Decision-making environments in which unboundedly rational decision makers choose to ignore relevant information

Nathan Berg

School of Social Sciences,
University of Texas at Dallas, GR 31 211300,
Box 830688, Richardson, Texas 75083 0688, USA
Fax: 1 972 883 2735 E-mail: nberg@utdallas.edu

Center for Adaptive Behavior and Cognition,
Max Planck Institute for Human Development,
Lentzeallee 94, 14195 Berlin

Abstract: This paper advances the claim that ignoring relevant information is sometimes consistent with good decision making. Although that finding is not new, the argument presented here is. In contrast with bounded rationality models, the decision-making model in this paper presupposes no cognitive constraints or costs associated with processing available information. The paper identifies a class of decision-making environments characterised by asymmetric payoffs and probabilities – a property which gives rise to optimal decision rules that ignore relevant information. In other words, optimal decision procedures used by omniscient agents are sometimes independent of variables that objectively predict future outcomes.

Keywords: ignoring; information; bounded rationality; optimising; adaptive; behavioural; decision.

Reference to this paper should be made as follows: Berg, N. (2005) 'Decision-making environments in which unboundedly rational decision makers choose to ignore relevant information', *Global Business and Economics Review*, Vol. 7, No. 1, pp.59–73.

Biographical notes: Nathan Berg is an Economist at the University of Texas at Dallas and Guest Researcher at Max Plank Institute's Adaptive Behaviour and Cognition Group in Berlin. Berg's work in behavioural economics has shown that systematically incorrect beliefs sometimes improve the market performance. His work has also demonstrated that peer comparisons by financial market participants can encourage risk taking and foster economic growth. Berg's writings on normative behavioural economics have underscored the crucial role of modelling assumptions and their empirical validity in the analysis of economic policy. Currently, Berg is studying *process* models that seek to go beyond the as-if task of matching statistical pattern of decision outcomes, and instead describe the cognitive structure of the decision process itself.

Copyright © 2005 Inderscience Enterprises Ltd.

1 Introduction

This paper seeks to advance the claim that ignoring relevant information is consistent with good decision making. Although that claim is not new, the argument presented here is, because it presupposes no cognitive constraints, i.e., no bounded rationality assumptions. The goal is to provide a general representation of decision making environments in terms of payoffs and information, and provide a characterisation of those environments in which optimal actions ignore, or are independent of, variables that are objectively useful for predicting future outcomes.

It seems paradoxical at first. Consider a stochastic outcome that gives rise to random payoffs. The decision-maker can influence the probability distribution of random payoffs by selecting actions from a set of available choices. Influencing the decision is an observed vector of cues that help predict the stochastic outcome. Cues are not redundant (i.e., co-linear) and have non-zero correlation with future payoffs. Therefore, the expected payoff, conditional on observed cues, is a nontrivial function of those cues. Moreover, the decision-maker faces no decision costs or limitations on the ability to optimise – computation is free. And yet, when the environment features payoffs and probabilities that roughly offset each other, actions that maximise the expected payoff function are independent of at least one cue.

In other words, if the function that maps relevant cues into the payoff-maximising strategy is independent of one cue, that means it is fully rational to ignore the cue, even though it is correlated and thus help predicts future payoffs. This paper shows that the scenario described above implies no paradox. The optimality of ignoring information is a straightforward consequence of asymmetric (e.g., J shaped) probability distributions and payoff functions.

This paper defines the decision-making *environment* as consisting of a joint probability distribution over states of nature and cues, and a payoff function that ranks stochastic outcomes conditional on observed cues and actions. Actions influence payoffs both directly, because of different costs and benefits associated with choosing different actions, and indirectly through the dependence of outcome-cue probabilities on actions. For a given environment, an action or decision rule is defined as *optimal* if it maximises the expected payoff function conditional on observable information.

In general, optimal actions depend on observable cues, because the expected payoff function inherits dependence on those variables through the conditional distribution function. Given choice variable a and a function $f(a, x)$ that depends on actions and exogenous variables (observed cues x), the maximiser of f depends on x: $\text{argmax}_a\ f(a, x) = a^*(x)$. However, for particular payoff and information structures, optimal actions assign zero weight to cues that are in fact relevant in the sense that the conditional distribution of payoffs depends on those cues. In other words, the focus of this paper is environments giving rise to expected payoff functions f that non-trivially depend on x while having maximisers a^* that do not depend on (some components of) x.

This paper describes a general condition, both necessary and sufficient, specifying precisely the environments which give rise to optimal actions that assign zero weight to a relevant cue. Within the class of such environments, a special subclass is identified that possesses a recognisable form which is consistent with previous experimental findings in psychology and decision making theory. Theoretical results reported in this paper help to explain the pervasive experimental finding that human subjects frequently make incomplete use of available information. Thus, the experimental record on incomplete information processing is consistent with the hypothesis that adaptive responses to particular types of payoff/information environments account for information processing anomalies known to be characteristic of human cognition.

The claims in this paper do not deny the importance of bounded rationality as an explanation for ignoring in the real world. In fact the presence of bounded rationality strengthens the paper's results, suggesting that it is the interaction of bounded rationality and the payoff-probability structure of the environment that jointly accounts for the pervasiveness of information-processing irregularities. This paper abstracts from bounded rationality in order to isolate the connection between decision-making environments and ignoring. The advantage of this strategy is that it establishes probability-payoff structure as an independent cause of ignoring. As discussed in later sections, the probability-payoff channel is strengthened by the presence of bounded rationality, although it does not depend on it to generate optimal decision rules independent of relevant information.

