Towards a living theoretical spine for (behavioural) economics

Gigi Foster

Abstract
The past decade has witnessed an explosion in lay applications of empirical insights credited to the field of behavioural economics. Organizations in the public and private sectors have increasingly spawned sub-units with names clearly signalling behavioural economics, e.g., “behavioural insights team”, ostensibly to harness the new knowledge of behavioural economics and apply it to problems of particular import to the organization or its stakeholders. In stark contrast to this impressive take-up of what behavioural economics is seen to offer by those outside the academy, the basic theory of human decision-making promulgated by academic economists for decades has not fundamentally shifted as a result of the efforts of behavioural economics researchers. This paper presents an argument for devoting effort toward the development of true theoretical advance in the core economic model of decision-making, motivated but not confined to the extensions suggested by the empirical results delivered to date by behavioural economists studying choice.

JEL Classification: B20; B41; D01; D91

Keywords
behavioral — decision-making — microeconomic theory

“’It’s said that science will dehumanize people and turn them into numbers. That’s false, tragically false… Science is a very human form of knowledge. We are always at the brink of the known, we always feel forward for what is to be hoped…’”

JACOB BRONOWSKI
The Ascent of Man, BBC 1973
Episode 11 “Knowledge or Certainty”

Introduction
The theory of individual decision-making behaviour taught to first-year economics students, and built upon in all subsequent courses in the discipline, has remained fundamentally unchanged for several decades. The core constituents of the theory are simple to enumerate. There is a single decision-maker who chooses in a deliberate fashion how to allocate his scarce resources (typically money) to alternative purposes (usually goods) based on an optimization problem where he aims to get the highest amount of utility (a catch-all for pleasure, happiness, and enjoyment, assumed to be generated for him to feel in varying degrees when he allocates resources to particular purposes) subject to his budget constraint. Not only his set of alternatives and constraints but the way in which he derives utility from his choices are fully known to him, are more or less stable through time and across decision contexts, and otherwise have a number of specific mathematical properties that catalyse the mathematical manipulation characterising much economic analysis. His budget constraint is set exogenously, at least in the short run, as is his utility function. His income, which may be exogenous in all periods in some more simplistic models, informs his budget constraint; in analyses of labour or public choice, the generation of income is often modelled in its own right, as a function of the individual’s utility-maximizing choices in the realm of leisure and other goods, and possibly his investments into education and training. With its wild success in influencing thought about the human predicament, from how companies understand their customers to how policy-setters think about resource-allocation problems, this is by far the most impactful model of human choice-making ever produced by social science. It should then come as little surprise that it has been around for so long, underpinning the view of society that our discipline promotes in a form that has been essentially unchanged since the 1950s.

Partly fuelled by frustration with this inertia, behavioural economics has gathered identity and strength over the past generation as it seems to offer an alternative to the strict
assumptions of the traditional model. Behavioural economists have pointed to many stylized empirical facts that appear to run counter to the predictions made by the model above, and outside the academy these facts have been gathered up and used as motivation to argue for interventions (e.g., “nudges” (Sunstein and Thaler, 2008)) allegedly designed to correct for the natural yet apparently economically unmodeled tendencies (e.g., “heuristics”) of human decision-making.

Although the core decision-making model has remained unchanged, the same frustration with its essential inertia that has fuelled the rise of behavioural economics as a field has also borne a proliferation within economics—some penned by behavioural economists—of slightly tweaked versions of the model, each of which is claimed to capture some area of behavioural complexity. Researchers concerned with “explaining” altruistic behaviour, for example, have added into the individual utility function the utility of others (Fehr and Schmidt, 2006); those interested in “explaining” conformist behaviour have added into the utility function various terms that supposedly capture the individual’s view of and proclivities in regard to the groups surrounding the individual (Akerlof and Kranton, 2000); those concerned with decisions in the realm of “non-market” production via unpaid labour have assumed that the consumer desires goods like child quality and household cleanliness (Becker 1965); and so on. The model is so flexible, as illustrated by applications like these and a hundred others, that it is easy to be seduced into believing that it represents all that economics need offer in the realm of modelling the individual choices that in aggregate make and constitute whole societies. Like an infinitely-headed drill, the model can seemingly be customized to fit any choice-making scenario one cares to name. Although in the past, and still today in other disciplines, this model has been seen as normative rather than positive (Chai, 2005)—i.e., a description of what should be done if it were possible, rather than what in fact is done—today the model is heavily used by economists for its predictive power, in part because for a wide range of situations, it will deliver a defensible answer to the question of “what would a human do in this situation?” . What more, it may reasonably be asked, must economists provide?

