

People and Policy: Behavioural economics and its policy implications

Alex Dickson and Marco Fongoni, University of Strathclyde

Abstract

The standard approach of economic modelling and policy analysis is to assume that people are *homo economicus*: a rational and self-interested economic man. However, a well-documented body of evidence suggests that human decision makers depart from this assumption in several interesting ways. This paper explores three key behavioural aspects – loss aversion; present-biased preferences; and social preferences. It considers the evidence to support them; how behavioural economists have modified economic models to capture the implications of each modification; and how they offer new insights on the scope and efficacy of a wide range of policy interventions.

Keywords

Behavioural economics; loss aversion; reference-dependent preferences; present-biased preferences; social preferences; procrastination; reciprocity; public policy

I Introduction and background

Economists seek to describe the complexity of the real world by using abstract models of reality. At the heart of most models in economics are people who make decisions, and economists therefore have to make assumptions about how they behave. The ‘standard economic model’ (“SEM”) in the neoclassical tradition is populated by individuals that satisfy the assumptions of ‘*homo (or femina) economicus*’ – rational economic (wo)man – people who understand their preferences; are rational; obey the axioms of ‘utility theory’; are self-interested; and discount for time in a consistent way. There are some good reasons to pursue such an approach. Economists try to develop general models of reality that help explain – and predict – (average) behaviour. Its behavioural assumptions don’t have to concur with how we conceive that people actually make decisions, but their predictions should be consistent with observed behaviour *on average*. Given the need for generally applicable models, the SEM has often served economists – and by extension – society rather well.

In recent decades, however, there has been something of a ‘behavioural revolution’ in economics. By observing how people make decisions in experimental settings, some well-documented anomalies of behaviour to the SEM have been identified, bringing into question the validity of some of its assumptions. These have given behavioural economists the inspiration to incorporate psychological insights into economic models in an attempt to better capture the way people make decisions. This paper presents the basic idea behind three key modifications to the standard model: loss aversion; present-biased preferences; and social preferences. For each we consider an illustration of such behaviour; discuss the evidence that suggests people do behave in this way; consider how economic models have been adapted to capture these behaviours; and discuss some of their policy implications.

First let us define rationality. For the purpose of this paper the following definition is useful: a rational individual, faced with multiple options, will choose the one(s) that give her the highest payoff. This does not rule out, for example, an individual acting in the best interests of society even if it is not in her material interest to do so; the drivers of 'value' to individuals might be more than simply their own material gain, which are turned on when pro-social decisions are taken. Using this definition allows us to distinguish between: a) behaviour that is inconsistent with the standard model because people are not optimising and are being irrational; and b) departures from the standard model that involve rationalisable behaviour where the carrier of 'value' is not simply a standard utility function, but something that is more behaviourally influenced. Our focus is very much on the latter.¹

Loss aversion is the idea that losses loom larger than equivalent-sized gains. If people are loss averse and they are exposed to losses then they face particularly sharp incentives to avoid incurring them. This seems to be a widely reported characteristic of human behaviour which gives rise to many interesting conclusions. For example, to incentivise someone to do something, rather than rewarding them upon completing the task, a loss averse individual is more likely to complete it if they are paid first and face being required to repay the payment if they fail.

Present-biased preferences capture the idea that when considering decisions that have consequences in the future, the present is particularly salient and has a large impact on the decision at hand. This can give rise to time inconsistency: people may plan to do something in the future, but after the passage of time when that future becomes the present they change their plans because decision making in the here and now takes a different shape to decision making regarding the future. Models that incorporate this behaviour can shed light on procrastination and self-control problems, and, more importantly, mechanisms to overcome these.

Social preferences allow economists to capture other-regarding elements to preferences that might also be drivers of behaviour. One of the most convincing of these is reciprocity: if an act of kindness is done to an individual they will gain some pleasure from reciprocating kindness with kindness, but will suffer displeasure if they are unkind in return. Reciprocity can temper the behaviour of people in pursuit of self-interest. For instance, in a situation in which there are negative externalities (i.e. where people, in doing what is best for themselves, do harm to others) the prediction of the standard model is market failure, promoting an institutional response to correct it; but reciprocity may provide a self-correction mechanism that limits the extent to which people harm others, particularly if this is encouraged by society.

By considering more descriptive accounts of people's behaviour, the analysis of economic environments often leads to different conclusions than those drawn from a standard analysis. Consequentially, consideration of behavioural ideas gives rise to implications for policy. Affiliated with the UK Cabinet Office is the Behavioural Insights Team (established in 2010) that works to consider the implications of insights that stem from behavioural economics for public policy², demonstrating the centre stage that behavioural economics is taking in the public policy arena. This paper aims to provide some insight into

¹ Very accessible discussions of the former are included in Thaler and Sunstein (2008); Kahneman (2011), which is a highly recommended read for anyone with an interest in behavioural economics; and Ariely (2009).

² For the interested reader, their 'MindSpace' report on "influencing behaviour through public policy" presents an interesting dialogue on the broad policy implications of many aspects of behavioural economics.

the economics behind behavioural policy interventions to help readers make an informed assessment of their importance.