2 Background

Awareness of the fundamental importance of information reduction and the cognitive processes underlying signal extraction was evident as early as the 1960s in the work of Gerard (1960), Walker and Bourne (1961) and Posner (1964). Much of the intuition for the phenomenon of ignoring information be found in Stroop (1935) and, some would claim (Mullainathan, 2002), in even earlier work, such as James' (1890) implicitly probabilistic interpretation of memory. Later studies of information reduction and signal extraction led to theories of automaticity (Hasher and Zacks, 1979; Kahneman and Treisman, 1984; Logan, 1980) implying that human perception is impervious to certain stimuli for good reason – so that cognitive resources can be allocated to other more important tasks. Recent evidence (Dishon-Berkovits and Algom, 2000; Hutchison, 2002; Godijn and Theeuwes, 2002) suggests that the presence of automaticity, perceptive imperviousness and Stroop effects are governed by complex dynamics which are partially manipulable, exhibiting systematic responses to a variety of experimental stimuli. These results motivate this attempt to model optimal cognitive processing in terms of payoffs and frequencies given exogenously as features of the decision-making environment.

A relatively large body of research has demonstrated how cognitive constraints, whether they derive from memory limitations, upper bounds on processing speeds, or a fixed number of channels for allocating attention, lead to cognitive strategies that economise processing resources by suppressing or ignoring particular categories of stimuli. For example Ruthruff et al. (2001) test models of activation suppression; part of theory that suppressing recently used information can facilitate readiness (i.e., improved performance) in attacking newly encountered tasks. Shiu and Pashler (1992) show how certain types of inattention can improve performance on perceptual tasks. Reingold et al. (2001) find some evidence of this phenomenon among expert chess players. Recent work (Schubo et al., 2001; Lippa and Goldstone, 2001; Most et al., 2001; Wentura, 2002) also shows that so-called pre-attentive features of information processing respond systematically to environmental variables, making the idea that cognition responds to payoff-probability structure given by decision making environments even more plausible. It is precisely this interplay between environment and decision making that motivates this paper's attempt to formalise the relationship between payoff-probability structure and the phenomenon of non-responsiveness or information suppression.

While the empirical record unambiguously testifies to the existence of the phenomenon of ignoring information, a variety of theories are available, making it difficult to point to a single mechanism. Schiller and Caramazza (2002) argue that there are distinct categories of information, each with correspondingly distinct modes of information processing. Rather than a homogeneous substance that can be quantified in standardised units, the theory holds that information is mediated through language, and is therefore parsed into subcategories which call forth appropriate cognitive responses. The implication is that mapping the taxonomy of subcategories is the key to predicting which cognitive response will likely prevail. Rodriguez-Fornells et al. (2002) similarly focus on language processing, demonstrating that distinct neural architectures support different kinds of information suppression, which correspond to different tasks.

In addition to cognitive research focussed on language, theories of brain and neural architecture are also implicated in information suppression, e.g., the motion perception study of McDermott et al. (2001). Some authors argue based on psycho-social theories that information suppression arises as a coping mechanism, permitting individuals to continue functioning without fully processing potentially stressful or traumatic information (Crozier and Skliopidou, 2002; Neuman, 2002). In the domain of conscious rather than automatic information suppression, Erber and Fiske (1984) postulate that the decision makers' overriding objective is self consistency; to eliminate dissonance, attention may focus explicitly on inconsistent facts in order to assimilate them, or may simply turn away. Erber and Fiske claim that environments with large-magnitude payoffs (or costs) should induce individuals to focus directly on dissonance rather than suppress it. However, the model in this paper implies no particular connection between payoff magnitude and the likelihood of ignoring information.

Understanding the mechanism underlying information suppression potentially illuminates decision-making problems connected to policy issues such as pension financing, lifestyle choices that lead to large public expenditures on healthcare, and immigration/migration decision (Mullainathan, 2002; Dow, 1991; Cutler et al., 1989). Conlisk (1996) surveys the bounded rationality literature within economics, describing how problem-solving costs and upper bounds on processing time have been incorporated into the neoclassical optimisation framework. Mullainathan (2002) acknowledges that the economic approach to memory does not deal at all with the possibility of beneficial or adaptive information suppression.

Within economics, bounded rationality is virtually always viewed as a limitation, standing in the way of full optimisation rather that as an adaptive tool that facilitates enhanced performance. Computer scientists Schein et al. (2002) Emran and Ye (2001) and decision theorists Jones and Brown (2002) Ferri et al. (2001) report findings from applications of models of memory that overlap with the basic idea of this paper, namely, that ignoring can be good. In other specialised applications such as listening to music (Bigand et al., 2000) and chess playing (Reingold et al., 2001), there is a clear association between ignoring information and high levels of performance, implying that models of ignoring may prove to be valuable prescriptive as well as descriptive tools.

3 A general decision problem

Let Y represent a continuous random variable on the set $\mathcal{Y}$ corresponding to an uncertain outcome of interest to the decision maker.[1] The vector of cues X represents observable environmental factors that are used to form an expectation of Y. All components of X are correlated with Y and imperfectly correlated with each other to ensure against perfect co-linearity, i.e., redundancy among the predictors. The variables Y and X, together with the decision variable, action a, which takes on values over the action space A, determine the decision maker's payoff function $\pi(Y, X, a)$.