In this short paper I argue that the stagnation of the core economic model of choice as outlined above is a problem for the discipline as a whole and in particular for the field of behavioural economics, and that it is eminently reasonable for those who fund our research and look to us for advice to ask us to try to dislodge our core model of choice from its current position at the centre of our positive analyses. My claim is that any living science is fed by the tolerance and openness embodied in its practitioners, who by definition in their chosen profession seek the truth about some matter of interest. Creating a stylized model of how things (should)

The problem of theoretical stagnation

To see that the model accepted as a descriptor for individual decision-making in economics has not altered for decades, one need only consult first-year undergraduate textbooks hailing from successive decades ago. Samuelson’s canonical economics text, first published in 1948, presents in essence the individual choice-making framework sketched above as the source of micro-level inputs into the production and investment behaviour described by the neoclassical macroeconomic model that forms the text’s primary pedagogical focus (Samuelson, 1948). After the smashing success of Samuelson’s book in the post-war period, micro-economic models of decision-making appearing in subsequent introductory and often intermediate texts have been built from the same constituents. There have naturally been attempts to introduce different material into introductory undergraduate economics, a task occasionally undertaken with sufficient energy to merit a journal article describing the effort (e.g., Zweig, 1972), but these attempts have typically seen as their main enemy either the increasing quantification of economics (as documented in Kamerschen, 1977) or the narrowness of focus upon firm and consumer behaviour within modern capitalist systems, to the exclusion of consideration of the broader institutional backdrop that sustains such systems.

While our model of individual choice has not been attacked wholesale from within the discipline—as such an attack would immediately place the attacker outside the discipline—many lunge have been made at its various sides and angles. Increasingly over the past decades, these attacks have originated in the camp of behavioural economics. Behavioural economists have suggested in turns that individual decision-
making is afflicted by systematic “biases” contra the standard model (for a list of such biases, see Blumenthal-Barby and Krieger, 2015); that it may be informed by other-regarding rather than only self-regarding preferences (e.g., Andreoni and Miller, 2002); that in different settings it occurs in different ways, each of which is best understood as a separate process (Kahneman, 2011); and that in a thousand other ways the standard model is incomplete. Many of these critiques have been defended based on robust scientific evidence, and even those that have been less convincingly empirically validated have typically been the product of thought and conjecture as befits a scientist. In other words, these critiques are not ill-conceived, half-baked notions dreamt up in a fit of pique just to irritate adherents to the standard canon. They represent reasonable guesses about things that might be part of decision-making that are not captured in what we teach our first-year students. Yet, because none is truly a wholesale critique of our standard model is incomplete. Many of these critiques have been defended based on robust scientific evidence, and even those that have been less convincingly empirically validated have typically been the product of thought and conjecture as befits a scientist. In other words, these critiques are not ill-conceived, half-baked notions dreamt up in a fit of pique just to irritate adherents to the standard canon. They represent reasonable guesses about things that might be part of decision-making that are not captured in what we teach our first-year students. Yet, because none is truly a wholesale critique of our standard model is incomplete. Many of these critiques have been defended based on robust scientific evidence, and even those that have been less convincingly empirically validated have typically been the product of thought and conjecture as befits a scientist. In other words, these critiques are not ill-conceived, half-baked notions dreamt up in a fit of pique just to irritate adherents to the standard model.

The challenge

Why has the simple view of individual choice behaviour sketched in the Introduction persisted for decades, essentially without innovation?

