SECURITY AND HUMAN BEHAVIOUR 2010
Michelle Baddeley

SECURITY: FOUNDATIONS FROM BEHAVIOURAL ECONOMICS

Rough preliminary draft. For Discussion only.

1. INTRODUCTION

The growth of the Internet has been enabled by rapid technological developments. Specifically for electronic commerce, there is pressure on financial systems to adapt to the increased volume of spending taking place over the Internet. Until now, most buyers have used credit arrangements or checking accounts as the principle means of paying for Internet purchases. For Internet transactions, a range of electronic solutions including electronic cash (ecash) are proving to be potentially superior substitutes for conventional monetary instruments. Significant problems have emerged however because alongside the positive innovations, significant abuses have grown concomitantly including anti-social behaviour and security/privacy abuses, e.g. spam attacks, phishing and identity theft. In preventing fraud, the virtue of traditional cash is its anonymity, security and lower transactions costs so in amelorating problems of fraudulent collection of personal financial information during electronic transactions, at least theoretically and technically, systems such as e-cash can be developed to harness the lower computational and/or transactions costs of electronic payments schemes whilst retaining in electronic cash the virtues of conventional cash (e.g. in terms of security and anonymity).

Hypothetically at least, the potential security of virtual money could be greater than that of conventional money given the sophisticated printing and counterfeiting methods used for conventional cash. In practice however, governmental constraints have meant that adoption of tamper resistant technology is limited, e.g. for e-money by the US export limits on long /complex keys (Swire,1997). Many existing e-cash systems, particularly those that can be used with a number of different merchants, are not completely anonymous because the monitoring of their use is actually essential to the proper operation of these systems in order to prevent the double spending of virtual coins. This monitoring may be very costly requiring collusion between institutions. The use of a conventional cash system allows direct interaction between buyer and seller and so it is not possible to monitor transactions taking place mediated using conventional cash. Anonymity is ensured. Conventional cash will be preferred by those involved with criminal activities as long as criminals and tax evaders believe that electronic transactions will always leave some trace (Goodhart, 2000). It can be argued that complete anonymity is not desirable from a social welfare point-of-view (de Solages and Traore, 1998). In theory, a system of anonymity that is only revoked by some trusted authority when criminal activities take place would mean that criminal activity could be more effectively monitored and punished in a world of e-cash. But, in practice, the whole point is that criminals would not use a system that they believe allows effective monitoring and punishment. Even with such a system, until complete anonymity can be assured electronic cash cannot substitute completely for conventional cash for illicit transactions and there will always be a demand for conventional cash, whether or not agents admit their real reasons for holding it.

More generally, in developing secure systems and protecting privacy, key policy question include: To what extent should governments intervene to prevent these abuses? To what extent are individuals able to control for themselves the personal and financial information that they release to the world via email and the Internet? To inform our understanding specifically about what individuals can do to protect themselves in a computerised world this paper outlines a series of insights from economics in general and behavioural economics in particular.

2. ECONOMIC MODELS OF RATIONAL CHOICE

Whether or not people have the inclination and/or ability to protect themselves is a key issue to explore in discussions of security and human behaviour. Modern orthodox economics focuses on models of behaviour which assume that people are selfish and independent maximisers, driven by objective factors rather than more diffuse psychological and sociological forces. The simplifying assumptions in orthodox economics allow an analysis of decision-making which is clear and simple, though often lacking in realism and empirical validity. The approach is usually associated, in normative terms, with justifying free markets and eschewing government intervention, though the orthodox recognition of sources of market failure can deepen the analysis in some areas. Orthodox models were used as the basis of moves towards "economic imperialism" - most will have witnessed the burgeoning of popular economics books analysing a wide range of human behaviours (e.g. The Armchair Economist, Freakonomics etc. etc.) and these are based in Becker's analyses of rational choice as applied to a wide range of ordinary decisions (e.g. marriage, divorce, crime, addiction etc. etc.) - the essential assertion is that most human choice can be understood as a balancing of the marginal benefits and costs of choices (Becker 1993). In such an approach, if the restrictive assumptions are satisfied, the implication would be that individuals should be left to decide for themselves whether or not they need protection. But this approach assumes perfect markets and it is difficult to understand within such a stark approach the full range of issues relevant to security and human behaviour, nonetheless a few themes can be understood whilst retaining the assumptions associated with rational choice, once sources of market failure are recognised as explanations for ubiquity of imperfect competition in the real world. Key sources of market failure affecting security and privacy include network effects and externalities, public goods and price discrimination.

2.1 Externalities and Network effects

In understanding interactions between security and human behaviour, it is important to recognise that these systems are networked goods: their utility emerges from the fact that they are accessed within a network of other users. For networked goods, high fixed costs generate economies of scale; low marginal costs and lock-in suppress competitive pressures and sustain oligopolistic industrial structures; forces of imperfect competition are encouraged further by the other distinctive but related characteristics of networked goods including complementarities, externalities and switching costs (Katz & Shapiro, Shy 2001, Anderson and Moore). Complementarities emerge because networked products have little value in isolation (Katz and Shapiro, 1994; Economides, 1996). Furthemore, given heterogeneity of preferences and preference shifts, profits can be made from price discrimination and so there are commercial incentives to erode privacy in order to target different groups in different ways (Anderson and Moore 2008, 2009)

Consumers of electronic money products for example will be looking for a system that supports their electronic payments and so compatibility and operating standards including security are important. Network externalities emerge because the utility derived from consumption of networked goods increases with the number of agents consuming that good (Katz and Shapiro, 1985; Economides, 1996). The utility derived from the use of an electronic payment system for example is dependent upon the fact that other consumers are using the same system: if other consumers are not using the same payment system, then the value of the system will be reduced accordingly. In a dynamic context, this means that multiple equilibria can exist in which a producer will have all the potential consumers within the network - or none of them. In the example described above: PayPal is an example of a system which attracts many consumers just because other consumers are using it; DigiCash is an example of a system which attracted few consumers and so could not reach the critical mass required for it to survive. As electronic payment networks grow, the utility derived by each consumer will grow with the growing acceptability of the system. In theory at least, acceptability of a system should be affected by the efforts it makes to secure privacy. However, Bonneau and Preibusch (2009) analyse evidence about social networks which shows, that whilst the industry is vigorously competitive, privacy is not a selling point to the ordinary user even though it is a concern for the hawkish privacy experts. Thus the provision of privacy becomes a privacy communication game in which the privacy hawks are kept happy whilst privacy issues are hidden in order to maximise sign-up, generating a dysfunctional market for privacy.