It must be assumed that the payoff function is twice differentiable in all its arguments, but nothing is assumed about its shape (the signs of the first and second derivatives). Although $\pi(Y, X, a)$ could be interpreted as a utility function representing a preference ordering over all outcomes in the space of (Y, X, a), it is simplest to think of it as an objective fitness function given by the decision-making environment (Lipman, 1991)[2].

The individual's decision-making *environment* is defined as a pair of functions

$$\{f_{Y|X}(Y, X; a), \pi(Y, X, a)\}, \tag{1}$$

where $f_{Y|X}$ is the probability distribution function (pdf) of Y conditional on X, and, π is the scalar valued payoff function defined above. The function π summarises the costs and benefits associated with different values of a in combination with cues and outcomes, and $f_{Y|X}$ describes the frequency distribution of $(Y|X)$, which depends on the decision maker's choice of a. Thus, frequencies and net payoffs are exhaustive determinants of what it means for a decision making rule to be adaptive, i.e., optimal with respect to the stochastic payoff structure implied by the pair of functions $\{f_{Y|X}(Y,X; a), \pi(Y,X, a)\}$.

An action is said to be *adaptive*, denoted a^*, if it maximises the expected payoff conditional on observable information X:

$$a^* \in \operatorname{argmax}_{a\in A} \int_{\mathcal{Y}} \pi(Y, X, a) f_{Y|X}(Y, X; a)\mathrm{d}Y. \tag{2}$$

The kth component of the vector X, X_k, is said to be *relevant* if the mean of Y, conditional on X, has a non-zero derivative in X_k somewhere in the domain of X_k. That is, X_k is relevant if and only if

$$\frac{\partial E[Y \mid X, a]}{\partial X_k} \neq 0 \text{ for some values of } X_k \text{ and } a. \tag{3}$$

Intuitively, relevance holds whenever the expected value of Y changes for some (but not necessarily every) change in X. The conditional mean function of Y, $E[Y|X, a]$ need not be linear in X.

Any mapping from X into the space of a is referred to as an *action function* $a(X)$. A member of this family of functions is said to *ignore information* whenever $\partial a(X)/\partial X_k = 0$ for all values of X_k.[3]

If all observable variables X are assumed to be relevant, then ignoring information is the same as ignoring relevant information. Thus, any function from X into A is an action function. A large subfamily of action functions ignores relevant information. It is assumed also that at least one action function maximises the expected payoff function and is therefore adaptive.

The question at the heart of this paper is whether there exists an action function that ignores relevant information and is adaptive. The investigation of this question proceeds in two steps. The first order condition necessary for an interior action to be adaptive is imposed first, characterising a^* as an implicit function of X. The second step examines the consequence of the independence of a^* from relevant cue X_k. The resulting condition defines a set of environments where payoffs and probabilities are such that adaptive action functions ignore relevant information. The following calculations seek to show that adaptive ignoring is not contradictory. The question remains, however, whether it is usual or to be expected in the real world, an issue that the paper returns to in a later section.

To repeat, the symbol a^* denotes not just a value in action space, but a mapping from X to the interior of A that is a (local) maximiser of the decision making objective. The decision-making objective is simply the conditional expected payoff

$$\int_Y \pi(Y,X,a) f_{Y|X}(Y,X;a)\mathrm{d}Y. \tag{4}$$

By applying the chain rule to the integrand in the expression above, the first-order condition that an adaptive action function $a.$ must satisfy is

$$\int_Y \left[\frac{\partial \pi(Y,X,a)}{\partial a} f_{Y|X}(Y,X;a) + \pi(Y,X,a) \frac{\partial f_{Y|X}(Y,X;a)}{\partial a} \right] \mathrm{d}Y = 0. \tag{5}$$

The first order condition defines a^* as an implicit function of X, which means that derivatives of a^* in X will, in general, be nonzero. That is, relevant information (variables that genuinely help predict Y) will in general influence adaptive decision rules. The next step is to find an additional condition under which the connection between a^* and X is broken.

When a component of X appears nowhere in the optimal action a^*, we describe it as the phenomenon of *ignoring relevant information* which, as stated earlier, occurs whenever there exists at least one information variable X_k such that

$$\frac{\partial a^*}{\partial X_k} = 0 \text{ and } \frac{\partial E[Y \mid X, a^*]}{\partial X_k} \neq 0. \tag{6}$$

The following theorem provides a condition that incorporates the first order condition (5) while imposing the further requirement that a^* is independent of X_k. The condition states precisely, which environments give rise to adaptive actions that ignore relevant information.

Theorem (When ignoring relevant information is adaptive**):** *Given a payoff function π and conditional probability distribution function $f_{Y|X}$ dependent on relevant information X, which satisfies the partial differential equation*

$$\int_Y \left[\frac{\partial^2 \pi(Y,X,a)}{\partial a \partial X_k} f_{Y|X}(Y,X;a) + \frac{\partial \pi(Y,X,a)}{\partial a} \frac{\partial f_{Y|X}(Y,X;a)}{\partial X_k} + \frac{\partial \pi(Y,X,a)}{\partial X_k} \frac{\partial f_{Y|X}(Y,X;a)}{\partial a} + \pi(Y,X,a) \frac{\partial^2 f_{Y|X}(Y,X;a)}{\partial a \partial X_k} \right] \mathrm{d}Y = 0, \tag{7}$$

adaptive action functions ignore the relevant variable X_k.