The first and most obvious answer is that ours is a pretty good model. While true, that very statement shows the tension between the ideals of economics as practiced today and the traditional ideals of science. Suppose that Albert Einstein –an approximate embodiment of the archetypal scientist– were to have succeeded in reconciling his theory of relativity with the tenets of quantum mechanics, thereby explaining what is known today as “quantum entanglement”; in Einstein’s day this was known as the Einstein-Podolsky-Rosen paradox (Einstein, Podolsky and Rosen, 1935). Would he have then sat back, satisfied with his “pretty good” effort, and devoted himself henceforth to industrial and/or commercial applications of his resolution of quantum entanglement, exploited in slightly different ways to suit the needs of different objectives and conditions? I claim that his nature as a true scientist would have compelled him to continue to attempt to perfect his understanding, such that he could inform applications with a more and more accurate mathematical model, drawing on new insights to reject old blueprints if they were demonstrably based on less-realistic assumptions he had previously held. He might well instruct apprentices in how to apply the existing model, and disseminate knowledge of how to apply it, but he would himself continue to strive for a more perfect holistic understanding of quantum entanglement that would inform subsequent applications. While Einstein was a scientist first and foremost and a technician only in service to his latest understanding, many academic economists today are technicians first and foremost, and scientists only when convenient.

To understand why this is so, we must consider the incentives facing today’s academic economist. Needing to eke out a living for himself in the publish-or-perish world of academia, the young tenure-track economist today will generally choose to embark on one of the following paths. He may target methodological advance; he may aim to apply existing economic models and analytical tools to novel problem settings in order to generate implications for welfare-enhancing interventions, broadly conceived; or he may join the exclusive cadre of technical economic theorists and attempt to deliver theoretical models that incrementally build upon the breathtaking array of existing models of highly stylized situations that have been produced by this cadre over the past several decades. Of these paths, methodological advance is perhaps the most scientific and yet also the most risky; applied economics is the least risky and most frequently-selected path, and consists of working mainly as a technician; and theoretical modelling is both risky and plied most successfully by expert technicians. Economic theorizing about the fundamental processes of the human agent or of how those processes interact with the society in which he lives, in the way that such modelling was known at the time of Adam Smith or even Alfred Marshall, is today pursued only by very few people, most of whom are of only marginal influence within the academy. Such theorizing has been crowded out in both letter and spirit.

Economists today are incentivized to care about publishing in journals that have high citation counts and otherwise-measured “impact”, meaning in large part journals that are popular in the circle of economists. This system of incentives creates both a clique of selfreferencing scholars and a very strong personal incentive for anyone outside that group to become a member of it if he wishes to gain status as an economist. To become a member of the clique of present-day economic theorists, in particular, requires that one have a great deal of expertise in mathematical techniques and a bank of knowledge about how to build mathematical models that is very difficult to acquire except through a heavy investment in apprenticeship. Once such an investment is made, the investor’s continued career depends upon the very system that originally challenged him, and the cycle begins again, with renewed commitment by the existing clique of theorists to the standard model of choice. It is in part from this socially-mediated situation – ironically or not, one that itself is rarely modelled in economics – that the inertia of our model of individual decision-making arises.

The consequences

In face of a stagnated model of choice and the many incremental and ultimately failed attempts from within the discipline to advance it, some individuals outside the academy – and even some within it – have recently found utility in selecting bits of those failed attempts and applying them to various real-world decision-making scenarios. We have seen from this origin a massive growth in “nudges”, popularized by the eponymous popular-science book (Sunstein and Thaler, 2008) and taken up as a policy-setting aide by governments in the UK, Australia, and beyond (The Economist, 2017). The compass being
used in such applications of behavioural science –whether it has more of a psychology flavour or more of an economics flavour– is not grounded in a new theory of choice, but rather derives from specific empirical regularities that have been unearthed by behavioural scientists and then publicized as generally applicable. Examples of such empirical regularities include people’s tendencies to opt for default alternatives (Carroll et al., 2009), their over-weighting of the probabilities of low-probability events (Kahneman and Tversky, 1979), their tendency to conform to what others in their social environment are doing (Bernheim, 1994), and the significant impact on decisions not only when too little but also when too much information is provided to decision-makers (Scheibeheen et al., 2010). None of these observed aspects of individual choice derives from our core choice model.

While some have worried about the potential paternalism of government-sponsored nudging (e.g., Sugden, 2017), many policy applications of regularities that behavioural economists have promoted have been clearly welfare-improving. From that perspective alone, behavioural economists have clearly found a market niche, serving consumers at the coal-face of policy-setting who demand useful tools of thumb. Even better, the use being made of scientific validation tools such as randomized controlled trials breathes fresh intellectual life into what might otherwise have been moribund government policy-making endeavours. Although such initiatives are slowly gaining acceptance even within economics for advancing the technique of policy-setting (e.g., Chetty, 2015), by themselves they do nothing to advance economics as a science.

