Entropy and the value of information for investors

Antonio Cabrales[*]   Olivier Gossner[†]   Roberto Serrano[‡]

November 2010

Abstract: Consider any investor who fears ruin facing any set of investments that satisfy no-arbitrage. Before investing, he can purchase information about the state of nature in the form of an information structure. Given his prior, information structure α is more informative than information structure β if whenever he rejects α at some price, he also rejects β at that price. We show that this complete informativeness ordering is represented by the decrease in entropy of his beliefs, regardless of his preferences, initial wealth or investment problem. It is also shown that no prior-independent informativeness ordering based on similar premises exists.

[*]Department of Economics, Universidad Carlos III de Madrid
[†]Paris School of Economics and Department of Mathematics, London School of Economics
[‡]Department of Economics, Brown University and IMDEA-Social Sciences Institute

1 Introduction

Consider a decision maker operating under uncertainty. When can one say that a new piece of information is more valuable to such an agent than another? The usual answer to this question is that such a ranking is in general hard, because it typically depends upon at least three considerations. (i) The agent's priors matter (an agent who is almost convinced that a serious crisis in the strength of the dollar is forthcoming versus one who is not may rank the appearance of bad financial news in China or in Europe very differently). (ii) The preferences and/or wealth of the agent matter (for the same prior, two agents who have different degrees of risk aversion may rank news that almost eliminate uncertainty in terms of a financial loss versus more uncertain news about it in distinct ways). And (iii) the decision problem in which information is applied matters (whether eliminating uncertainty is to be used in order to make a financial investment or to find out about a terrible disease).

We model uncertainty by a finite set of states of nature and a prior probability distribution on the set of states. Before making a decision, the decision maker can access additional information about the state. This is done via an information structure, a finite set of signals that, when observing each of them, leads to an update of the prior into the corresponding posterior. With this terminology, our question is when an information structure provides

more information than another.

The first answer to this fundamental question was provided in the seminal work of Blackwell (1953). Blackwell showed that an information structure is more informative than another whenever the latter is a garbling of the former, i.e., there exists a stochastic matrix –interpreted as noise– such that the matrix of conditional probabilities of each signal for the former structure is the one for the latter multiplied by the stochastic matrix. Moreover, this happens if and only if a decision maker with *any* utility function would prefer to use the former information structure over the latter when facing *any* decision problem. This is of course an extremely strong result, which not surprisingly, provides only a partial ordering of information structures. That is, most of the time two information structures cannot be ordered in the sense of Blackwell.

Following recent developments in the theory of riskiness, our attempt here is to provide an approach based on decision-theoretic principles in order to complete Blackwell's ordering.[1] Restricting attention to a class of no-arbitrage investment decisions first studied in Arrow (1971) and to a specific class of utility functions that like to avoid bankruptcy in a sense specified below, we postulate the following informativeness ordering. Fixing a prior over the states, we say that an information structure is more informative than

---

[1]In particular, we follow closely a recent paper by Hart (2010), in which two orderings are proposed to justify the Aumann and Serrano (2008) index of riskiness and the Foster and Hart (2009) measure of riskiness.

another if, over the allowed class of problems and preferences, whenever the first one is rejected at some price, so is the second. This seems a minimum desideratum for a notion of informativeness.

Our main result is that this informativeness ordering is represented exclusively by the decrease in entropy of the agent's beliefs. Specifically, if one considers the prior and the collection of posteriors generated by the information structure, we show that the informativeness of the information structure is the difference between the entropy of the prior distribution and the expected entropy of the conditional posterior distributions. This is referred to as "rate of transmission" in Arrow (1971), who works with a set of Arrow securities and a logarithmic utility function.[2]

By restricting our scope to specific classes of problems and preferences, our approach based on total rejections of information structures suggests entropy as the unique "objective" way to speak of the informativeness of information structures when dealing with preferences, wealth levels and decision problems in the classes we consider. On the other hand, the approach is limited in one important respect, in that it is dependent on the prior. Indeed, we also show that there does not exist any prior-independent informativeness ordering based on the same postulate, as the entropy ordering is affected by priors. Therefore, revisiting the difficulties described in the first

---

[2]Unlike the riskiness papers mentioned in the previous footnote, our decision-theoretic considerations here do not uncover a new index, but provide a new support to the classic concept of entropy.

paragraph, we provide a framework in which our complete informativeness ordering takes care of considerations (ii) and (iii), but cannot possibly do the same with (i).