By definition a^* maximises expected payoffs conditional on X, and X is relevant, which implies that X_k helps to predict Y, or $\partial E[Y \mid X, a^*] / \partial X_k \neq 0$. The theorem says that whenever an environment satisfies equation (7), a^* ignores X_k, i.e., a^* is independent of X_k. A proof is provided in Berg and Hoffrage (2004).

4 A simple ignoring environment

Suppose the joint pdf of Y and X is

$$f_{YX}(Y, X; a) = Y + X \quad \text{for } Y, X \in [0, 1], \text{ and 0 otherwise.} \tag{8}$$

In this example, $f_{YX}(Y,X; a)$ happens to be independent of a, although this is not required for (7) to hold. By computing the marginal density f_X and computing the ratio f_{YX}/f_X, it is easy to show that the conditional pdf of Y, given X is

$$f_{Y|X} = \frac{Y + X}{(1/2) + X}. \tag{9}$$

Computing the conditional expectation of Y given X,

$$E[Y \mid X] = \int_0^1 Y \frac{Y + X}{(1/2) + X} \mathrm{d}Y = \frac{1}{(1/2) + X}\left(\frac{1}{3} + \frac{X}{2}\right), \tag{10}$$

it is clear that X is relevant because E[$Y|X$] is dependent on X (i.e., has nonzero derivative for some, in this case all, values of X):

$$\partial E[Y \mid X] = \frac{1}{12(1/2 + X)^2} \neq 0, \quad \textit{for } \text{all } X \in [0,1]. \tag{11}$$

Suppose the payoff function is given by the product of the reciprocal of $f_{Y|X}$ and a simple function of a with a global maximum:

$$\pi(Y, X, a) = \frac{(1/2) + X}{Y + X}(a - a^2). \tag{12}$$

Then the objective function

$$\mathrm{E}[\pi(Y, X, a)|X] = a - a^2 \tag{13}$$

is independent of X, and $a^* = 1/2$. In this case, there is a unique adaptive action, $a = 1/2$, and it remains constant, irrespective of the relevant information X.

Because $\partial a^*/\partial X = 0$, the environment given by f_{YX} and π satisfies condition (7) from the theorem, demonstrating that the class of adaptive ignoring environments is nonempty.

In the preceding example, the joint distribution of X and Y was independent of a, the action space A was the real line, the payoff function itself had a global max (was non-monotonic) in a, and the optimal action happened to be a constant function. Berg and Hoffrage (2004) provide more complex examples which demonstrate that those special features from the example above are not required for equation (7) to hold in general.

A number of important questions remain, however, concerning the nature of environments where ignoring information is adaptive. Just how broad is the class of distribution and payoff functions that support adaptive ignoring, and what do those environments look like? The next section provides a more intuitive characterisation of those environments.

5 The role of asymmetry

Equation (7) is necessary for adaptive ignoring to occur. The condition covers a precise and complete description of environments where the phenomenon occurs. However, it is not obvious how it is to be interpreted or verified empirically. How would one check to see if a real-world decision-making environment satisfied equation (7)?

To better understand what underlies adaptive ignoring, a simplification is proposed. Within the larger set of adaptive-ignoring environments defined by (7), there is a subset where it is simple to see the payoff-probability structure that is described. By focussing on this subset, insight is gained into empirical clues that suggest that (7) holds approximately. Because it is a subset, the characterisation offered below is incomplete. However, the subset is large enough to be interesting in its own right, and produces a new result concerning J-shaped distributions, a class of asymmetric probability distributions that plays a key role in many areas of social sciences.

The subset of adaptive-ignoring environments is defined by the condition that the product $f_{Y|X}\pi$ is independent of X_k:

$$S \equiv \{\{f_{YX}, \pi\} | f_{Y|X}\, \pi = g\,(Y, X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots X_k, a)\}, \tag{14}$$

where $g\,(X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots X_k, a)\}$ is any scalar valued function that does not depend on X_k for some $k \in \{1, \ldots, K\}$.

It is easy to verify that S is a subset of the class of adaptive ignoring environments defined by condition (7). The independence of the product $f_{Y|X}\pi$ from X_k implies that

$$\frac{\partial}{\partial X_k}[f_{Y|X}(Y, X; a)\pi(Y, X, a)] = 0, \tag{15}$$

from which it follows:

$$\frac{\partial^2}{\partial X_k a}[f_{Y|X}(Y, X; a)\pi(Y, X, a)] = 0. \tag{16}$$

Expanding the equation above by two applications of the chain rule produces the four terms in the integrand of equation (7) which must sum to zero. Thus, if an environment belongs to S, it satisfies the equation above and therefore satisfies the more general condition (7). In other words, every member of S is an optimal ignoring environment.