Switching cost and lock-in may apply if existing a payment system is relatively more costly than entering it (Shapiro and Varian, 1999) which, as explained above, is a characteristic that applies to an extent to PayPal because it is easier to set up an account with PayPal than to get money out. Finally, economies of scale will mean that whilst there are high sunk or fixed costs involved in developing an electronic payments infrastructure, the marginal costs of copying and distributing electronic payment devices or tokens will be low. This generates a natural monopoly in which the average cost function declines sharply and limits the operation of competitive forces. These limits are likely to be more important for electronic payment system producers if the costs of developing new privacy and security infrastructure have to be born by private institutions.

2.2 Security as a Public good

Network externalities are also linked to the fact that security is a quasi public good. From consumers' point-of-view, if others in the network are adopting security controls disabling and therefore deterring a large volume of fraudulent activity, then there is no incentive for an individual to adopt those security controls themselves. When a network is already highly secure, then that security provision exhibits many of the key characteristics of a public good i.e. non-depletability, non-rivalry and non-excludability in consumption. This means that the provision of a good or service does not diminish because of consumption by an additional person; consumption by one person does not preclude consumption by others; and no one can be prevented from consuming the good. As a public good, a secure Internet is susceptible to the free-rider problem: consumers are able to free ride on the benefits without incurring any of the costs, generating a Prisoner's Dilemma type game (Anderson and Moore 2008, 2009). The value of access to additional users of the Internet is generally very small and so the costs involved in using credit/check payment systems are not easily justifiable. This gap cannot be easily filled by involving financial intermediaries because the financial transactions are too costly relative to the value gained from the exchange of information. The costs of using credit and checking payment mechanisms are not justifiable.

3. BOUNDED RATIONALITY

The security issues discussed above are analysed within a rational choice approach and these models by definition neglect socio-psychological forces affect security and human behaviour though the bridge between the two is bridged to an extent by approaches which recognise the constraints on rational choice in a world of risk, uncertainty and imperfect information. Most importantly and significantly Herbert Simon softened economists' conceptions of rationality by introducing models of bounded rationality (Simon 1955) and distinguishing substantive versus procedural rationality (Simon 1979, Baddeley 2006). Bounded rationality occurs when individuals' rationality is constrained by imperfect information, cognitive limitations, and time pressures. Substantive rationality occurs when sophisticated agents use mathematical algorithms to maximise their payoffs; procedural rationality is more likely to be associated with satisficing (i.e. sticking with the current situation because it's comfortable even if it's not an optimum) and involves blunter, broader approaches to information-processing.

In either the substantive or the procedural approach, some assumption or hypothesis must be formed to explain how people form their expectations about the future. Prediction is particularly complex when it comes to economic processes because the economic world is changeable: peoples' beliefs about economic structure have the capacity to change that economic structure, as emphasised in the mainstream literature on dynamic inconsistency (e.g. Kydland and Prescott, 1977) and the heterodox literature on non-ergodicity (e.g. Davidson, 1991). This suggests that Classical statistical or 'frequentist' approaches to the analysis of probability which assume repeatable events, complete information and/or an understanding of the data-generating mechanism, will be of little use in understanding the predictions of fixed asset investors for three reasons.

First, information is incomplete and the data-generating processes dictating economic outcomes are often unknown; an investment decision is not like dealing a card from a pack of 52 cards or buying a lottery ticket when you know that one million tickets are being sold. Secondly, many human decisions are about non-repeatable and unprecedented events and this means that information about past outcomes (e.g. as might be captured by frequency data) will be of little use. Thirdly, endogeneity means that economic realities are complex and mutable; expectations affect economic events that determine expectations (e.g. stock prices go up because people believe they will go up because stock prices are going up). Future outcomes will be affected by current decisions based on expectations of the future formed today: inter-temporal feedbacks between past, present and future will determine reality. Given these three sources of complexity, the objective basis for probability judgements may be missing or unknowable and the third source of complexity will undermine even more subjectively based Bayesian probability concepts.

3.1 Substantive Rationality and Imperfect Information

Simon defines substantive rationality as focusing on the achievement of objective goals given constraints (Simon, 1979, p. 67). If people are substantively rational, then they will form quantifiable expectations of the future and will make their decisions, e.g. about their security and privacy, using constrained optimisation techniques to balance the marginal benefits and costs. In other words they will use algorithms, e.g. entrepreneurs will use relatively sophisticated mathematical rules involving discounted cash-flow calculations incorporating assumptions about stable inter-temporal preferences. Algorithmic approaches assume that people using the same information set, will form identical expectations centred about some objective probability distribution of outcomes. They will be forward looking in incorporating a rate of time preference (i.e. discount rate) into their decisions and so will be optimising some objective function.

