (1954). The theory of decision making. *Psychological bulletin*, 51(4), 380.
- Filippin, A., & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, 62(11), 3138-3160.
- Frerichs, S. (2011). False promises? A sociological critique of the behavioural turn in law and economics. *Journal of Consumer Policy*, 34(3), 289-314.
- Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of political Economy*, 56(4), 279-304.

- Friedman, M., & Savage, L. J. (1952). The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy*, 60(6), 463-474.
- Foucault, M. (1977) *Discipline and punish*. London: Allen Lane.
- Galizzi, M. M. (2012). Label, nudge or tax? A review of health policies for risky behaviours. *Journal of public health research*, 1(1), jphr-2012.
- Gelman, A., & Loken, E. (2014). The statistical crisis in science: data-dependent analysis--a "garden of forking paths"--explains why many statistically significant comparisons don't hold up. *American scientist*, 102(6), 460-466.
- Gigerenzer, G. (2015). On the supposed evidence for libertarian paternalism. *Review of philosophy and psychology*, 6(3), 361-383.
- Guala, F. (2005). *The methodology of experimental economics*. Cambridge University Press.
- Guala, F. (2019). Preferences: neither behavioural nor mental. *Economics & Philosophy*, 35(3), 383-401.
- Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. *The foundations of positive and normative economics: A handbook*, 3-42.
- Hands, D. W. (2013). Foundations of contemporary revealed preference theory. *Erkenntnis*, 78(5), 1081-1108.
- Hands, D. W. (2020). Libertarian paternalism: taking Econs seriously. *International Review of Economics*, 67(4), 419-441.
- Harrison, G. W., & Ross, D. (2018). Varieties of paternalism and the heterogeneity of utility structures. *Journal of Economic Methodology*, 25(1), 42-67.
- Harrison, G., & Swarthout, J. T. (2023). Cumulative prospect theory in the laboratory: A reconsideration. In G. Harrison, & D. Ross (Eds.), *Models of risk preferences: Descriptive and normative challenges*. Emerald. Forthcoming
- Harstad, R. M., & Selten, R. (2013). Bounded-rationality models: tasks to become intellectually competitive. *Journal of Economic Literature*, 51(2), 496-511.
- Hausman, D. M. (2000). Revealed preference, belief, and game theory. *Economics & Philosophy*, 16(1), 99-115.
- Hausman, D. M. (2012). *Preference, value, choice, and welfare*. Cambridge University Press.
- Hausman, D. M. (2016). On the econ within. *Journal of Economic Methodology*, 23(1), 26-32.
- Heidl, S. (2016). *Philosophical problems of behavioural economics*. Routledge.
- Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, 12(6), 973-986.

- Heukelom, Floris. 2014. *Behavioral Economics: A History*. New-York: Oxford University Press.
- Infante, G., Lecouteux, G., & Sugden, R. (2016a). 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.
- Infante, G., Lecouteux, G., & Sugden, R. (2016b). ‘On the Econ within’: a reply to Daniel Hausman. *Journal of Economic Methodology*, 23(1), 33-37.
- Juille, T., & Jullien, D. (2017). Narrativity and identity in the representation of the economic agent. *Journal of Economic Methodology*, 24(3), 274-296.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Knights, D., & Morgan, G. (1991). Corporate strategy, organizations, and subjectivity: A critique. *Organization studies*, 12(2), 251-273.
- Leamer, E. E. (2009). *Macroeconomic patterns and stories*. Berlin Heidelberg: Springer-Verlag.
- Lecouteux, G. (2015). In search of lost nudges. *Review of Philosophy and Psychology*, 6(3), 397-408.
- Lecouteux, G. (2018). Bayesian game theorists and non-Bayesian players. *The European Journal of the History of Economic Thought*, 25(6), 1420-1454.
- Lecouteux, G. (2021a). Behavioural welfare economics and consumer sovereignty, In Heilman, C. & Reiss, J. (Eds), *Routledge Handbook of Philosophy of Economics*.
- Lecouteux, G. (2021b). Welfare economics in large worlds: welfare and public policies in an uncertain environment, In Kincaid, H. & Ross, D. (Eds), *Elgar Modern Guide to the Philosophy of Economics*.
- Loewenstein, G., & Chater, N. (2017). Putting nudges in perspective. *Behavioural Public Policy*, 1(1), 26-53.
- McCloskey, D. N. (1998). *The rhetoric of economics*. Univ of Wisconsin Press.
- Mehta, J. (2013). The discourse of bounded rationality in academic and policy arenas: pathologising the errant consumer. *Cambridge Journal of Economics*, 37(6), 1243-1261.
- Morgan, M. S., & Wise, M. N. (2017). Narrative science and narrative knowing. Introduction to special issue on narrative science. *Studies in History and Philosophy of Science Part A*, 62, 1-5.