The reason for considering the subclass S is that the condition $f_{Y|X}\pi = 0$ makes transparently clear what is going on in such an environment. There is cancellation between probabilities (i.e., values of the pdf, or likelihoods, which, technically speaking, are not probabilities when f_{YX} is continuous) and payoffs, in response to a change in the kth cue. Consider a fixed value $Y = y^+$ that is better than average in terms of fitness considerations as summarised by π, i.e., $\pi(y^+, \bar{x}, a) > \pi(\bar{y}, \bar{x}, a)$. When the average cue value $X_k = \bar{x}_k$ is observed, the likelihood of $Y = y^+$ is $f_{Y|X}(y^+, \bar{x}, a)$ and the payoff is $\pi(y^+, \bar{x}, a)$. Now suppose that in another realisation of the same environment a different value of X_k is observed, $X_k = x_k^+$, which makes the outcome y^+ more likely. That means the value of $f_{Y|X}$ increases by the percentage

$$p^+ = \left[\frac{f_{Y|X}(y^+, \bar{x}_1, \ldots, \bar{x}_{k-1}, x_k^+, \bar{x}_{k+1}, \ldots, \bar{x}_K, a)}{f_{Y|X}(y^+, \bar{x}, a)} - 1 \right] \times 100. \quad (17)$$

If the payoff associated with the average value $Y = y^+$ decreases approximately proportionately, falling by p^+ percent, i.e., if

$$\pi(y^+, x_k^+, a) / \pi(y^+, \bar{x}, a) - 1 = -p^+, \quad (18)$$

then the environment belongs to S and adaptive ignoring prevails. The following scenarios illustrate how probability-payoff cancellation could occur in the real world.

Strawberry/raspberry gathering

The stochastic outcome Y represents the amount of raspberries that can be harvested in an hour of gathering. Action a is the amount of time the individual devotes to raspberry gathering. Among various environmental cues that help predict how dense raspberry growth at a particular time of year is expected to be is the quality of strawberry harvests over the preceding month. Strawberry and raspberry harvests reflect a common set of weather patterns, but strawberries ripen earlier and therefore precede raspberries in manifesting signs of the current season's growing conditions.

Thus, better than average strawberries predict better than average raspberries. It follows that decision makers should choose raspberry-harvest intensity a in accordance with the observed quality of strawberries.

Consider however what happens when better than average strawberry harvests leave gatherers more nutritionally fortified, with more stored energy in their bodies, therefore leading to lower payoffs for raspberry consumption. Strawberry harvests being equal, gatherers will spend more or less time on raspberries depending on whether the expected harvest is high or low. But good strawberry harvests cause the nutritional payoff of raspberry consumption to reduce proportionally. Likewise, low strawberry harvests lead to hungrier gatherers who receive higher payoffs from an extra unit of raspberry consumption. The probabilistic change in harvest is offset by a corresponding change in payoffs based on nutritional need.

Stock market investing with two periods of consumption

Current economic conditions X_k are observed at t_0, when a consumption/savings decision a must be made. Whatever is saved today goes into the stock market. Y represents the stochastic value of the stock portfolio at a later date t_1. In general, action functions $a(X)$ depend on the observed value of X_k, because better current economic conditions (higher values of X_k) predict higher expected values for stock portfolio outcomes.

However, consider what happens when good economic times at t_0 (better than average current economic conditions X_k) lead to a reduced need for future income Y. For example, if the decision maker's primary motivation to save is to pay for a child's education at t_1, which can be paid out of pocket when an unexpected raise is received

(X_k is better than average), then the environment is likely belongs to S. If the utility of income was constant with respect to X_k, then higher expected future returns should induce increased savings. However, the case described above is one in which the increased likelihood of high stock values coincides with a lower subjective payoff associated with those high stock values, leading to an adaptive ignoring environment in which X_k predicts Y but does not influence the decision of how much to invest.

Storms in April predict storms in May

Suppose the severity of storms in May, Y, can be predicted to some extent by the severity of storms in April, X. Costly actions such as leaving town or investing in window and roof reinforcement can reduce physical harm (raise payoffs) in the event of a storm. However, once a storm occurs in April and causes damage, the additional damage from storms in May is reduced.

That is, if $Y = y_0$ represents the event of a storm in May, $X = x_0$ is a storm in April, and $X = \bar{x}$ is the average value of an aggregate of wind speed and precipitation for the month of April, we have:

$$f_{Y|X}(y_0, x_0, a) > f_{Y|X}(y_0, \bar{x}, a) \quad \text{and} \quad \pi(y_0, x_0, a) < \pi(y_0, \bar{x}, a) \text{ for all } a. \tag{19}$$

A storm in April makes a storm more likely in May, but the payoff to preventive measures in May is lower once the storm occurs in April. Because of probability-payoff cancellation, it is adaptive to observe a storm in April (and therefore revise the probability of storms in May upward) yet not increase preventive measures a at all.

6 Is cancelation too strict?

The defining condition for optimal ignoring and, as a special case, the cancellation principle, may appear to be insufficiently general. After all, how often does one find environments that feature exact cancellation? In many parameterised environments, the optimal ignoring subset would indeed have the measure zero, and therefore generically fail to exist.

This alone is not a valid reason to ignore the possibility of optimal ignoring, as economic theory provides abundant examples of measure-zero phenomena which are considered to nevertheless be informative and interesting. The unit root in econometrics, the saddle-path equilibrium in macroeconomics, and constant-returns-to-scale models in theories of industrial organisation are three prominent examples. One might even argue that any model in which it is assumed that decision makers' beliefs are objectively correct (i.e., rational expectations models) imposes a generically untrue restriction. In a continuous probability space, the event that subjective beliefs and objective frequencies match is as likely as a spinner landing on a single number, or the probability that a normal random variable is exactly equal to 1.80.