Risk, Uncertainty and Limits on Quantification

A significant problem for models of behaviour based on substantive rationality is in capturing how people deal with risk and uncertainty when making choices that have future consequences. For example, in using the Internet and in particular when using an online payments system or a social network, consumers must form an expectation of the likelihood of the information that they reveal will be used against them in some way in the future, e.g. via online fraud, being fired or ostracised for indulging in indiscrete online gossip, identity theft. In understanding how people form expectations, the basic distinction common to several frameworks of probability and uncertainty found in different academic disciplines is that between subjective versus objective probability. Subjective probabilities describe opinions or beliefs. Statistical literature makes the distinction between statistical probability and inductive probability (Carnap 1950; Bulmer 1979). A statistical probability is the limiting value of the relative frequency of an event over many trials. Statistical probability is therefore an empirical concept about some objective reality, and can be verified via observation and experiment (Bulmer 1979, p. 4). Statistical probabilities or frequencies are usually associated with some *ex post* calculation and/or a complete knowledge of a data-generating process; they may therefore have little to do with fundamental forms of uncertainty emerging from incomplete knowledge. Classical or frequentist statistical approaches have tended to assume implicitly that probabilities are statistical. By contrast, inductive probabilities describe rational expectations of a future event. They act as a guide to life and are formed even when an anticipated event is unprecedented; they therefore have no necessary association with frequency ratios. In contrast to statistical probabilities, inductive probabilities are to do with *ex ante* prediction; they are formed in the face of uncertainty and incomplete knowledge. In most areas of academic investigation, inductive probabilities are of greater practical importance than statistical probabilities because knowledge of an underlying objective reality is either limited or absent. With incomplete knowledge, statistical probabilities based upon past outcomes and an assumption of stationarity, are often inappropriate to the analysis of expert judgement in complex situations, either in natural scientific (such as in geoscience) or social scientific (such as in economics) contexts.

In analysing some of the limitations on quantification of economic probabilities, Keynes (1921) distinguishes between Knightian risk (the quantifiable risks associated with frequentist concepts) and Knightian uncertainty (which is unquantifiable). Under Knightian uncertainty people can say no more than that an event is probable or improbable; they cannot assign a number or ranking in their comparison of probabilities of different events. In the simplest terms the probabilities of Knightian risk and statistical/objective probabilities can be understood to be the same thing: Knightian risk events can easily be calculated using the frequency concepts associated with Classical statistical theory. These events tend to be repeatable and the outcome of a deterministic and immutable data generating mechanism, such as an unloaded die or a lottery machine. In a world of Knightian risk and quantifiable uncertainty it may be easy to assess and monitor expert judgement just by understanding the mathematics of the data generating process. Keynes (1921) argues, however, that in only a limited number of cases can probabilities be quantified in a cardinal sense; in some cases, ordinal comparisons of probability may be possible, but often, particularly in the context of unique events, probabilities may not be quantifiable at all. In reality there may be little consensus in expert (or amateur) opinions – particularly in economic decision-making. Keynes (1921) therefore argues that events characterised by Knightian uncertainty are more common than those characterised by Knightian risk, at least in the economic and social sphere.

Such issues are of particular importance in economics because much economic behaviour is forward looking, experiments may not be repeatable, and conditions cannot be controlled. People often make subjective probability judgements about events that have not occurred before, for which the data generating mechanism cannot be known. This makes the quantification and assessment of probabilities particularly problematic because it becomes impossible to match subjective probability judgements with an objective probability distribution. Also, endogeneity (i.e. the path a system takes is determined by events within the system) will limit the accuracy of probabilistic judgements of future events when beliefs about the future are affected by beliefs about the present. Shiller (2003) analyses such phenomena in the context of feedback theory, describing the endogeneity in belief formation: beliefs about the system determine the path of that system (e.g. stock prices go up because people believe they will go up). In this sort of world, no matter how much experts know there are no objective probability distributions waiting to be discovered; probabilistic judgements will always concern subjective beliefs rather than an immutable reality. If people are assumed to be consistent, rational and not prone to making systematic mistakes, then the distinction between conceptual probabilities and statistical probabilities disappears as uncertainty is reduced and as experts increase their knowledge of underlying data generating processes. However, experts can never be assumed to possess such qualities, as we show below.

Subjective beliefs are important in a world of conceptual uncertainty, and subjective probabilities can be analysed more effectively within a Bayesian approach than within a classical statistical approach. Bayesian analysis focuses on the subjective confidence that people have in a hypothesis about a single event and can be used to analyse the process by which subjective probabilities or judgements of confidence are updated as new information arrives. Subjectivity can be thought of as a negative quality, particularly in science. However, the formation of subjective judgements is not necessarily problematic if these subjective judgements are derived in a consistent way (Cox 1946). If any given set of information always generates the same set of probability judgements, then judgements can be said to have formed in a systematic way. The recognition of this insight has made the old subjectivist versus frequentist debates somewhat redundant as focus has shifted towards Bayesian methods, and thus the sting of the term ‘subjective’ has been drawn.

There are, however, a number of problems with the Bayesian approach. First, there are practical problems in its application, e.g. in economics, there is often a paucity of data that can be used to quantify subjectively formed probability judgements (Kennedy 1998, p. 205). Also, human intuitive cognitive processes do not deal well with Bayesian concepts. Anderson (1998) argues that this is a consequence of the nature of memes (the cultural analogy of genes – see below). Anderson suggests that Bayesian approaches can be refined using the advantages of a frequentist approach, e.g. using mental /visual imagery. For example, consistent methods should be developed: probabilistic information should be represented in graphical or pictorial form, and more generally frequentist approaches should be adopted in the presentation of information, attention should be paid to devices for cognitive interpretation, and Bayesian analysts should develop conventions for graphic displays. In other words, some frequentist methods can be used effectively within a Bayesian framework such that human cognition will process subjective probabilities more effectively.