1.1 Related Literature

The notion of entropy has appeared extensively in information theory. Many results can be found, including useful characterizations that rely on its many good properties. Shannon (1948), for example, shows that when expressing information as a Markov process, only three properties (continuity, monotonicity, and decomposability) characterize an entropy index of informativeness. Other characterizations rely on the informativeness of a pair of independent random variables being the sum of their respective levels of informativeness, or on the communication cost of conveying the information (see Gossner, Hernández, and Neyman, 2006). These authors study a repeated game between a forecaster, an agent, and nature. They prove that the information constraint (the agent cannot know more than the information he receives) expressed with an entropy function fully characterizes the set of implementable action triples (state of nature, forecaster's action, agent's action). The major difference between these existing characterizations and our work is that we show that entropy characterizes the *value* of information, hence establishing that any information pricing should reflect the entropy measure. Another important difference with our approach is that we ar-

rive at the entropy idea from completely different considerations, rooted in economic and decision-theoretic arguments.[3]

One of the first proponents of entropy as an adequate measure of the value of information in economics is Marschak (1959). His article contains also one of the first formal discussions on the demand and cost of information. Arrow (1971) considers an investor who has access to a set of securities which pay a positive amount in only one state of nature. He shows that, if the value of information about the state is independent of the returns, then necessarily this value is given by the entropy informativeness.

In economics, entropy has more recently also been used by Sims (2003) to model limits in human information processing capabilities, which he called "rational inattention". This approach has been applied to many different economic problems. For instance, Peng (2005) has explored its implications for asset price dynamics and consumption behavior; Kondor (2010) has shown that rational inattention can lead to counterintuitive effects on higher order expectations; see Sims (2007) for a summary of other contributions in this area.

Measuring the amount of information is a common problem in economics and decision theory.[4] Most of the work on this area follows the path-breaking

---

[3]See also Hershey (2009) for a general discussion of the use of entropy and the applications of the second law of thermodynamics in many different contexts, including physical, biological, informational and the business world.

[4]Veldkamp (2011) provides a good summary of ways in which economists have measured informativeness and its applications.

work of Blackwell (1953). For Blackwell, an information structure α is more informative than another one β, if every decision maker prefers α to β in any decision problem. As noted, the main drawback of his approach is that this criterion does not provide a complete ordering. The literature has progressed by focusing on decision makers having preferences in a particular class. Lehmann (1988), for instance, restricts the analysis to problems that generate monotone decision rules (and hence satisfy single-crossing conditions). Athey and Levin (2001) and Jewitt (2007) extend Lehmann's analysis to more general kinds of monotone problems. We follow this tradition with two main differences. Unlike the measures in those papers, our measure of informativeness provides a complete order of all information structures. We achieve this through a different kind of restriction on admissible preference orderings, and we characterize decision problems in terms of investment opportunities, thereby restricting the framework.

Azrieli and Lehrer (2008) take a reverse approach to the one just described. Rather than choosing a class of decision problems, and then providing an ordering of information structures, they characterize the orderings that are possible for any prespecified class of decision problems. They show that reducibility, weak order, independence, continuity and convexity characterize all binary relations on information structures induced by decision problems.

A recent paper by Ganuza and Penalva (2010) provides a different way to

measure informativeness (also a partial order) not based on decision-theoretic considerations, but rather on various measures of dispersion of distributions (many of those measures are presented in Shaked and Shanthikumar, 2007). They show that while some of their measures are implied by notions of informativeness based on the value of information, the strongest of their criteria, supermodular precision, is strictly different (it neither implies nor is implied) by those notions of informativeness. Then they study the implications of greater informativeness (in their sense) to auction problems, and show that while greater informativeness improves allocational efficiency, the auction organizers are not always interested in increasing informativeness since that may increase buyers' information rents.

1.2   Plan of the Paper

Section 2 introduces the investment problems that we study. Section 3 investigates valuable investment opportunities in terms of two conditions imposed jointly on utility functions and investment problems: no investment under no information (NINI) and only investors who are sometimes certain always invest (SCAI). The section characterizes NINI and SCAI by means of ruin averse utility functions and no-arbitrage investment sets. Section 4 introduces our informativeness ordering using the identified preferences and investment sets, and proves the main result. Section 5 offers several points of discussion, and the Appendix collects the proofs.