The positive reasons for considering the possibility of optimal ignoring follow from the impressive empirical evidence in favour of ignoring, revealed in experimental studies (e.g., those cited in Section 2). Two frequently cited experiments in which prior probabilities were ignored are Kahneman and Tversky's (1973) lawyer-doctor problem and Tversky and Kahneman's (1982) taxi cab problem. In both studies, respondents were asked to state posterior probabilities, given a prior probability, hit rate and false positive rate. Varying the prior probability, referred to as the base rate, these studies showed a surprising lack of sensitivity to changes in the base rate, a phenomenon which came to be known as base rate neglect or the base rate fallacy.

More broadly, one thinks of many applied business-decision problems and questions of economic policy that are both consequential and involve decision makers' ability to discern and use all the relevant information. Schwartz (1987) shows that business leaders often ignore the full range of suppliers of inputs, and that some international firms ignore exchange rate risk. Do these firms behave this way at their firms' peril, or is there theoretical justification for their frugality with respect to information? In the realm of economic policy, social security reform clearly depends on how workers use currently available information about their lifetime earnings prospects to make savings/consumption decisions today. Government outlays for healthcare depend on how the population uses available information about factors that condition future health outcomes, such as smoking, exposure to environmental toxins, diet and exercise. And political movements depend in part on perceptions about whether the future will significantly change as a result of investments in activism, movements and campaigns in the present. (See Berg, 2003, for a broader discussion of normative behavioural economics).

In contrast to neoclassical economics' exclusively pathological interpretation of ignoring, this paper makes no claim as to whether ignoring in any of theses contexts is necessarily good or bad. The claim is simply that ignoring makes sense in certain environments, but not in all environments. Thus, prescriptive claims about how decision makers ought to process information must be tied to particular decision-making contexts rather than resting on universal or axiomatic claims that directly characterise the use of information under uncertainty.

Given the abundant evidence that decision makers ignore information, the question turns to explaining why this happens. Time, memory and problem-solving costs figure prominently in the theory of bounded rationality and comprise a highly convincing category of explanations for why ignoring occurs. However, as Simon pointed out, cognitive processes should not be evaluated in a vacuum. A context is required to determine how adaptive decision rules are. This requires one to specify a particular decision-making environment. When the payoff and likelihood structure of the environment favour decision rules that ignore or approximately ignore an otherwise relevant piece of information (in terms of its predictive power), one expects to see evolutionarily successful decision makers adopt relatively frugal action rules.

Thus the central claim proposed here is that environmental payoff and likelihood structure should be considered as a new category of explanations, consistent with and complementary to bounded rationality, seeking to understand why human decision makers employ the rules they do. There is a normative claim too. While ignoring may be detrimental to decision makers' success in some environments, it can help in others. The reason why ignoring is prevalent is not merely because of hard-wired constraints, but also because there exist environments where ignoring is smart. This complicates the normative analysis of ignoring, and suggests that entirely negative normative interpretations are premature and possibly misplaced. Instead of proclaiming ignoring to be categorically pathological, the existence of optimal ignoring environments argues instead for a mixed, environment-specific mode of analysis that respects the principle of ecological rationality.

Finally, the case for the relevance of optimal ignoring and the ignoring theorem can be strengthened by introducing the concept of just noticeable differences in payoffs. If a one unit change in a cue variable X leads to no perceptible change in $\pi(Y, X, a)$ or $f_{Y|X}(Y, X, a)$, for all values of a, then $E\pi\ (Y, X, a)$ is independent of a, and the link between decision rules and a is broken. A less trivial case is where X perceptibly influences π and $f_{Y|X}(Y, X, a)$ but, because of *approximate* cancellation, there is no perceptible change in $E\pi\ (Y, X, a)$. In such a case, by requiring only approximate cancellation, there would indeed exist dense subsets of parameter space in which ignoring occurs, implying generic existence.

7 J-shaped distributions

In the light of previous work on so-called J-shaped environments in connection with limited rationality, it is intriguing that the subclass of rational ignoring environments S implies a J-shaped conditional density $f_{Y|X}$. Although different definitions of J-shaped environments have been put forth, this paper defines the conditional density $f_{Y|X}$ as J-*shaped* whenever it is monotonic in Y (holding X constant). In combination with a payoff function that is monotonic in Y, where high payoffs correspond to low probabilities and low payoffs correspond to high probabilities, the cancellation principle follows. If in addition, the magnitudes of the inverse movement in $f_{Y|X}$ and π are approximately equal, then the J shaped environment belongs to S.

Within S, cancellation between $f_{Y|X}$ and π with respect to Y occurs whenever the product $f_{Y|X}\ \pi$ is independent of both X_k and Y, in which case the conditional pdf can be expressed as:

$$f_{Y|X}(Y, X; a) = \frac{g(X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_k a)}{\pi(Y, X, a)}. \tag{20}$$

Thus, for an observed vector of cues and a given action, if the decision maker is better off with higher than average realisations of Y , i.e., payoffs are monotonic in Y , and condition (20) holds, then Y cannot be symmetrically distributed about its expected value of Y . With condition (20) satisfied, the conditional density $f_{Y|X}$ inherits monotonicity in Y from the monotonicity of the payoff function π.