The implications for security and human behaviour relate to legal issues, e.g. in insuring against the consequences of an internet attack for example, the basic principle would be that risks should be born by those who control the risk (Anderson and Moore 2008, 2009). But for decisions relating to internet use for example, the risks are interdependent, uncertain and to an extent unknowable; this profound uncertainty means that it is difficult to design efficient insurance. Imperfect information and misaligned incentives. This relates into the literature on adverse selection e.g. Akerlof’s (1971) lemons principle. Adverse selection is a pre-contractual problem of hidden information. This is relevant to security and human behaviour because people could select technical products to protect their privacy and security but as the technical sophistication of products increases, the ordinary consumer has far less information than the vendors about how effectively these products will work. Uncertainty may mean that even the vendor does not know how secure their software is (Anderson and Moore 2009). Whilst to an extent these problems might be overcome by learning, the search costs of investigating privacy products available are likely to be very high. One of the standard ways to overcome adverse selection problems is to devise a certification system but if the dodgy firms are the ones buying certification and/or if all firms are buying the easy certification then certification is unlikely to lead to efficiency gains (Anderson and Moore 2009).

Asymmetric information also leads to a post-contractual problem of moral hazard, i.e. hidden action, which occurs when incentives are misaligned: specifically, principal-agent problems lead to inefficiencies when the incentives of a principal and agent are different. For example, a firm providing security products aims to maximise profits and minimise costs; the consumer wants the best protection they can afford. Most consumers cannot monitor effectively whether or not their ISP or social network is doing what they promise to do. Principal-agent problems are also relevant for any area involving team effort. Security protection often depends on the efforts of many agents and the outcome may depend on either the minimum effort, best effort or aggregate effort (Varian 2004, ANderson and Moore 2009, Hirschleifer). For team work it is by definition difficult to identify who is responsible for output and this means that members of the team have an incentive to free-ride on the effort of others generating another Prisoner's Dilemma style game in which collective efforts are constrained. This could imply that minimum efforts are not dissimilar from best efforts and aggregate efforts; overall limited efforts will be made; the implication for security systems being that they become particularly vulnerable to attack.

3.2 Procedural Rationality and Cognitive Limitations

The above analyses do not abandon the assumption of maximising behaviour; they are just analysing behaviour in a world of imperfect information. In analysing procedural rationality Simon (1979, p. 68) responds to the mathematical approaches subsumed within the substantively rational approaches of rational behaviour by arguing that economic decisions are often the product of a 'procedurally rational' process. Procedurally rational behaviour is based on a broad reasoning process rather than the achievement of given representative agent's goals (Simon, 1979, p.68). The behaviour of the procedurally rational investor is guided by 'appropriate deliberation' and does not involve the optimisation of some objective function in the face of constraints. A procedurally rational person will use common sense rather than complex mathematical techniques in assessing their current and future choices. In contrast to the substantive approach, this implies that different people, even if they're using the same information will form different expectations reflecting arbitrarily assigned margins of error. It is not necessarily the case that these errors cancel out because mimetic heuristics (devices such as herding and following the crowd which were first hypothesised Keynes 1930, 1936, 1937 and later developed by Scharfstein and Stein 1990, amongst others) will form part of investment appraisal tool-kits. Whilst these heuristic devices are procedurally rational, they may nonetheless foster systematic mistakes and encourage path dependency.[1]

As explained above, for many decisions, people are operating in a world of bounded rationality in which the sensible application of clear and objective mathematical rules will be impossible because the existence of immeasurable Knightian *uncertainty* precludes the quantification of probabilities of future events. People will be forced to rely on appropriate deliberation by doing the best that they can, given the circumstances. They will use simple heuristics (or common-sense rule of thumb based on experience) in deciding whether or not to make particular choices. Limits on rationality are likely to be profound if the world is mutable and economic reality reflects endogenous processes. In this case, a consistent, immutable and objective reality may be missing; reality will be changing as expectations change. In such a world, subjective probabilities do not necessarily coincide with objective probability distributions, even on average so investors will be forced to adopt a procedurally rational approach. Without an objective path to follow, procedurally rational investors will use simple heuristics, e.g. they will look to others in deciding what to do, learning from the behaviour of

[1]The substantive approach of using algorithms and the procedural approach of using heuristics are not completely independent of each other. There is evidence from business decision-making that using simple heuristics may in some cases approximate the results that would emerge if a business had a more sophisticated algorithmic approach to investment appraisal (Baddeley 2006).

others. In a world of incomplete information it will be procedurally rational to follow the crowd, as explained in Topol (1991), and/or to learn from past output signals about what other investors are doing (Acemoglu, 1993). This sort of approach will lead to rationally justifiable herding, mimetic contagion and path dependency suggesting that any errors in expectations will not be random but instead may follow systematic trends.

If people are procedurally rational and the logical link between the objective and the subjective is broken, then a range of subjective probability judgements may be defensible. But if these turn out to be wrong, is it because businesses are misguided or is it because the economic reality changed unexpectedly? A large literature has developed analysing the first possibility: that cognitive limits on human information processing mean that individuals' subjective probability estimates are fallible (e.g. see Tversky and Kahneman 1982, Baddeley, Curtis and Wood 2004). If the second possibility holds true, will any predictive tool be unequivocally superior to all others? If complexity and endogeneity operate within limits, then the solution may lie with predictive tools that incorporate fuzzy logic methods, in which the binary concepts of 'true' and 'false' are replaced by *degrees* of truth.