2   Investments and Uncertainty

We measure the value of information according to its relevance to investment choices. To this end, we rely on a standard model of investment under uncertainty *à la* Arrow (1971)[5].

We consider an agent with initial wealth w, and with increasing monetary and twice differentiable utility function $u \colon \mathbb{R}_+ \to \mathbb{R}$. The coefficient of relative risk aversion at wealth $z > 0$ is:

$$\rho(z) = -\frac{u''(z)z}{u'(z)}$$

and we assume that the agent has weakly increasing relative risk aversion (IRRA), namely that ρ is non-decreasing on $\mathbb{R}_+$. (This class includes the constant absolute risk aversion -CARA- and the constant relative risk aversion -CRRA- classes, used often in applications.) We denote by $\mathcal{U}_0$ the set of such monetary utility functions. For $u \in \mathcal{U}_0$, we let $u(0) = \lim_{z \to 0} u(z) \in \mathbb{R} \cup \{-\infty\}$.

Let K be the finite set of states of nature. The investor has a prior belief p with full support, fixed throughout the paper.[6] An investment opportunity or asset is $b \in \mathbb{R}^K$, with the interpretation that if b is taken, the agent's wealth once uncertainty realizes is $w + b_k$ in state k. We do not allow for bankruptcy (the possibility for negative wealth) and say that an asset b is

---

[5]See Mas-Colell, Whinston, and Green (1995), Section 19.E on a general equilibrium model with assets paying in a numeraire commodity.

[6]Except for Subsection 5.1, where we discuss the impossibility of a prior-independent ordering

feasible at wealth w when $w + b_k \geq 0$ in every state $k \in K$.

The investor has the opportunity to choose from an investment set B of assets, from which he must take one. Among the choices is the possibility to opt out, namely to keep his wealth w in a safe asset. This assumption can be formulated as $0_K \in B$, where 0_K is the null vector of $\mathbb{R}^K$. An investment set thus consists of a subset of $\mathbb{R}^K$ containing 0_K. The set B can for instance consist of a set of Arrow securities, or any other complete or even incomplete markets asset structure. Elements in B can be either divisible (for every $b \in B$, $\lambda \in [0, 1]$, $\lambda b \in B$) or indivisible. We say that an investment set B is feasible at w when all its elements are feasible at w.

An information structure α is given by a finite set of signals S_α, together with transition probabilities $\alpha_k \in \Delta(S_\alpha)$ for every k. When the state of nature is k, $\alpha_k(s)$ is the probability that the signal observed by the agent is s. It is standard to represent any such information structure by a stochastic matrix, with as many rows as states and as many columns as signals; in the matrix, row k is the probability distribution $(\alpha_k(s))_{s \in S_\alpha}$. We assume that every signal s has positive probability under at least one state k, otherwise, such signals can be deleted from the set S_α.

It is useful to think of α in terms of a distribution over posterior probabilities. Signal s has a total probability $p_\alpha(s) = \sum_k p(k)\alpha_k(s)$, and the agent's posterior probability on K given s is q_α^s derived through Bayes' formula:

$$q_\alpha^s(k) = \frac{p(k)\alpha_k(s)}{p_\alpha(s)}$$

Information structures are ranked according to the partial Blackwell (1953) ordering. A most informative information structure, denoted $\overline{\alpha}$, is one that perfectly reveals the state of nature k, hence such that for any s, there exists a unique k such that $\overline{\alpha}_k(s) > 0$. The least informative is any information structure $\underline{\alpha}$ with no informational content $((\underline{\alpha}_k(s))_{s \in S_{\underline{\alpha}}}$ is the same distribution for all k).

3   Valuable Information

We make two assumptions on the agent's utility function u together with the set B of available choices that ensure that the class of utility functions and investment sets are suitable to rank informativeness. Both assumptions are formulated in terms of investment decisions under incomplete information.