Monotonicity of π with respect to Y arises in many real world contexts. Whenever higher values of Y are unambiguously good or unambiguously bad, π is monotonic in Y, and π will therefore be J-shaped anytime rational ignoring occurs. The construction of S provides an intuitive characterisation of rational ignoring environments: if conditional probabilities of Y are highly skewed, then conditions are favourable for rational, i.e., adaptive, ignoring.

$$\frac{\partial \pi}{\partial Y} > 0 \Rightarrow \frac{\partial f_{Y|X}}{\partial Y} = -(g/\pi^2)/\frac{\partial \pi}{\partial Y} < 0. \tag{21}$$

8 Conclusion

The phenomenon of ignoring information is frequently encountered in the experimental laboratory across a broad range of settings. This phenomenon is traditionally interpreted in one of two ways. Ignoring can be viewed as an artefact of hard-wired constraints on the decision maker's cognitive capacities, e.g., memory, channels of attention, or processing speed. The second interpretation while sharing the basic premise that ignoring is the consequence of cognitive limitations, views ignoring in a more positive light as a strategic or efficiency-enhancing response to limited capacity.

In contrast to these explanations in terms of cognitive constraints, this paper offers a new theoretical account as to why suppression of information is pervasive. According to this view, ignoring is optimal (in particular environments) simply because of the manner in which payoff functions and probability distributions interact. When those components of the decision making environment exhibit the inverse relationship described above, adaptive ignoring is predicted. In other words, when low probability events are associated with exclusively good payoffs or exclusively bad payoffs, and the majority of the payoff distribution clusters around either its maximum or minimum value, adaptive ignoring is likely to occur.

The main theorem in this paper provides a sufficient condition, equation (7), for expected-payoff -maximising, i.e., adaptive, ignoring to prevail. Thus, the formal model succeeds in generating a new prediction that can be tested, provided payoffs and conditional distributions ($Y|X$) are observable. In other words, where systematic ignoring is observed in the field, one may investigate whether the cancellation principle described here is a plausible cause. Future work should refine these predictions, test them, and pursue the theoretical question of just how widespread adaptive-ignoring environments are in the larger space of decision-making environments.

Acknowledgement

The author thanks an anonymous referee for providing valuable feedback and criticism. Research assistance from Yu Xue is gratefully acknowledged.

References

Berg, N. (2003) 'Normative behavioral economics', *Journal of Socio-Economics*, Vol. 32, No. 4, pp.411–427.

Berg, N. and Hoffrage, U. (2004) 'Do unboundedly rational decision makers ignore relevant information?', *Working Paper*, University of Texas at Dallas.

Bigand, E., McAdams, S. and Foret, S. (2000) 'Divided attention in music', *International Journal of Psychology*, Vol. 35, No. 6, pp.270–278.

Conlisk, J. (1996) 'Why bounded rationality?', *Journal of Economic Literature*, Vol. 34, No. 2, pp.669–700.

Crozier, W.R. and Skliopidou, E. (2002) 'Adult recollections of namecalling at school', *Educational Psychology*, Vol. 22, No. 1, pp.113–124.

Cutler, D., Porterba, J. and Summers, L. (1989) 'What moves stock prices?', *Journal of Portfolio Management*, Vol. 15, No. 3, pp.4–12.

Dishon-Berkovits, M. and Algom, D. (2000) 'The Stroop effect: it is not the robust phenomenon that you have thought it to be', *Memory & Cognition*, Vol. 28, pp.1437–1449.

Dow, J. (1991) 'Search decisions with limited memory', *Review of Economic Studies*, Vol. 58, pp.1–14.

Emran, S.M. and Ye, N. (2001) 'Robustness of canberra metric in computer intrusion detection', *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, pp.80–84, West Point, NY. The paper is available online at: http://216.239.59.104/search?q=cache:Y2-hOGHJXDUJ:www.itoc.usma.edu/Workshop/2001/Authors/Submitted_Abstracts/paperT2A1(02).pdf+robustness+of+canberra+metric&hl=en

Erber, R. and Fiske, S.T. (1984) 'Outcome dependency and attention to inconsistent information', *Journal of Personality and Social Psychology*, Vol. 47, No. 4, pp.709–726.

Ferri, C., Flach, P. and Hernandez-Orallo, J. (2001) *Learning Decision Trees using the Area under the ROC Curve*, Working Paper, University of Valencia.

Gerard, R.W. (1960) 'Neurophysiology: an integration', in Field, J., Magoun, H.W. and Hall, V.E. (Eds.): *Handbook of Physiology*, Vol. III, pp.1919–1965, Amer. Physiol. Soc., Washington DC.

Godijn, R. and Theeuwes, J. (2002) 'Programming of endogenous and exogenous saccades: evidence for a competitive integration model', *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 28, No. 5, pp.1039–1054.

Hasher, L. and Zacks, R.T. (1979) 'Automatic and effortful processes in memory', *Journal of Experimental Psychology: General*, Vol. 108, pp.356–388.

Hutchison, K.A. (2002) 'The effect of asymmetrical association on positive and negative semantic priming', *Memory and Cognition*, Vol. 30, No. 8, pp.1263–1276.

James, W. (1890) *The Principles of Psychology*, Reprinted, Harvard University Press 1983, Cambridge MA.

Jones, D.R. and Brown, D.R. (2002) 'The division of labor between human and computer in the presence of decision support system advice', *Decision Support Systems*, Vol. 33, pp.375–388.

Kahneman, D. and Triesman, A. (1984) 'Changing views of attention and automaticity', in Parasuraman, R. and Davies, D.R. (Eds.): *Varieties of Attention*, Orlando, Academic Press, FL, pp.29–61.

Kahneman, D. and Tversky, A. (1973) 'On the psychology of prediction', *Psychological Review*, Vol. 80, pp.237–251.