II Loss Aversion: losses loom larger than gains

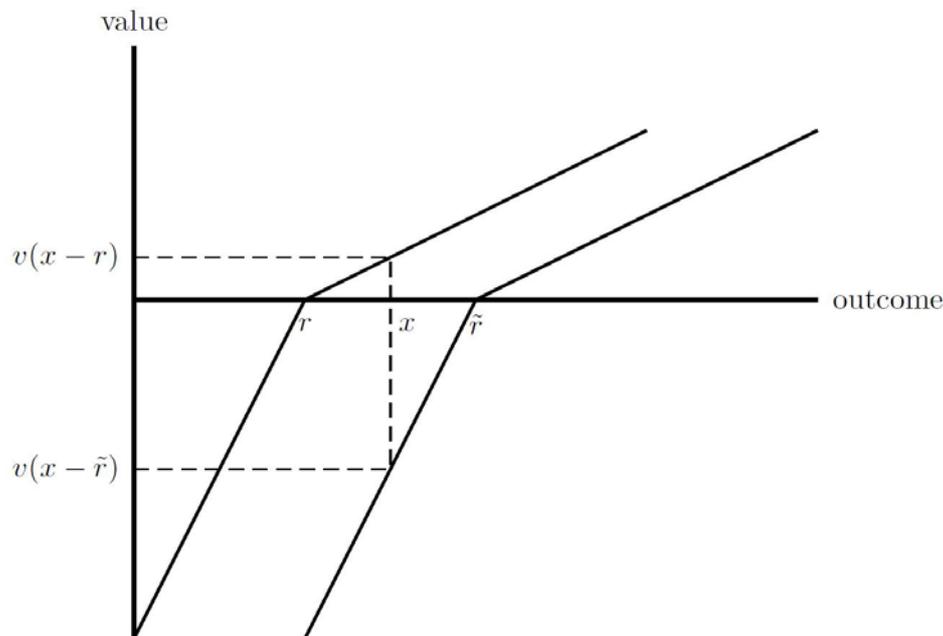
A well-documented feature of human perception is that losses loom larger than equivalent-sized gains: the pleasure felt from gaining an amount of money is less than the displeasure of losing that same amount of money. This cannot be consistently captured in the SEM because it assumes that only final outcomes are the carriers of value. Loss aversion suggests that how an outcome compares to a reference point influences how it is valued.

To capture this reference dependence, a model of preferences needs to not only consider the magnitude of outcomes, but also how they compare to a reference point. An outcome is coded as a gain if it exceeds the reference point, and as a loss if it falls short of it, and individuals seem to intrinsically assess gains and losses differently. Loss aversion was introduced as a central feature of Kahneman and Tversky's (1979) 'Prospect Theory' where the reference point was taken to be *status quo*, the individual's current state.

The way loss aversion is captured in an economic model is straightforward. In standard preference representation models individuals have a utility function defined over outcomes that assigns to each outcome x a utility number $u(x)$. Decision making involves comparing the utility of different options and choosing the one that gives the highest utility. To allow for loss aversion the outcome x is compared to a reference point r , and a standard utility function is augmented with a 'gain-loss' utility $v(x-r)$ which can differ depending on whether x exceeds the reference point r (in which case $x-r > 0$) or falls short of it (in which case $x-r < 0$). An individual is loss averse if, for some $y > 0$, $-v(-y) > v(y)$: the gain-loss utility is steeper in the domain of losses than in the domain of gains, so there is a kink at the reference point.³ An individual's overall payoff from the outcome x is then a combination of the utility of that outcome and the gain-loss utility: $U(x;r) = \eta u(x) + v(x-r)$ (where η is just a weighting parameter).

Figure 1 plots two gain-loss utility functions for different reference points, r and $\tilde{r} > r$. Suppose an individual's reference point is r , then if an outcome represents a gain relative to this reference point she moves up the relatively shallow value function to the right of r ; conversely, if an outcome falls short of r and so is judged to be a loss she moves down the relatively steep value function to the left of r . The disutility from a loss is larger than the utility from an equivalent-sized gain, which means that individuals face a sharp incentive to avoid losses. This figure also illustrates that how an outcome is evaluated depends on the reference point to which it is compared: relative to the reference point r the outcome x in the figure is a gain, but relative to the reference point \tilde{r} it is a loss and is evaluated as such.

³ Prospect theory also incorporates 'diminishing sensitivity': the effect of marginal gains and losses is smaller the larger is the gain or loss under consideration, which implies the gain-loss utility function is concave in the domain of gains and convex in the domain of losses. For simplicity of exposition, we consider only loss aversion here.

Figure 1: reference dependence and loss aversion.