Given a utility function u, initial wealth w, a feasible investment set B and a belief $q \in \Delta(K)$, the maximal expected utility that can be reached by choosing an investment opportunity $b \in B$ is

$$v(u, w, B, q) = \sup_{b \in B, \ b \text{ feasible}} \sum_k q(k)u(w + b_k)$$

with the convention that $0.(-\infty) = 0$. The ex-ante expected payoff before receiving signal s from α is

$$\pi(\alpha, u, w, B) = \sum_s p_\alpha(s)v(u, w, B, q_\alpha^s).$$

The possibility to opt out ensures that both $v(u, w, B, q)$ and $\pi(\alpha, u, w, B)$ are always larger or equal than $u(w)$. The gain from investment opportunities

in B and information α is the difference

$$V(\alpha, u, w, B) = \pi(\alpha, u, w, B) - u(w).$$

It is often useful to rewrite

$$V(\alpha, u, w, B) = \sum_s p_\alpha(s)(v(u, w, B, q_\alpha^s) - u(w)).$$

and this last expression shows that $V(\alpha, u, w, B) > 0$ if and only if there exists s such that $v(u, w, B, q_\alpha^s) > u(w)$.

Our first assumption is *No Investment under No Information*, or NINI for short. According to this assumption, under the absence of information except for the prior p, the agent prefers to opt out than investing in risky elements of B. It is a joint assumption on the possible investment set B and the agent's utility function u. It can be viewed as a normalization: for a decision maker that is considering improvements of his information before investing, we define his initial position as "not being ready to invest" if he gets no new information.

The NINI assumption expresses that B is such that V is null unless α has some informational content:

NINI: $\mathcal{B}$ is the class of investment sets B such that for

$$V(\underline{\alpha}, u, w, B) = 0 \ \text{ for every } u \in \mathcal{U}_0, w \in \mathbb{R}_+.$$

Now we discuss under what circumstances information is valuable to the agent. First, note that if B doesn't contain feasible elements b such that

$b_k > 0$ in some state k, the agent always weakly prefers to opt out. More generally, an agent fully learning that k is the state of nature cannot take advantage of such information, unless there exists a feasible b offering a gain in state k. We say that the investment set B is *investment prone* if, for every $k \in K$, there exists $b \in B$ such that $b_k > 0$.

What information quality ensures that every investor takes advantage of investment prone sets? We say that α is *sometimes certain* if $q_\alpha^s(k) = 1$ for some k and s, that is, when the true state of nature k is revealed for sure with positive probability. If α is not sometimes certain, we call it *always uncertain*.

The next lemma shows that *sometimes certain* information structures are always advantageous, provided B is feasible and investment prone.

Lemma 1 *If B is investment prone and feasible at wealth level w, then $V(u, w, \alpha, B) > 0$ for every α that is* sometimes certain *and $u \in U_0$.*

We assume that only *sometimes certain* investors are *always inclined to invest*, or SCAI for short. According to SCAI, there exists a feasible and investment prone set of investment opportunities such that the agent weakly prefers to opt out whenever α is *always uncertain*. Note that, for instance, SCAI excludes risk-neutral agents, as such agents can take advantage of every feasible investment prone set B as long as α is not completely uninformative. Although it is again a joint restriction on investment sets and utility

functions, we consider this second assumption mainly as a restriction on the set of potential monetary utility functions u.

The SCAI assumption is expressed as

SCAI: $\mathcal{U}$ consists of the elements u of $\mathcal{U}_0$ such that, there exists a wealth level w and an investment prone set B of feasible investment opportunities such that $V(\alpha, u, w, B) = 0$ for every always uncertain α.

Both NINI and SCAI are expressed jointly in terms of the possible utility functions u and sets B of investment opportunities. We show that these assumptions uniquely define a set $\mathcal{U}$ of utility functions and a class $\mathcal{B}$ of investment sets, and characterize both $\mathcal{U}$ and $\mathcal{B}$. Our characterization relies on two more definitions.

We call an asset $b \in \mathbb{R}^k$ *belief supported* (given initial belief p) if $\sum_k p(k) b_k \leq 0$, and we let B^* be the set of all belief supported assets. An investment set B is belief supported if it contains only belief supported assets ($B \subseteq B^*$), and we let $\mathcal{B}^*$ be the class of belief supported investment sets. Thus, the belief supported assets are the ones that are not preferable to opting out for a weakly risk-averse or risk-neutral agent with belief p. They are also characterized by the absence of arbitrage opportunities (see, e.g., Duffie, 1996, theorem 1 in page 4 and the later discussion in section 1.B. of risk-neutral probabilities).