Lipman, B.L. (1991) 'How to decide how to decide how to …: modeling limited rationality', *Econometrica*, Vol. 59, No. 4, pp.1105–1125.

Lippa, Y. and Goldstone, R.L. (2001) 'The acquisition of automatic response biases through categorization', *Memory and Cognition*, Vol. 29, No. 7, pp.1051–1060.

Logan, G.D. (1980) 'Attention and automaticity in stroop and priming tasks: theory and data', *Cognitive Psychology*, Vol. 12, pp.523–553.

McDermott, J., Weiss, Y. and Adelson, E.H. (2000) 'Beyond junctions: nonlocal form constraints on motion', *Perception*, Vol. 30, pp.905–923.

Most, S.B., Simons, D.J., Scholl, B.J., Jiminez, R., Clifford, E. and Chabris, C.F. (2001) 'How not to be seen: the contribution of similarity and selective ignoring to sustained inattentional blindness', *Psychological Science*, Vol. 12, No. 1, pp.9–17.

Mullainathan, S. (2002) 'A memory based model of bounded rationality', *Quarterly Journal of Economics*, Vol. 117, No. 3, pp.735–774.

Neuman, W. (2002) 'The presence of the past: using memory work to search for psychological traces of the Nazi past in contemporary Germans', *British Journal of Guidance and Counseling*, Vol. 30, No. 1, pp.5–16.

Posner, M.I. (1964) 'Information reduction in the analysis of sequential tasks', *Psychological Review*, Vol. 71, No. 6, pp.491–504.

Reingold, E.B., Charness, N., Schultetus, R.S. and Stampe, D.M. (2001) 'Perceptual automaticity in expert chess players: parallel encoding of chess relations', *Psychonomic Bulletin and Review*, Vol. 8, No. 3, pp.504–510.

Rodriguez-Fornells, A., Rotte, M., Heinze, H., Nosselt, T. and Munte, T.F. (2002) 'Brain potential and functional MRI evidence for how to handle two languages with one brain', *Nature*, Vol. 415, pp.1026–1029.

Ruthruff, E., Remington, R.W. and Johnston, J.C. (2001) 'Switching between simple cognitive tasks: the interaction of top-down and bottom-up factors', *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 27, No. 6, pp.1404–1419.

Schein, A.I., Popescul, A., Ungar, L.H. and Pennock, D.M. (2002) 'Methods and metrics for cold-start recommendations', *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *SIGIR'02,* 11–15 August, Tampere, Finland, pp.253–260.

Schiller, N.O. and Caramazza, A. (2002) 'The selection of grammatical features in word production: the case of plural nouns in German', *Brain and Language*, Vol. 81, pp.342–357.

Schubo, A., Meinecke, C. and Schroger, E. (2001) 'Automaticity and attention: investigating automatic processing in texture segmentation with event-related brain potentials', *Cognitive Brain Research*, Vol. 11, No. 3, pp.341–361.

Schwartz, H. (1987) 'Perception, judgment, and motivation in manufacturing enterprises: findings and preliminary hypotheses from in-depth interviews', *Journal of Economic Behavior and* Organization, Vol. 8, No. 4, pp.543–565.

Shiu, L.P. and Pashler, H. (1992) 'Improvement in line orientation: discrimination is retinally local but dependent on cognitive set', *Perception and Psychophysics*, Vol. 52, pp.582–588.

Stroop, J.R. (1935) 'Studies of interference in serial verbal reactions', *Journal of Experimental Psychology*, Vol. 18, pp.643–662.

Tversky, A. and Kahneman, D. (1982) 'Evidential impact of base rates', in Kahneman, D., Slovic, P. and Tversky, A. (Eds.): *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, pp.153–160.

Walker, C.N. and Bourne Jr., L.E., (1961) 'The identification of concepts as a function of amounts of relevant and irrelevant information', *American Journal of Psychology*, Vol. 74, pp.410–415.

Wentura, D. (2002) 'Ignoring 'brutal' will make 'numid' more pleasant but 'uyrvu' more unpleasant: the role of a priori pleasantness of unfamiliar stimuli in affective priming tasks', *Cognition and Emotion*, Vol. 16, No. 2, pp.269–298.

Notes

[1]For simplicity, *Y* is discussed throughout as a scalar and *Y* is a subset of the real line. This could be generalised to a vector of outcomes with few alterations to the results that follow.

[2]If interpreted as an expected utility function, a number of delicate details regarding π and its subjective content (beliefs about the probabilities of different states of *Y*) must be addressed. For example, are beliefs about probabilities correct? Are there impossible (zero probability) events assigned positive weight, or positive probability events ignored altogether? These issues are discussed by Lipman (1991) who demonstrates that subjectivity can in principle be handled using a bounded rationality framework with explicit decision costs. Addressing these issues here, however, distracts from the main objective, which is to characterise environments where optimal ignoring arises with no decision costs whatsoever, simply as a result of the probability and payoff structure of the environment.

[3]It may be argued that even when the condition holds, optimisers do not 'ignore' information, rather they must carefully analyse all cues in order to determine which can be left out of the optimal action rule. However, this criticism applies only when the payoff function is interpreted as a subjective utility function. When instead the payoff function is viewed as an objective summary of adaptive fitness, then the theorem in this paper says that the environment will reward individuals with decision rules that do not depend on *X* and penalise those who condition *a* on *X*.