The evidence base that supports the concept of loss aversion is large and convincing. Early evidence came in the form of the ‘endowment effect’ (Kahneman et al., 1990), which noted that, on average, individuals require more compensation to give up a good that they have gained entitlement over than they would be willing to pay to acquire it. After being endowed with a good an individual perceives ownership of it so if they were to sell it they would suffer a loss, losses are felt strongly, and so a high level of compensation is required to sell the item. This is easily seen in the diagram: if r is the individual’s initial situation and \tilde{r} incorporates acquisition of the good, the individual’s willingness to pay for the good is given by the height of the left value function evaluated at \tilde{r} ; by contrast, their willingness to accept if they are endowed with the good and have incorporated possession of it into their *status quo*, is given by the amount by which the right value function reduces when moving from \tilde{r} to r . Loss aversion implies the willingness to accept exceeds the willingness to pay.

Loss aversion has been found in many experimental settings, as well as in field studies: for example, in seller behaviour in the housing market (Genesove and Mayer, 2001); in physician behaviour (Rizzo and Zeckhauser, 2003) and police performance (Mas, 2006) in relation to pay; and in labour supply decisions (Camerer et al., 1997; Crawford and Meng, 2011). The evidence by no means suggests that every individual is loss averse; indeed, there may be distinct groups who are not. For example, List (2004) found that experienced market participants did not exhibit loss aversion whereas those with little experience did, but professional golfers – who have vast experience and are subject to large stakes – do tend to exhibit loss aversion (Pope and Schweitzer, 2011). The phenomenon seems to be robust enough that it should be given careful consideration by economists when undertaking modelling exercises so that they can infer whether people do exhibit loss aversion, and if so assess its implications.

People make decisions based on the incentives they face: they weigh up the benefits and the costs of undertaking an action, compare the two, and act appropriately. If people are loss averse, they evaluate gains and losses differently, which means they face different incentives in the domain of losses to those in the domain of gains: the incentives to avoid losses are sharper. By altering the decision-making environment in a way that changes the reference point, people can be exposed to losses which may consequently lead them to change their behaviour. As such, if people are loss averse, the 'choice architecture' within which decisions are made may matter for the decisions that are taken.

'Drip pricing' is one example of companies trying to influence reference points that recently commanded the attention of an Office of Fair Trading Report (OFT, 2010). Drip pricing involves the buyer not seeing the total price for a good when making a purchase decision; a base price is augmented with additional charges before the purchase can be made (for example, airline tickets with compulsory debit and credit card fees or theatre and concert booking fees). An experienced consumer obeying the assumptions of *homo economicus* would see through such pricing and make decisions based on the total price. For a loss averse consumer, once they have made the decision to make a purchase their decision shifts their reference point; subsequently pulling out of the purchase would imply a loss that the consumer is keen to avoid, making them more willing to pay the additional charges than they would be before they made the decision to purchase. Logically, consumers who are subjected to drip pricing may end up paying a total price for the good that is more than they would have chosen to pay had the total price been cited before their purchase decision was made.⁴

For further illustration, a recent study considers the effect of appealing to loss aversion when trying to incentivise teachers to be more efficacious (Fryer et al., 2012). Whilst financial rewards following good performance had no effect on teacher performance, when teachers were paid up front and asked to return the money if their students under-performed, students achieved significantly better results in a test.⁵

Later developments of models of reference-dependent preferences have allowed the reference point to be influenced by various features of the environment in which decisions are made: for example, by expectations (Shalev, 2000; Koszegi and Rabin, 2006); and by the outcome achieved by a reference group (e.g. Card et al. (2012)). This has further implications for behaviour and policy. For example, what seems like innocuous behaviour by firms might have non-obvious consequences. If a firm offers a good at a sale price this might act to manipulate a consumer's future willingness to pay if they buy at the sale price and their expectations about consumption are influenced by what they consume. The reason is that if, at the next purchase opportunity, the consumer does not consume the good they will suffer a loss relative to their expectations, so are willing to pay a higher price to avoid the loss; their willingness to pay has increased, perhaps to a level that exceeds the regular price (Heidhues and Koszegi, 2014). A further example concerns aspirations that can be influenced by the reference group to which individuals compare themselves. If a person's reference point is influenced by those around her and that group is currently enjoying better outcomes than the individual, then by remaining in her current state the individual would be suffering a loss relative to her aspirations. Since realising this loss would give

⁴ For the reader interested in competition issues, there is a further OFT survey on consumer behavioural biases in competition (OFT, 2011).

⁵ Following a similar logic, the website StickK.com allows users to expose themselves to both monetary losses and social pressure with the aim of helping people achieve their goals; simple, but effective!

substantial displeasure, she faces a sharp incentive to achieve her aspirations. As such, it is more likely that lower performing groups will make achievements if they are exposed to influences that cause them to aspire to more.⁶

The implications of loss aversion for behaviour and therefore policy are numerous and varied. The long-term consequences of policy interventions should be borne in mind, however, because loss aversion implies there may be an asymmetric response to the introduction and elimination of policy measures if those policy measures generate entitlement effects that get incorporated into reference points, as might conceivably be the case with changes to minimum wages (Fehr et al., 2009).