We call a monetary utility u *ruin averse* whenever $u(0) = -\infty$. A ruin averse agent is thus one who prefers to opt out rather than taking any invest-

ment that leads to ruin with positive probability. Let $\mathcal{U}^* \subset \mathcal{U}_0$ be the set of all ruin averse utility functions in our domain. Following the analysis of Hart (2010), we show in Lemma 8 that ruin aversion is equivalent to $\rho(z) \geq 1$ for every $z > 0$.

The following Theorem characterizes NINI and SCAI. It shows that, although the assumptions are expressed jointly in the pair $\mathcal{U}, \mathcal{B}$, they actually impose separate restrictions on $\mathcal{U}$ and $\mathcal{B}$. Further, $\mathcal{U}$ and $\mathcal{B}$ are uniquely defined by the NINI and SCAI assumptions.

Theorem 1 *$\mathcal{U}$ and $\mathcal{B}$ satisfy NINI and SCAI if and only if $\mathcal{U}$ is the set $\mathcal{U}^*$ of ruin averse utility functions, and $\mathcal{B}$ is the class $\mathcal{B}^*$ of belief supported investment sets.*

4 Entropy as an Ordering of Information for Investment Problems

Next, our attempt is to arrive at an index of informativeness for information structures, an objective way to talk about an information structure being more informative than another, based on the investment framework of the previous section. The result below characterizes entropy as such an index, independent of the utility function of the decision maker, of his wealth and of the specific investment decision considered. In contrast, as we also show, such an index cannot be independent of the decision maker's prior.

4.1 Information Purchasing

In order to understand the value of information for the agent, we consider a situation in which the agent has the possibility to purchase information from information structure α, before making an investment decision in B. Decisions whether to purchase information or not are based on the comparison of the expected payoff under information with the sure payoff $u(w)$. The agent with utility function u and wealth w *purchases information* α at price $\mu < w$ given an investment set B when:

$$\pi(\alpha, u, w - \mu, B) \geq u(w).$$

Otherwise, the agent rejects information α at price μ.

Our information ordering is defined as follows:

Definition 1 *Information structure α investment-uniformly dominates (or investment dominates, for short) the information structure β whenever, for every wealth w and price $\mu < w$ such that α is rejected by all agents with utility $u \in \mathcal{U}^*$ at wealth w for every opportunity set $B \in \mathcal{B}^*$, so is β.*

Lemma 2 *Given an information structure α, price μ and wealth level $w > \mu$, α is rejected by all agents with utility $u \in \mathcal{U}^*$ at wealth level w given every opportunity set $B \in \mathcal{B}^*$ if and only if α is rejected by an agent with $\ln$ utility at wealth w for the opportunity set B^*.*

4.2 Entropy Ordering

Following Shannon (1948), the entropy of a probability distribution $q \in \Delta(K)$ is the quantity:

$$H(q) = -\sum_{k \in K} q(k) \log_2(k)$$

where $0 \log_2(0) = 0$ by convention.[7] The entropy of p is a measure of the level of uncertainty of the investor with belief p on the state of nature. The entropy is always non-negative, and is equal to 0 only in the case of certainty, i.e., when q puts weight 1 on some state k. It is concave, representing that distributions that are closer to the extreme points in $\Delta(K)$ correspond to a lower level of uncertainty.

Recall that following information structure α, the agent's signal is s with probability $p_\alpha(s)$, and that the posterior probability on K following s is q_α^s. The *entropy informativeness* of information structure α is the expected reduction of entropy of the investor's beliefs due to the observation of s. It is the quantity:

$$I(\alpha) = H(p) - \sum_s p_\alpha(s) H(q_\alpha^s).$$

As shown in Section 5.1, $I(\alpha)$ depends on p as well as on α. (For notational simplicity, we drop p from the arguments of I.) The informativeness is minimal when α is $\underline{\alpha}$ with no informational content, and $I(\underline{\alpha}) = 0$. It is maximal

---

[7]The specific function $\log_2(\cdot)$ stems from the uncertainty about a bit of information carrying information or not, a problem with only "two states." For us, we can use any log function, for example $\ln(\cdot)$.

when α fully reveals the state of nature k, and takes the value $H(p)$ in this case. Fixing p, by being a numeric index, note that I provides a complete ordering of information structures.

4.3   Main Result

Our main result establishes that the ordering of information structures of investment dominance is complete, as it coincides with the ordering according to the entropy informativeness.