3.3 Psychological factors

Analyses of real-world behaviour often reveal that people's decisions are driven by non-rational forces such as gut feel. The term non-rational implies here that information is not being used in any systematic way which does not necessarily imply that behaviour is stupid or misguided. The classic example is gut feel, a force that demonstrably drives entrepreneurs' decision-making (see Baddeley 2004 for survey evidence). Keynes's animal spirits - non-rational urges to act rather than remain idle - is a similar concept. Developing Keynes (1936), Akerlof and Shiller define animal spirits broadly, as the psychological factors affecting human behaviour. One of Akerlof and Shiller's animal spirits that are central to the issue of financial security is corruption: they argue that financial instability is exacerbated by the corruption that grows during boom phases – throughout history, corruption has accompanied the business cycle –from the Prohibition and the Great Depression of the 1930s, to the Savings and Loan crisis associated with the 1991 recession, to the Enron scandal associated with the 2001 recession and to the sub-prime mortgage crisis which precipitated today's recession.

Many of these non-rational forces are caught up with socio-psychological motivations and whilst conceptually are woolly and therefore difficult to analyse, there is increasing evidence that they are relevant (e.g. Loewenstein on animal spirits). A conceptual question could be whether or not behaviour is the outcome of irrational mistakes or procedurally rational time/cost-saving devices. Economists have traditionally been preoccupied by such distinctions between the rational and irrational but there is increasing recognition, particularly by some neuroeconomists, that this dichotomous approach is spurious. Neuroeconomics has a lot to offer in increasing our understanding of the neurological foundations of reward processing (e.g. see Schultz). It also escapes specious distinctions between rational, irrational and non-rational behaviour and enhances our understanding of evolutionary processes / proximate mechanisms, e.g. those that lead to procrastination as discussed below.

Are biases and heuristics procedurally rational but blunt decision-making tools i.e using information in a very rough way to cut costs and save time? Whether or not these are psychological biases rather than procedurally rational decision-making tools, is to a large extent a semantic debate that is not relevant here and instead we will outline some psychological and sociological factors

3.3.1 Cognitive Bias and Heuristics

Research has shown that most ordinary people make common mistakes in their judgements of probabilities (e.g. Anderson 1998) generating individual and group biases (Baddeley, Curtis and Wood 2005).[2] This links into bounded rationality because it reflect cognitive limitations in the processing ability of the human mind (Gould 1970; Tversky & Kahneman 1974; Anderson 1998). The problem originates in the input format of data, and in algorithms used: if prompted by clear signals, the human brain is able to deal with probabilities effectively (Anderson 1998). For example, if students are asked to judge the probability of two coincident events within the context of a statistics class, then they will know what to do. However, if outside their classes they are confronted with a problem requiring probability judgements in a situation in which it is not obvious that this is what is required, then they may make a judgement using instincts and intuition rather than statistical reasoning (see Kyburg 1997). The key sources of inconsistency emerge from either individual bias or group bias.

At least two main types of individual bias can be distinguished: motivational bias and cognitive bias (Skinner 1999). Motivational biases reflect the interests and circumstances of the expert (e.g., does his or her job depend on this assessment? If so, s/he may be overconfident in order to appear knowledgeable). Motivational biases such as these can often be significantly reduced or entirely overcome by explaining that an honest assessment is required, not a promise. Also, it may be possible to construct incentive structures encouraging honest assessments of information. Motivational biases can be manipulated because they are often under rational control. Cognitive biases are more problematic because they emerge from incorrect processing of the information; in this sense they are not under conscious control. Cognitive biases are typically the result of using heuristics, the common-sense devices or rules of thumb derived from experience, used by people to make relatively quick decisions in uncertain situations. They are used because a full assessment of available information is difficult and/or time consuming or when information is sparse. For example, when thinking about buying/selling shares from their portfolio, a potential investor may have little real knowledge about what is going to happen to share prices in the future; given this limited information, they will adopt the heuristic of following the crowd, i.e. buying when the market is rising and selling when it is falling.

At least four types of heuristics that produce cognitive bias are commonly employed: availability, anchoring and adjustment, representativeness, and control (Kahneman et al. 1973; Tversky & Kahneman 1974). Availability is the heuristic of assessing an event's probability by the ease with which occurrences of the event are brought to mind. This often works quite well, but can be biased by the prominence of certain events rather than representing their frequency. For example, headline news of airplane crashes will be brought to mind more readily than bike crashes, even though the latter are far more frequent. For security and human behaviour, the

[2] See also Thaler and Sunstein (2008) for an accessible introduction to this literature.

availability heuristic combined with an overoptimism bias may lead people to decide that security is not a problem because they haven't recently had a problem with it.

Anchoring and adjustment is a single heuristic that involves making an initial estimate of a probability called an anchor, and then revising or adjusting it up or down in the light of new information (Tversky & Kahneman 1974). This typically results in assessments that are biased towards the anchor value. For example, in deciding about an appropriate wage demand to make in the face of an uncertain economic environment, workers will tend to anchor their demands around their existing wage.

The control heuristic is the tendency of people to act as though they can influence a situation over which they have no control. People value lottery tickets on which they have chosen the numbers more highly than those with random number selection, even though the probability of a win is identical in both cases. The representativeness heuristic is where people use the similarity between two events to estimate the probability of one from the other (Tversky and Kahneman 1982). The classic example is the "Linda problem". In experiments, a large proportion of people will judge it to be more likely that Linda is a social worker active in the feminist movement, than that she is just a social worker even though the former is a subset of the latter. If this problem were to be expressed in probabilistic / statistical terms, anyone with a basic knowledge of probability would realise that two events happening together is less probable than each event happening irrespective of whether the other occurred. However, when confronted with the details about Linda, most people find the first option more likely than the second, simply because the description appears to be more representative of a feminist stereotype. This is a conjunction fallacy. In the same way that the probability of events compounded using logical AND is often overestimated, the probability of events compounded using logical OR is often underestimated (Bar-Hillel 1973). Such biases can also create an unbounded probability problem: subjects tend to over-estimate each probability in a set of exhaustive and mutually exclusive scenarios, so that the estimated sum of all probabilities is greater than one (Anderson, 1998, p.15). Also, people do not correct their probability estimates when the set of exhaustive but mutually exclusive outcomes is augmented, again leading to an estimate of total probability in excess of one.