III Present-biased preferences: the best laid plans; what happens ‘when tomorrow comes’

We now turn to discuss inter-temporal decision making, in which the consequences of a decision occur in multiple periods. When outcomes occur over time, economists discount those in the future to reflect the idea that people are impatient: other things being equal, people prefer to receive rewards sooner rather than later. The standard ‘discounted utility model’ proscribes that the present discounted utility of a stream of outcomes occurring between now and time period T is

$$u(x_0) + \delta u(x_1) + \delta^2 u(x_2) + \dots + \delta^T u(x_T)$$

An individual's discount factor is $\delta \leq 1$: patient individuals have large discount factors and don't discount the future too heavily; impatient individuals have small discount factors so outcomes that occur in the future are heavily discounted. Regardless of the level of impatience, the discount factor for some period t is δ^t , which is smaller the larger is t : outcomes farther into the future are more heavily discounted. A key feature of this model is that the additional impatience between any two consecutive periods is the same regardless of the time horizon: for example, the discount factor applied today is 1 (no discounting) and that in the following period is δ (the additional discount applied is $\delta/1 = \delta$); the discount factor 7 periods hence is δ^7 and 8 periods hence is δ^8 (the additional discount applied to the later period is $\delta^8/\delta^7 = \delta$).

A feature of real-life decision making is that many people exhibit a present bias: when considering decisions at short time horizons that involve the present people tend to exhibit a high degree of impatience, but when thinking about consequences at long time horizons people tend to be less impatient.⁷ An example will help to illustrate. When asked whether they would choose to receive £100 now or £110 next week, people often choose the ‘smaller sooner’ reward revealing that they are relatively impatient at short time horizons; when asked whether they would prefer £100 in 1 year or £110 in 1 year and 1 week (separated by the same amount of time as the first choice), people tend to choose the ‘larger later’ reward. However, if £110 in 53 weeks is preferred to £100 in 52 weeks then $\delta^{53}110 > \delta^{52}100$ (δ is the weekly discount factor in this example), so $\delta > 10/11$, but if £100 now is preferred to £110 next week then $100 > \delta 110$, so $\delta < 10/11$. But in the standard discounted utility model individuals have a single discount factor that is meant to capture their time preferences; such behaviour is inconsistent with the standard model.

⁶ Genicot and Ray (2014) undertake a careful analysis of aspirations in relation to aggregate outcomes, and in particular consider the effect on inequality of only moderate aspirations being effective in incentivising people to achieve more (due to diminishing sensitivity in the gain-loss value function).

⁷ Frederick et al. (2002) discuss the body of evidence that supports this idea.

This pattern of behaviour is an example of 'time inconsistency' because an individual that makes such choices would plan that in a year's time they would wait and take the 'larger later' reward; however, after the passage of a year if asked the same question again they would choose to take the 'smaller sooner' reward, since that is what they would choose now. What seems to be the case is that for decisions at short time horizons, and particularly for decisions that involve the present, people are more impatient than they are at longer time horizons. Individuals exhibit a 'present bias' that gives rise to the time inconsistency: plans can be made at long time horizons, but when it actually comes to making the decision to enact that plan (after the passage of time) this happens in the present and the evaluation of that decision changes.

To try to capture this present bias behavioural economists have turned to alternative models of discounting for time. Perhaps the most appealing, because of its simplicity, is Laibson's (1997) 'quasi-hyperbolic' discounting model.⁸ This introduces a single additional parameter, $\beta \leq 1$, into the discounting model that lowers the weight of all periods in the future by changing the discount applied to any period $t \geq 1$ from δ^t to $\beta\delta^t$; put another way, it increases the (relative) weight attached to the present. Under so-called (β, δ) –preferences the present discounted utility of the same stream of outcomes considered above is:

$$u(x_0) + \beta\delta u(x_1) + \beta\delta^2 u(x_2) + \dots + \beta\delta^T u(x_T)$$

This is a really simple, but very clever, modification since it captures the present bias that gives rise to time inconsistency in a single present-bias parameter, β .

Take the example used previously: if an individual prefers £110 in 53 weeks to £100 in 52 weeks then $\beta\delta^{53}110 > \beta\delta^{52}100$ so $\delta > 10/11$ (the β s cancel). If £100 is chosen now over £110 in 1 week then $100 > \beta\delta 110$, which is consistent with $\delta > 10/11$ so long as β is sufficiently small, i.e. there is a strong enough present bias to preferences. By capturing present bias, quasi-hyperbolic discounting allows us to understand the extent to which people might plan to do things ("in one year I will wait another week to get £110") but then renege on those plans after the passage of time (when one year has passed "I want £100 now").

The behavioural insight of present bias, which implies time inconsistency and is captured by the quasi-hyperbolic discounting model, is important because it can explain the mechanism through which people suffer from self-control problems that are associated with the behaviours of '*procrastination*' (putting off until later a costly activity) and '*prepropiation*' (bringing forward rewarding activities for instant gratification). To see this, consider a costly activity that brings subsequent benefits (such as exercising)⁹: the cost is c and the benefit which is received in the following period is b . An individual, when planning to do exercise in, say, 5 days' time will do so if the (appropriately discounted) cost is less than the benefit, i.e. if $\beta\delta^5 c < \beta\delta^6 b$, which will be the case so long as the benefit is sufficiently greater than the cost and the individual is not too impatient ($c < \delta b$). However, after the passage of 5 days when it actually comes to doing the exercise the cost to be incurred on that day is particularly salient, and they

⁸ A different approach to modelling present-biased preferences considers individuals as having dual selves – a long term planning self and a sequence of short term decision-making selves – see Fudenberg and Levine (2006).