A second and more sobering consequence of the stagnation of our fundamental choice-making model has been the creation of more enemies of mainstream economics as a whole, who are sufficiently disillusioned with the inertia of its precepts that they choose to mount some form of public resistance. In this vein we have seen the round denigration of “rationality”, a poorly defined characteristic of our discipline’s most obvious whipping-boy, Homo Economicus (e.g., Thompson, 2013); popular science books disparaging the discipline (e.g., Keen, 2001, Rist, 2011); and new models of teaching economics led more by students’ interests in current problems than by the discipline’s need to train a new generation of open-minded, innovative theorists (of which the most successful example today is the Curriculum Open-access Resource for Economics, core-econ.org).

An alteration in the path that economics is now on may happen slowly, through an evolution of understanding and the development of a new common purpose amongst existing and emerging economists, or it may happen quickly, as a result of a revolution in which economics-as-we-know-it is rejected by the broader tribe of scientists, leaving a vacuum to be filled. If the latter alternative eventuates, it is hard to predict what will happen because no clear contender for the core model of choice exists to replace it. Chaotic competition amongst many competing models is the almost sure result of a revolution, and it is unclear whether the winner emerging from that competition would truly offer an advance over the model we have today.

The paradox of holistic overhaul
Aspiring theorists wishing to meet the above challenge may be tempted to carve out a piece of the choice puzzle and develop a more realistic model of just that part. Indeed, this has been the aim of most if not all serious modern attempts to advance economic theory beyond the simple sketch provided in the first paragraph of this paper. One might argue that the success of economics on the back of that recent history proves this incremental approach to be a viable one. Yet without a conceptual unification of the various extensions that are attempted, no wholesale overhaul of the human choice model can be produced: each extension relaxes one set of assumptions but retains assumptions that are relaxed in other extensions, and no one working in the academy of economics has a strong incentive to reconcile approaches that are prima facie mutually inconsistent.

One could instead opt for a direct approach, attempting a holistic extension to the individual choice-making problem by building into it many larger parts of the human puzzle all at once. This option is inevitably accompanied by an abandonment of mathematical formalism, due to the complexity of the problem. Attempts at this (e.g., Frijters, 2013) are generally published in forms other than journal articles and do not make much of an impact on the work of other theorists, in part because mathematical formalism is seen as the mark of good and useful theory in the modern day. Inevitably, the product of such an extension is moreover dangerously close to true interdisciplinarity, further reducing its potential impact on what economists working within their academic silo have reason to busy themselves with.

Promising frontiers
How then to begin? I briefly suggest below a few areas as promising grounds for nurturing a non-holistic and yet viable evolution of our theory of individual choice. Theoretical development in these directions may yield a re-furbished, internally consistent economic theory of choice that is able to accommodate the empirical regularities uncovered by behavioural economics. I have noted, where applicable, some very recent papers that indicate how these areas are starting to be recognized by economic theorists.

Models of attentional focus
Humans are only able to direct limited attention to items in their environment at any one time, yet our world is full of virtually infinite available streams of information. Our inability to attend to an infinite number of things combined with the incredible cost of mental effort to decipher incoming information (i.e., mental processing) means that the human organism has good reasons to have evolved mechanisms of efficiently allocating its attention.
These observations have already led to economists’ consideration of the concept of rational inattention, based on the scarcity of mental effort (e.g., Woodford, 2012, Dean et al., 2017). Going further in this direction, we could recognize that all decisions are based not on some objective reality but on our brains’ perception of that reality, a perception that itself derives in part from whatever mechanisms allocate our mental effort to information processing.