Other well-known biases introduced by the representativeness heuristic include the gambler's fallacy and base-rate neglect. The gambler's fallacy is the belief that when a series of trials have all had the same outcome then the opposite outcome is more likely to occur in the following trials, since random fluctuations seem more representative of the sample space. Base-rate neglect is neglect of the relative frequency with which events occur. Consider the following example: The group was told that Dick came from a population of 30 engineers and 70 lawyers: Dick is a 30-year-old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues. This description provides no information about Dick's profession, but when subjects were asked to estimate the probability of Dick being an engineer, the median probability estimate was 50%, whereas the correct answer is 30%. Subjects ignored the base rate and judged the description as equally representative of an engineer and a lawyer (Tversky & Kahneman 1974, p. 1126).

An interesting and commonly used combination of the gambler's fallacy and base rate neglect is called probability matching, a heuristic known to be used by humans and some other primates (e.g. Bliss et al. 1995). This is where a reaction from a given range is chosen in proportion to the probabilities of occurrence of various consequences. An example given by Lo (2001) was from World War Two. Bomber pilots were allowed to carry either a flak jacket or a parachute, but not both because of the extra weight. They knew that their probability of getting strafed by enemy guns (requiring a flak jacket for protection from shrapnel) was three times that of being shot down (requiring a parachute). Pilots were observed to take flak jackets three times out of every four and parachutes on the fourth occasions. This is not an optimal assessment of the probabilities. Pilots were more likely to have survived if they had taken a flak jacket 100% of the time because the probability of getting strafed by enemy guns was always more likely than the probability of being shot down – the flak jacket was always more likely to be of use. Probability matching might occur in a geological context if a human was asked to estimate the type of geology that is most likely at a set of sub-surface locations in a reservoir, knowing that wells would be drilled at those locations and their estimates checked. If all they knew about the geology in the reservoir was that it was either of a type 1 or a type 2, and type 1 was three times as likely to occur as type 2, then it is possible that they would posit geology type 1 as the most likely on average three times out of every 4. Although this would be a non-optimal prediction, this would be a natural tendency for any human, including an expert, who is not intimately familiar with basic probability theory.

Other cognitive biases may reflect emotional responses. Framing effects are about how people's responses will be determined by the way / context in which questions or problems are framed. People are often overconfident about their knowledge. Overconfidence is especially a problem for extreme probabilities (close to 0% and 100%) which people tend to find hard to assess, and in choices about security and privacy may lead to people being overly sanguine about the chances of identity theft for example. Other forms of emotional response affecting the heuristics employed include mood: people in a happy mood are more likely to use heuristics associated with topdown processing, i.e. relying on pre-existing knowledge with little attention to precise details. By contrast, people in a sad mood are more likely to use bottom-up processing heuristics, paying more attention to precise details than existing knowledge (Shwarz 2000, p. 434). Minsky (1997, p. 519) analyses some of the emotional constraints in the case of expert knowledge, arguing that the 'negative knowledge' associated with some emotional states may inhibit whole strategies of expert thought. Of all of the biases described above, the most prevalent may be overconfidence and base-rate neglect (Baecher 1988).

Many other cognitive biases have been identified too including status quo bias, attribution error, endowment effects and loss aversion. Some of these biases can be manipulated to encourage people to engage in more efficient behaviour - for example the status quo bias which is about the fact that people tend to favour the existing situation and will tend to avoid the effort involved in changing their choices. Setting default options cleverly can exploit this bias e.g. if the default option is the maximum privacy protection then a large number of consumers may be too lazy to change these options thus protecting them from security violations.

3.3.2 Psychological factors: present bias and time inconsistency

Another cognitive bias that deserves particular attention is the present bias. People's behavior may be inconsistent over time: plans to do something constructive (e.g. giving up smoking) in the future change as the future becomes the present, people lack self control. This can be captured theoretically by a small tweek to the standard orthodox economic assumptions about exponential discounting: by introducing a present bias parameter into standard discount functions, preference reversals and time inconsistency are the outcome. There is a wide literature demonstrating the relevance of present bias to a wide range of microeconomic and macroeconomic behaviours (Laibson 1997; Angeletos *et al.* 2001, Frederick, Loewenstein and O'Donoghue 2002). Present bias may not be irrational i.e. may reflect a procedurally rational approach i.e. if people are treating different financial decisions in different ways i.e. using different 'mental accounts'. Experimental evidence shows that people, experiencing a windfall gain of $2,400, will save different proportions depending on the circumstance of the windfall and the context in which the windfall is received: they spend $1,200 if the windfall is spread over a series of monthly payments, £785 if it's a single lump sum and nothing if is an inheritance. Rather than treating economic decisions together as one, gigantic maximization problem, people assign different events to separate mental accounts (Thaler 1999, Akerlof and Shiller 2009).

For security and human behaviour, Acquisti (2004) and Acquisti and Grossklags (2006) have analysed the implications of present bias for people's choices about privacy and security building on the behavioural economics literature on procrastination and self control (O'Donoghue and Rabin 1999, 2001; Dellavigna and Malmendier 2006). When using the internet people will procrastinate about setting up effective security systems in much the same way as many ordinary people procrastinate about backing-up files. Procrastination is potentially a key policy issue particularly if the most effective privacy and security solutions are to be driven by individual choices. Assuming that people suffer present bias but are sophisticated enough to realise that this might generate security / privacy problems in the future, then they can be encouraged to set-up pre-commitment devices which will force them to take a little more effort to protect themselves from security violations in the short-term i, e.g. identity verification systems.