⁹ The notation of DellaVigna (2009) is followed in this discussion.

will exercise only if $c < \beta\delta b$. If the individual suffers from a present bias and $\beta < 1$ it may very well be the case that whilst $c < \delta b$ implying the individual plans to do exercise, $c > \beta\delta b$ so when the planned exercise day arrives they choose not to.

People with a present bias plan to do things that subsequently don't get done: they suffer from a self-control problem. When it comes to carrying out the plan, costs must be incurred now, and whilst there are subsequent benefits, the salience of the cost due to the present bias means the benefit may be insufficient to motivate them to do it, even if they planned to. People procrastinate; they put off until later something that should, in their long-term interests, be done today.

The extent to which people understand their self-control problems can be captured by considering an individual's perception of their true present bias, $\hat{\beta} \geq \beta$. Someone is said to be sophisticated if they perfectly understand their present bias and subsequent self-control problem (i.e. $\hat{\beta} = \beta$), and is said to be (partially) naïve if they over-estimate their self-control ($\hat{\beta} > \beta$). A sophisticated individual understands that they have a self-control problem and will not be surprised when they subsequently procrastinate. Naïve individuals will be surprised in some circumstances: they will make plans fully anticipating following them through, not recognising that they over-estimated their ability to stay in control.

In an interesting study by DellaVigna and Malmendier (2006), the gym attendance patterns among members of health clubs in the US were considered. Despite there being a pay-as-you-go membership option, the *ex post* average price per visit of members who signed up to a monthly contract was substantially higher than the pay-per-visit fee. This suggests that at the planning phase, when deciding which contract to purchase, members over-estimate their gym usage, consistent with people exhibiting present-biased preferences and being partially naïve to their self-control problem.

By accounting for present bias in economic models we can understand and begin to think about the consequences of procrastination. One fairly obvious consequence is that, if people are faced with a decision problem in which a choice has to be made, they may procrastinate over making a decision even though the transaction cost is small and there are substantial subsequent benefits from making a decision (such as joining a pension scheme). This implies that default options for decisions will matter and that careful consideration should be given to what the default option is. In an early study of pension choices, Madrian and Shea (2001) considered the effect of an opt-out system (with a default saving plan), rather than an opt-in system: under the opt-out system 90% joined immediately and 98% after 3 years; whereas with the opt-in system only 20% joined after 3 months and 65% after 3 years. Thaler and Sunstein (2008) in their book *Nudge* champion the idea of default options as a simple, almost costless, policy tweak that can have dramatic effects, and there is a significant body of evidence in the recent literature to suggest that this is so.

By a similar logic to the idea behind procrastination, people may also engage in enjoying the instant gratification of a pleasurable activity that gives rise to subsequent costs, even though they plan not to do so, i.e. they preappropriate rewards. It seems, then, rather surprising that the UK Government in the 2014 Budget announced flexibility for pensioners in respect of their retirement savings, allowing them to withdraw their pension pot upon retirement. All very libertarian, but whilst pensioners may *plan* to be sensible with their money at the point of retirement, when it comes to acting on their plan they will require great strength of will; the temptation to preappropriate for instant gratification is clear and present.

For instance, a recent study linking the level of credit card borrowing with present bias suggests that people with present bias do prepropriate consumption, more so if their present bias is stronger (Meier and Sprenger, 2010). Understanding the extent of people's present bias through the lens of a model incorporating quasi-hyperbolic discounting could allow economists to predict the additional spending that will take place around retirement, and the subsequent gap in income that may emerge in later years. Such analysis should certainly be borne in mind by policy makers when assessing the implications of such radical policy changes in, for example, pension arrangements – especially as any future income shortfalls may in all likelihood be borne in part by the State.

A further point of consideration for policy makers to bear in mind relates to commitment. If people are sufficiently aware that they suffer from a self-control problem then at the planning stage they will have some idea that they will engage in procrastination/preappropriation at the point of carrying out decisions, and might be able to take actions to help them stay in control. For example, a field study of savings behaviour by Ashraf et al. (2006) reported in an aptly-titled paper “Tying Odysseus to the mast” found that savings rates of those individuals offered special accounts that allowed for some commitment to monthly saving were significantly higher than for individuals left to their own devices. In relation to gym attendance discussed earlier, another explanation is that people buy a monthly membership to sink the financial cost of going to the gym that encourages them to go more than they otherwise would. If people are presented with an opportunity at the planning stage to either tie their hands in respect of their future decision, or to alter their future incentives making it more favourable to complete the task they plan to undertake when the time arrives, individuals who are sufficiently aware of their self-control problems can overcome them. Providing such commitment opportunities is a simple way to help people make better decisions that are in their long-term interests.