- Moscatti, I. (2019). *Measuring utility: From the marginal revolution to behavioral economics*. Oxford Studies in History of Economics.
- Nelson, J. A. (2014). The power of stereotyping and confirmation bias to overwhelm accurate assessment: the case of economics, gender, and risk aversion. *Journal of Economic Methodology*, 21(3), 211-231.
- Nelson, J. A. (2015). Are women really more risk-averse than men? A re-analysis of the literature using expanded methods. *Journal of economic surveys*, 29(3), 566-585.
- Ortmann 2021 On the foundations of behavioural and experimental economics. In Kincaid, H. & Ross, D. (Eds). *Elgar Modern Guide to the Philosophy of Economics*.
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on psychological science*, 7(6), 528-530.
- Patterson, M., & Monroe, K. R. (1998). Narrative in political science. *Annual review of political science*, 1(1), 315-331.
- Rabin, M. (2013). Incorporating limited rationality into economics. *Journal of Economic Literature*, 51(2), 528-43.
- Ross, D. (2005). *Economic theory and cognitive science: Microexplanation*. MIT press.
- Ross, D. (2014). *Philosophy of Economics*. Palgrave Macmillan.
- Ross, D. (2021). Neo-Samuelsonian methodology, normative economics, and the quantitative intentional stance. In *Methodology and History of Economics* (pp. 90-116). Routledge.
- Ross, D. (2023). Economics is converging with sociology but not with psychology. *Journal of Economic Methodology*, 1-22.
- Savage, L. J. (1954). *The foundations of statistics*. Courier Corporation.
- Schimmack, U., Heene, M. & Kesavan, K. (2017). Reconstruction of a Train Wreck: How Priming Research Went off the Rails. *Replication Index*, URL: <https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-off-the-rails/>
- Sugden, R. (2023). How Hayekian is Sunstein's behavioral economics?. *Behavioural Public Policy*, 7(1), 189-198.
- Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 1159-1202.
- Sunstein, C. R. (2015). *Choosing not to choose: Understanding the value of choice*. Oxford University Press, USA.

- Sunstein, C. R. (2017). Nudges that fail. *Behavioural public policy*, 1(1), 4-25.
- Sunstein, C. R. (2023). Hayekian behavioral economics. *Behavioural Public Policy*, 7(1), 170-188.
- Thaler, R.H. 2015a. *Misbehaving: The Making of Behavioral Economics*. New York, NY: W.W. Norton & Company.
- Thaler 2015b. Unless you are Spock, irrelevant things matter in economic behavior, *New York Times*, <https://www.nytimes.com/2015/05/10/upshot/unless-you-are-spock-irrelevant-things-matter-in-economic-behavior.html>
- Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of political Economy*, 112(S1), S164-S187.
- Thaler, R. H., & Sunstein, C. R. (2008a). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin
- Thaler, R. H. & Sunstein, C. R. (2008b). Human frailty caused this crisis. *Financial Times*. <https://www.ft.com/content/ac2839ac-b015-11dd-a795-0000779fd18c>
- Thoma, J. (2021a). Folk psychology and the interpretation of decision theory. *Ergo*, 7, 904-936.
- Thoma, J. (2021b). In defence of revealed preference theory. *Economics and Philosophy*, 37(2), 163-187.
- Tomlinson, J. (2007). *The culture of speed: The coming of immediacy*. Sage.
- Truc, A. (2021). Becoming paradigmatic: the strategic uses of narratives in behavioral economics. *The European Journal of the History of Economic Thought*, 1-23.
- Truc, A. (2022). Forty years of behavioral economics. *The European Journal of the History of Economic Thought*, 29(3), 393-437.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2), 105.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4), 297-323.
- Von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behavior* (2nd rev ed.). Princeton, NJ: Princeton University Press.
- Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.
- Wakker, P., & Tversky, A. (1993). An axiomatization of cumulative prospect theory. *Journal of risk and uncertainty*, 7(2), 147-175.