3.4 Sociological Forces

Orthodox economics assumes that people act as atomistic agents. In reality, many of our decisions are subjective to social influence - whether normative (e.g. peer pressure) or informational (e.g. learning from the actions of others). These group interactions generate more complex forms of bias as people interact and copy each other thus spreading misjudgements quickly through groups of people. Anchoring effects may operate in a social dimension too if one individual's judgements is 'anchored' to another's (Tversky & Kahneman 1974; Eichenberger 2001). An literature assessing various possibilities for such thought contagion has developed in economics, beginning with Keynes's analysis of uncertainty, rationality, subjective probabilities, herd behaviour and conventions (Keynes 1921, 1936, 1937). In Keynes's analysis, herding behaviours are linked back into an analysis of probabilistic judgement in a Bayesian setting. Differences in posterior judgements of probable outcomes may not reflect irrationality but instead may emerge as a result of differences in prior information. Rational economic agents may have an incentive to follow the crowd and herding will result as a response to individuals' perceptions of their own ignorance. This herding will be rational if an individual has reason to believe that other agents' judgements are based upon better information than their own: other people's judgements become a data-set in themselves. In this way, people will incorporate others' opinions into their prior information set and their posterior judgements may exhibit herding tendencies. Shiller (2000, 2003) analyses these ideas in the context of feedback theories of endogenous opinion formation in which beliefs about the system determine the path of that system, e.g. as is seen in stock markets. These ideas are also developed in Topol (1991), Schleifer (2000), Brunnermeier (2001) and Sornette (2003), amongst others.

However, whilst herding behaviour can be explained as a rational phenomenon, the existence of herding may still contribute to instability if the herd is led down the wrong path. Stable outcomes will only be achieved if the herd can be led along a path of increasing the stock of common (real) knowledge. In such cases, increases in the stock of reliable prior information will contribute to convergence in posterior probabilities. If, however, the herd path fosters increasing noise within the system then the process of opinion formation will become unstable. Path dependency in the evolution of scientific beliefs can be described using evolutionary biological analogies, e.g. those based around the concept of a meme, the cultural equivalent of a gene (Dawkins 1976). Imitation is a distinguishing characteristic of human behaviour and a meme is a unit of imitation (Blackmore 1999). These ideas are related to analyses of theory of mind. The discovery of 'mirror neurons' (neurons in the pre-motor areas of primate brains that are activated without conscious control and generate imitative behaviour in primates) has lent some scientific support to these biological explanations for imitative behaviour (Rizzolatti et al. 2002). This biological approach is also consistent with the use of neural networks for information processing: i.e. mathematical approaches that emulate adaptive learning processes observed in human brains. Biological insights can be applied in the analysis of belief formation in a human context. Anderson (1998) asserts that successful memes survive (that are remembered) and reproduce (are transmitted) effectively when they a) map effectively onto human cognitive structures, b) incorporate a standardised decision structure, and c) have been reinforced by dominant members of the scientific community. Lynch (1996, 2003) applies these insights in his analysis of the evolutionary replication of ideas and argues that 'thought contagion' affects a wide range of human behaviours and beliefs.

The implications for security and human behaviour are that if group leaders can be identified and encouraged to adopt appropriate protection then others will follow their example. This links into issues of social capital, social norms and social networks. In understanding these issues the role of trust and reputation is crucial. For security and human behaviour, Clark (2010) argues decisions are made in a multidimensional space and reflect contradictory goals and so trust and control are central; effective security and privacy systems will allow transparent communication between trusted parties but will be closed to the "bad guys". Another relevant factor in this context is social norms; e.g. norms affecting privacy and security are changing; for example, it is widely believed that the younger generation is more vulnerable to identity theft because they are far more willing to reveal important personal information. In terms of policy implications, social norms can be manipulated in various ways including advertising, sanctions and rewards.

There is much more that can be said about social influences, reciprocity, inequity aversion, reputation and cooperation. Lessons can be learnt from other applications of behavioural economics, e.g. to manipulating household energy consumption (see Shulz et al. 2007). There are many other topics of interest too: for example, Personality and individual differences: personality will be an important variable affecting how individuals protect themselves; Anderson and Moore (2009) draw on Baron-Cohen's distinction of systematisers from empathizers arguing that systems must be designed to suit both.

4. CONCLUSIONS AND POLICY IMPLICATIONS

In designing effective policies efficiently to protect privacy and enhance security a key policy debate is the relative roles to be played by government regulation versus private initiative. Since 911, geopolitical factors have necessitated a cautious approach to the development of systems which enable the cheap and anonymous electronic movement of money. For phishing attacks, the marginal costs are very low for the perpetrators and the chances of being caught are slim, a significant problem will be formulating strategy proof designs given the very small costs faced by perpetrators. Is it ever going to be possible to manipulate their incentives to prevent spam and phishing? Fines and penalties might be more effective but, for both phishing and online fraud therefore, the capacity for governments effectively to police these violations is limited. So effective solutions will necessarily have to concentrate on encouraging people to take a more responsible attitude towards protecting their privacy. Sophisticates who are well-informed about the dangers of identity theft etc. can be encouraged to use pre-commitment devices much as sophisticates who realise the long-term benefits