IV Social Preferences: self-interest or the common good; competition or co-operation?

Economic models have traditionally made the assumption that individuals care only about their own material self-interest. There is, however, a body of convincing evidence suggesting that some people care about more than simply themselves.¹⁰ Behavioural economists have worked to capture this by considering, in addition to material self-interest, ‘other-regarding elements’ to preferences. The pleasure someone derives from an outcome depends not only on the material value to that individual, but how achieving that outcome sits within a societal context. For example, people might derive pleasure from acts of altruism or displeasure from taking actions that lead to increased inequality or that contrast with accepted norms of behaviour. These issues, and their implications, are perhaps most easily illustrated using the ‘ultimatum game’ (Güth et al., 1982). In the ultimatum game there are two individuals, a proposer and a responder; the proposer is given a ‘pie’ (e.g. an amount of money), and has to make an offer of a share of this pie to the responder; the responder then has to decide whether to accept this share, in which case the proposer and responder leave with the suggested shares, or reject it, in which case both leave with nothing (players know the rules before play begins). Under the SEM the responder should accept any share, no matter how small; the proposer should reason that this is the case and so should offer a tiny share of the pie retaining the vast majority for herself, and the responder should accept.

¹⁰ This literature is too vast to survey here; Fehr and Fischbacher (2002) give a good overview.

In general, this is far from what is observed when the ultimatum game is played in experimental settings (see, e.g., Thaler (1988)). Proposers tend to make significant offers to responders, and responders sometimes reject substantial offers. This behaviour can be explained in a number of ways. 1) Proposers may be *altruistic* and derive a 'warm glow' from the act of giving something to the responder (Andreoni, 1990); they are willing to sacrifice their personal material payoff to experience the warm glow. 2) Both proposers and responders might be *inequity averse*: receiving more or less than others in society gives some disutility (Fehr and Schmidt, 1999). This implies people are willing to sacrifice their material payoff to achieve more equal outcomes, so proposers may propose substantial shares to ensure the outcome is more equal, and responders might reject particularly unequal shares to avoid increased inequality despite foregoing the share that was offered. 3) Responders might exhibit *reciprocity*: some pleasure is gained by reciprocating fair behaviour with fair behaviour (positive reciprocity) and unfair behaviour with unfair behaviour (negative reciprocity), and likewise displeasure is felt if an individual behaves unfairly to those they perceive have treated them fairly, and vice versa (Rabin, 1993; Falk and Fischbacher, 2006). If a responder engages in reciprocity and she judges the proposer's offer to be unkind she may reject the offer since the pleasure she derives from reciprocating with an unkind act (rejecting leaves the proposer with nothing) is larger than the material payoff from acceptance of the small offer. Fearing that unfair offers will be rejected tempers the behaviour of the proposer since she needs to ensure her offer is perceived to be fair for it to be accepted. 4) Both proposers and responders might be influenced by social norms (see, e.g., Young (2007)); some disutility will be felt by straying from what is deemed by society to be acceptable behaviour. If sharing is a social norm, both proposers and responders may take actions to concord with this accepted behaviour (proposers by making substantial offers, responders by rejecting insubstantial offers) even if it reduces their material payoff.

Having summarised some of the main ideas that behavioural economists have advanced in considering social elements of behaviour, we now turn to consider some of the implications of taking social preferences into account. By accounting for the influence of such motivations we can capture the idea that individuals are potentially constrained in their actions by a comparison with both internal and external norms of behaviour. Reciprocity and social norms seem particularly convincing explanations: if an individual is motivated by reciprocity she may not pursue activities that are in her self-interest if this involves being unfair to someone who is perceived to have acted in a fair way, as this violates an internal norm of behaviour. Likewise, if individuals are motivated by social norms they may not pursue acts that are in their self-interest if the actions contrast with what is believed to be acceptable behaviour by society, i.e. external norms.

It is useful at this point to outline two scenarios from game theory that provide an effective representation of many of the environments encountered in settings where social behaviour is important. The first is the classic 'Prisoners' Dilemma', in which two (or more) individuals have to choose whether to 'cooperate' or 'defect', all individuals face a material incentive to defect regardless of what they think others will do, and when any individual acts in their own best interests they do substantial harm to others (i.e. there are negative externalities). If individuals pursue their own self-interest they choose to defect, but then each will be doing harm to the other and the collective outcome will be inferior to what could be achieved by cooperating. This is represented in the left panel of Figure 2; given the incentives the equilibrium involves both individuals defecting so each achieves a payoff of 1, even though both players could achieve a payoff of 2 if they cooperate. In the right panel of Figure 2 a somewhat different scenario

is considered, in which the only change is that the payoff from defecting when the other cooperates has been reduced from 3 to 1. This is a 'Stag Hunt' in which each player can achieve a relatively low payoff regardless of what their adversary does by defecting (the analogy is to go hare hunting; a small but guaranteed gain); larger gains are possible by cooperating, but they can only be achieved if *both* players cooperate (the analogy is to go stag hunting; the gain is large, but a stag cannot be caught single-handedly). From a situation of mutual cooperation, there is no incentive for any player in this game to defect, so in contrast to the prisoners' dilemma mutual cooperation is an equilibrium. Mutual defection is also an equilibrium, so in the parlance of game theory there is a coordination problem, but one that should be relatively straightforward to solve since the equilibrium in which players cooperate payoff-dominates that in which they defect: if people have sufficient confidence that others will cooperate, they will choose to cooperate as well.