**The role of the unconscious mind**

Even if it is not seen as a threat to the principle of free will, the control of our unconscious mind over our attentional focus and framing of problems implies that any realistic model of decision-making should accommodate that part of our mental process. Accommodating something that lies outside our conscious control necessitates understanding what it –i.e., the unconscious mind itself– reacts to. An example of something to which there is evidence that the unconscious mind reacts, but that is rarely directly accommodated by economic theorists, is incoming information that is mis-aligned with the mental image of the self. One conjecture is that such information has the potential to harm our self-esteem and through that psychological mechanism to reduce our likelihood of success, and that this threat thereby creates a motivation for the unconscious mind to dream up ways to filter out and/or turn attention away from the information. The self-deceptive mechanisms investigated by brain scientists (e.g., Troisi, 2011) are consistent with this conjecture.

More generally, the unconscious mind may have evolved in part as a machine to understand or justify the actions we have already taken (Wegner and Wheatley, 1999), implying that internal reflection is overlaid upon the problem of choice. Even more radically, as recently proposed in the neuropsychology literature (Bear and Bloom, 2016), neither our conscious nor our unconscious minds may have what a layperson would define as free will.

With more effort towards learning what scientists in other areas are discovering about these and other unconscious processes, economists could capitalise on the discipline’s analytical strength to revolutionise our understanding of choice in the presence of an unconscious mind. Not only can accommodating a role for unconscious information filtering help to deepen our understanding of how individuals mentally frame problems, but adding the ability to reflect internally and relaxing the principle of free will could result in a view of choice that is less encumbered by the hubris implicit in economists’ much-maligned, allegedly autonomous and amoral individual agent.

**The dynamic self**

While regime switching in the area of preferences, information, or other components of a choice problem is a brute-force way to accommodate the fluidity of self, like many other attempts at advance beyond the core choice model, it is in essence tacked on. Simple introspection reveals that in reality a person has no fixed preference map in her head, her sense of self and hence her desires vary continuously over time, and her constraints evolve minute-by-minute. The absence of a fixed self in particular, while anathema to the assumptions used in most micro-models of choice over time, characterises not only our acute developmental and learning phase (childhood) but our adult lives as well. Our fluidity of self enables us to fall into and out of love –an area of human cognition crucial for economists to understand, as argued in Frijters and Foster, 2017– and is arguably a core element enabling us to adapt so successfully to our ever-changing environment that we have emerged as the dominant species on Earth. To move our model of individual decision-making away from the concept of a fixed self while retaining its analytical tractability would represent a major theoretical advance.

**Conclusion**

Our discipline began with a broad but non-technical view of society crafted by thinkers like Adam Smith, John Locke, Jeremy Bentham and Alfred Marshall, each of whom implicitly acknowledged many inputs into individual decision-making, of which only a small number have been incorporated into today’s standard canon. While behavioural economics has recently delivered a bountiful array of empirical insights that excite and fascinate curious people in and out of the discipline, a gauntlet is thrown down by these findings to economic theorists: how can we explain these findings, to a degree that is both intellectually satisfying and internally consistent?

I argue here that while a broad, empirically-validated expansion on multiple fronts might be the ideal, such an attempt is unlikely to be observed and even less likely to have significant impact in the discipline of economics due to the siloization of social scientific research that characterizes the scholarly arena today. As a second-best option, particular behavioural fronts are suggested that seem *ex ante* promising as conceptual theatres for expanding our model of choice, slowly but steadily, in the direction of the realism championed by behavioural economists and in service to the preservation of economics as a science. Such a theoretical advance would enable the discipline to continue to deliver socially helpful applications into the field, long after the current nudges-and-heuristics-led “behavioural insights” buzz has waned.

As early as 1987, Nobel Laureate Herbert Simon urged that economic theories be tested and refined in light of empirical evidence from the nascent sub-field of behavioural economics, instead of constructed based on armchair assumption, in order to ensure “the progress of economics” (Simon, 1987, p. 225). In his words from thirteen years later (Simon, 2000):

> There remains a large task of organizing our picture of economic and social processes and adding the new facts needed to shape the theory in an empirically sound way. It is also urgent that new tools now available for conducting empirical inquiry and constructing models be incorporated in
social science graduate education.

Acknowledgments

I thank Paul Frijters for comments on an early draft of this paper.

References


Bronowski, Jacob (1973). The Ascent of Man, Episode 11: “Knowledge or Certainty”. BBC.


Thompson, Derek (2013). “The irrational consumer: Why economics is dead wrong about how we make choices”. The Atlantic, January 16.