REFERENCES

1. Akerlof G, Shiller R (2009), Animal Spirits: How Human Psychology Drives the Economy and Why it Matters for Global Capitalism', Princeton: Princeton University Press.
2. Angeletos G-M, Laibson D, Repetto A, Tobacman J and Weinberg S (2001) The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation, Journal of Economic Perspectives, Vol. 15(3), pp. 47-68.
3. Acquisti A (2004), Privacy in Electronic Commerce and the Economics of Immediate Gratification, EC'04.
4. Acquisti A and Grossklags J (2006) What can behavioral economics teach us about privacy, ETRICS 2006.
5. Anderson R, Moore T (2009) Information security: where computer science, economics and psychology meet, Phil. Trans. R. Soc A, 367, 2717-2727.
6. Anderson R, Moore T (2008) Information Security Economics – and Beyond, Information Security Summit 2008.
7. Baddeley M (2010) Herding, Social Influence and Economic Decision-Making: Socio-Psychological and Neuroscientific Analyses, Phil. Trans. R. Soc B, 365, pp 281-290.
8. Baddeley M (2006) Behind the Black Box: a survey of real-world investment appraisal approaches, Empirica, vol. 33, no. 5, pp. 329-350.
9. Baddeley M (2004) Using e-cash in the New Economy: An economic analysis of micropayments systems', Journal of Electronic Commerce Research, 2004, 5(4), pp. 239-53, http://www.csulb.edu/web/journals/jecr/issues/20044/Paper3.pdf.
10. Baddeley M, Burke C, Schultz W, Tobler P (2010) Impacts of personality on herding in financial decision-making, Cambridge Working Papers in Economics CWPE No. 1006, Faculty of Economics, Cambridge January 2010.
11. Baddeley M, Curtis A, Wood R (2005) An introduction to prior information derived from probabilistic judgments; elicitation of knowledge, cognitive bias and herding, Geological Prior Information: Informing Science and Engineering, Royal Geological Society, London, Special Publications No. 239, pp. 15-27.
12. Baddeley M, Fontana G (2006) Monetary Policy in the Information Economy: Old Problems and New Challenges, 2006, in P. Arestis, M. Baddeley & J. McCombie (eds) (2006) 'New' Monetary Policy: Implications & Relevance, Cheltenham: Edward Elgar.
13. Becker, Gary (1993), `The Economic Way of Looking at Behavior', Journal of Political Economy, vol. 101, no. 3, pp. 385-409.
14. Bernheim BD, Rangel A (2004): "Addiction and Cue-Triggered Decision Processes", The American Economic Review, Vol. 94, No. 5, pp. 1558-1590
15. Bonneau J and Preibusch S (2009) The Privacy Jungle: On the market for Data Protection in Social Networks, WEIS 2009.
16. Burke C, Tobler P, Schultz W, Baddeley M (2010) Striatal BOLD response reflects the impact of herd information on financial decisions. With C Burke, P Tobler and W Schultz. Forthcoming in Frontiers - Neuroscience.
17. Camerer C, Loewenstein G and Prelec D (2005), `Neuroeconomics: How Neuroscience Can Inform Economics", Journal of Economic Literature, vol. 43, no. 1, pp. 9-64.

18. Cialdini, R. (2007). Influence: The psychology of persuasion. New York: Collins Business.
19. Clark (2010), A social embedding of network security: trust, constraint, power and control, Security and Human Behaviour 2010.
20. DellaVigna S and Malmendier U (2006), Paying Not to Go to the Gym, American Economic Review, vol. 96(3), pp. 694-719.
21. Fehr E. and Gächter, S. (2000), Cooperation and Punishment in Public Goods Experiments, American Economic Review, vol. 90(4), pp. 980-94.
22. Frederick S, Lowenstein G and O'Donoghue T (2002) 'Time Discounting: A Critical Review', Journal of Economic Literature, vol. 40(2), pp. 351-401.
23. Kahneman, D. (2003) 'Maps of bounded rationality: Psychology for behavioral economics', American Economic Review, vol. 93, no. 5, pp. 1449-1475.
24. Kahneman D and Tversky A (1979), 'Prospect Theory -An Analysis of Decision under Risk'', Econometrica, vol. 47, no. 2, pp. 263-292.
25. Loomes, G and Sugden R (1982): "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty", The Economic Journal, Vol. 92, No. 368, pp. 805-824.
26. Thaler, RH (1999) 'Mental Accounting Matters', Journal of Behavioral Decision Making, vol.12, no. 3, pp. 183-206. Reprinted in Camerer et al. Advances in Behavioral Economics (2004), Ch. 3.
27. Laibson, D., 1997. Golden eggs and hyperbolic discounting. Quarterly Journal of Economics 112,443–477.
28. McClure SM, Laibson DI, Loewenstein G and Cohen JD (2004). ' Separate Neural Systems Value Immediate and Delayed Rewards', Science, vol. 306, pp. 503-7.
1. O'Donoghue and Rabin (2001) Choice and Procrastination, QJE
2. O'Donoghue, Ted and Rabin, Matthew (1999): "Doing It Now or Later", American Economic Review, 1999, 89(1), pp. 103—24.
3. Schultz, W.: Behavioral theories and the neurophysiology of reward. Ann. Rev. Psychol. 57: 87-115, 2006.
4. Schulz, P., J. Nolan, R. Cialdini, N. Goldstein and V. Griskevicius (2007). The constructive, destructive and reconstructive power of social norms. Psychological Science 18: 429-34.
5. Simon H 1979, 'From Substantive to Procedural Rationality', in. Hahn, Frank. H and Martin Hollis (eds), Philosophy and Economic Theory, pp. 65-86.
6. Simon, H (1955) 'A behavioural model of rational choice', Quarterly Journal of Economics, vol. 69, pp. 99-118.
7. Thaler (1999) Mental accounting matters
8. Thaler, R.H. and Shefrin H.M. (1981) 'An Economic Theory of Self-Control', Journal of Political Economy, vol. 89(2), pp.392-406.
9. Thaler and Sunstein (2008), *Nudge*.
10. Tversky, A & Kahneman, D (1986) 'Rational Choice and the Framing of Decisions', Journal of Business, vol. 59, no. 4, pp. 5251-78.