Figure 2: A Prisoners' Dilemma in the left panel, and a Stag Hunt game in the right panel. In each game two players A and B decide whether to cooperate or defect. A chooses between the rows of the matrix, B the columns. Each cell of the matrix corresponds to a potential outcome, and in each cell the payoffs to the two players are listed, the first number being the payoff to player A, the second that of player B.

A Prisoners' Dilemma		A 'Stag Hunt'	
		B	
		Cooperate	Defect
A	Cooperate	2,2	0,3
	Defect	3,0	1,1
A	Cooperate	2,2	0,1
	Defect	1,0	1,1

In many scenarios in the real world pursuing self-interest harms the common good and the material incentives people face concur with those in the prisoners' dilemma. If people do pursue their material incentives this is at the expense of society, which leaves society with a problem. Prisoners' dilemma-type scenarios imply a market failure, to which there should be an institutional response to enforce the cooperative outcome to make society better off: decisions should be taken out of the hands of the members of society through acts of formal governance.

However, if we take the view that people are motivated not only by material self-interest but also by internal and external norms of behaviour, the picture is perhaps rosier. In a prisoners' dilemma, when an individual pursues their self-interest by defecting they do harm to others (since there are negative externalities). If doing so conflicts with either internal or external norms of behaviour there will be a psychological cost. This reduces the payoff of defecting from 3 to, say, $3 - c$; if this cost is large enough the overall payoff will fall short of 2, in which case the incentive to defect disappears. If the psychological cost exceeds 1 then the structure of the game in psychological payoffs is not a prisoners' dilemma, but a stag hunt, in which the cooperative outcome is an equilibrium. So long as the material benefits from defecting on cooperation are not too large and the psychological cost of defecting on cooperation is large enough, members of society will find it in their own interests to act in the best interests of society, so long as they believe others will. If this is the case then the institutional response should be to promote cooperation by supporting reciprocity, nurturing social norms and communicating the extent of

cooperative behaviour, which is a very different response to formal governance that requires legal or other formal sanctions.

Whether individuals will act pro-socially in situations that materially look like a prisoners' dilemma but where their actions are tempered by the constraints of internal and external norms of behaviour depends on the size of the material gain, and the psychological cost of defecting (amongst many other factors). If people are unconstrained by social preferences they will never act pro-socially. However, if people are subject to fairness concerns embodied in reciprocity, and are influenced by social norms that contrast with pursuing self-interest at the expense of others, society may very well be able to be self-regulating without recourse to formal governance. From a vast body of research investigating this issue, we have some understanding of the drivers behind pro-social behaviour (see, for example, Ostrom (2005)), but there is no widely applicable universal truth. In situations where policy makers identify a market failure and consider an institutional response of formal governance, they should also carefully consider whether people's intrinsic motivation can be appealed to and strengthened to enable society to resolve its own problems without recourse to formal sanctions. In addition, where people are behaving pro-socially, policy makers should be careful not to introduce material motivations or issue instructions on how to behave that may 'crowd-out' existing intrinsic motivation and hence inadvertently discourage people from behaving pro-socially (Frey, 2012).

V Conclusions

The standard economic model does a good job of explaining how some people behave, and given its generality and tractability has served economists – and society – well. However, the well-documented and consistent evidence pointing out its flaws should not be ignored if economists are to provide more accurate models of behaviour and hence help policy-makers design policy 'that works'. By incorporating insights from psychology, experiments and field observations, more accurate models of behaviour can be constructed that can be used to better understand how people act in economic environments and how they might respond to changes in those environments through policy interventions.

However, behavioural economics is not without its critics, some of whom are behavioural economists themselves, who are critical of some of the methodologies used and their consequent insights. There is, for example, concern that the findings from experiments that are conducted in decision laboratories are not generalisable to the real world and so only deductions from field studies should be used to inform insight into how people actually behave. This is a healthy debate that promotes high-quality academic work in behavioural economics; hence there are an increasing number of field studies, and the issue of control in experimental settings is under particular scrutiny.

Developing a better understanding of people's behaviour is critically important in assessing the economic value of policy interventions and in identifying new policy approaches and levers through which change can be enacted. A main focus of this paper has been to explain the economic models that capture behavioural insights, because an understanding of the mechanisms at work is crucial to understanding the importance of taking departures from the standard approach into account and the applicability of ideas to new and different settings.

Behavioural economics has certainly come of age as an academic sub-discipline in economics over the past two decades. Most encouragingly it has extended its influence into a wide variety of critically important policy fields, be that in labour market policy, social policy, pensions and savings policy or competition policy. Given its relatively recent existence, behavioural economics has a strong and growing presence in the analysis and evaluation of public policy in the UK and internationally – and this will only grow as fundamentally most public policy is about effecting change in peoples' behaviours and in societal outcomes.

References

- Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *The economic journal*, 100, 464-77.
- Ariely, D. (2009). *Predictably irrational*. Harper Collins.
- Ashraf, N., D. Karlan and W. Yin. (2006). Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. *The quarterly journal of economics*, 121, 635-72.
- Camerer, C., L. Babcock, G. Loewenstein and R. Thaler. (1997). Labor supply of New York City cabdrivers: one day at a time. *The quarterly journal of economics*, 112, 407-41.
- Card, D., A. Mas, E. Moretti, and E. Suez. (2012). Inequality at work: the effect of peer salaries on job satisfaction. *American economic review*, 102, 2981-3003.
- Crawford, V. and J. Meng. (2011). New York City cab drivers' labor supply revisited: reference-dependent preferences with rational-expectations targets for hours and income. *American economic review*, 101, 1912-32.
- DellaVigna, S. (2009). Psychology and economics: evidence from the field. *Journal of economic literature*, 47, 315-72.
- DellaVigna, S. and U. Malmendier. (2006). Paying not to go to the gym. *American economic review*, 96, 694-719.
- Falk, A. and U. Fischbacher. (2006). A theory of reciprocity. *Games and economic behavior*, 54, 293-315.
- Fehr, E., L. Goette and C. Zehnder. (2009). A behavioral account of the labor market: the role of fairness concerns. *Annual Review of Economics*, 1, 355-84.
- Fehr, E. and U. Fischbacher. (2002). Why social preferences matter: the impact of non-selfish motives on competition, cooperation and incentives. *The economic journal*, 112, C1-C33.
- Fehr, E. and K. Schmidt. (1999). A theory of fairness, reciprocity and competition. *The quarterly journal of economics*, 114, 817-68.
- Frederick, S., G. Loewenstein and T. O'donoghue. (2002). Time discounting and time preference: a critical review. *Journal of economic literature*, 40, 351-401.
- Frey, B. (2012). Crowding out and crowding in of intrinsic preferences. In: *Reflexive governance for global public goods*, E. Brousseau, T. Dedeurwaerdere and B. Siebenhüner (eds).
- Fryer R., S. Levitt, J. List and S. Sadoff. (2012). Enhancing the efficacy of teacher incentives through loss aversion: a field experiment. *NBER working paper*.
- Fudenberg, D. and D. Levine. (2006). A dual-self model of impulse control. *American economic review*, 96, 1449-76.
- Genicot, G., and D. Ray. (2014). Aspirations and inequality. *NBER working paper*.
- Güth, W., R. Schmittberger and B. Schwarze. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior and organization*, 3, 367-88.
- Heidhues, P. and B. Koszegi. (2014). Regular prices and sales, *Theoretical economics*, 9(1) 217-251.
- Kahneman, D. (2011). *Thinking, fast and slow*. Allen Lane.
- Kahneman, D., J. Knetsch and R. Thaler. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of political economy*, 98, 1325-48.

- Kahneman, D., and A. Tversky. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47, 263-91.
- Koszegi, B. and M. Rabin. (2006). A model of reference-dependent preferences. *The quarterly journal of economics*, 121, 1133-65.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The quarterly journal of economics*, 112, 443-477.
- Madrian, B. and D. Shea. (2001). The power of suggestion: inertia in 401(k) participation and savings behavior, *The quarterly journal of economics*, 116, 1149-87.
- Mas, A. (2006). Pay, reference points, and police performance. *The quarterly journal of economics*, 121, 783-821.
- Meier, S., and C. Sprenger. (2010). Present-biased preferences and credit card borrowing. *American economic journal: applied economics*, 2, 193-210.
- OFT (2010) The impact of price frames on consumer decision making.
- OFT (2011) Consumer behavioural biases in competition: a survey.
- Ostrom, E. (2005). *Understanding institutional diversity*, Princeton University Press.
- Pope, D. and M. Schweitzer. (2011). Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *American economic review*, 101, 129-57.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American economic review*, 83, 1281-302.
- Rizzo, J. and R. Zeckhauser. (2003). Reference incomes, loss aversion, and physician behavior. *Review of economics and statistics*, 85, 909-22.
- Shalev, J. (2000). Loss aversion equilibrium. *International journal of game theory*, 29, 269-87.
- Thaler, R. H. (1988). Anomalies: the ultimatum game. *Journal of economic perspectives*, 2, 195-206.
- Thaler, R. and C. Sunstein. (2008). *Nudge*, Penguin.
- UK Cabinet Office (undated) Mindspace, available from <http://www.behaviouralinsights.co.uk/publications/mindspace> (accessed 19/02/15).
- Young, P. (2008). Social norms. *The New Palgrave Dictionary of Economics* (2nd ed), by S. Durlauf and L. Blume (eds). Palgrave Macmillan.

Author Details

Alex Dickson,
Department of Economics
University of Strathclyde
alex.dickson@strath.ac.uk

Marco Fongoni
Department of Economics
University of Strathclyde
marco.fongoni@strath.ac.uk