Theorem 2 *Information structure α investment dominates information structure β if and only if $I(\alpha) \geq I(\beta)$.*

5   Discussion

5.1   Prior-Independent Ordering

For a given prior p, the informativeness ordering presented is represented by the decrease in entropy. An information structure α is more informative than another β if and only if α causes a larger reduction in entropy (from the entropy of the prior to the expected entropy of the posteriors) than does β.

We now want to prove that there can be no index that orders information structures in the same sense, but without using priors. In order to do this, let us define the following.

Definition 2 *We say that an information structure α prior-independent investment dominates β whenever for all priors p α investment dominates β.*

Proposition 1 *There exists no informativeness index that orders information structures according to the prior-independence investment dominance ordering.*

Proof. By Theorem 2 and a fixed prior p the only possible index is $H(p) - \sum_s p_\alpha(s)H(q_\alpha^s)$. Therefore, it suffices to construct an example to show that this index orders two information structures in different ways for two different priors. The example follows.

Let $K = \{1, 2, 3\}$. Let $p_1 = (1/2, 1/2, 0)$ and $p_2 = (1/3, 1/3, 1/3)$ and an agent with $u(x) = \ln(x)$.[8]

Let information structures α_1 and α_2 be described by the two-signal three-states matrices:

$$\alpha_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, \ \alpha_2 = \begin{bmatrix} 0.3 & 0.7 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Clearly, the expected utility for the agent with logarithmic utility under α_1 is larger than that for α_2 when priors are p_1 as the former gives her full information while the latter does not. Thus,

$$H(p_1) - \sum_s p_{\alpha_1}(s)H(q_{\alpha_1}^s) > H(p_1) - \sum_s p_{\alpha_2}(s)H(q_{\alpha_2}^s).$$

---

[8]We make our computations below on the basis of this utility function, but recall that for all x, $\log_2 x = \ln x / \ln 2$, a positive transformation.

What is the expected entropy of the posteriors generated by α_1 and α_2 under p_2? First, the utility for a ln agent of prior p_2 is $\ln(1/3)$. Then for α_1 the expected utility is:

$$\left(\frac{2}{3}\right)\ln\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\ln\left(\frac{1}{3}\right) = -0.63651$$

Therefore,

$$H(p_2) - \sum_s p_{\alpha_1}(s)H(q^s_{\alpha_1}) = \frac{(2/3)\ln 2}{\ln 2} = \frac{0.462\,10}{\ln 2} = 2/3.$$

As for α_2, the (conditional on p_2) probability of either signal is $13/30$ and $17/30$. Upon observing each signal, her posterior is respectively $(3/13, 0, 10/13)$ and $(7/17, 10/17, 0)$. Thus, her expected ln utility from α_2 is:

$$\left(\frac{13}{30}\right)\left(\left(\frac{3}{13}\right)\ln\left(\frac{3}{13}\right) + \left(\frac{10}{13}\right)\ln\left(\frac{10}{13}\right)\right)$$

$$+ \left(\frac{17}{30}\right)\left(\left(\frac{7}{17}\right)\ln\left(\frac{7}{17}\right) + \left(\frac{10}{17}\right)\ln\left(\frac{10}{17}\right)\right) = -0.618$$

Noting that $\ln(1/3)$ is the expected utility from the prior, one gets:

$$H(p_2) - \sum_s p_{\alpha_2}(s)H(q^s_{\alpha_2}) = \frac{0.480\,61}{\ln 2}.$$

That is,

$$H(p_2) - \sum_s p_{\alpha_1}(s)H(q^s_{\alpha_1}) < H(p_2) - \sum_s p_{\alpha_2}(s)H(q^s_{\alpha_2}).$$

Hence, whereas for prior p_1 information structure α_1 is more informative than α_2, the opposite is true for prior p_2. ∎

20

5.2   Uniformity in Wealth

We have defined information dominance as wealth-independent. This is not really a restriction. To see this, consider the following alternative definition:

Definition 3 *Information structure α* investment dominates *the information structure β for wealth w if for all price $\mu < w$ such that α is rejected by all agents with utility $u \in \mathcal{U}^*$ for every opportunity set $B \in \mathcal{B}^*$, so is β.*

With this definition we would have

Theorem 3 *Information structure α investment dominates information structure β for wealth w if and only if $I(\alpha) \geq I(\beta)$.*

The result follows clearly because Lemma 2 works for either definition and the ordering $I(\cdot)$ induced by logarithmic preferences is independent of wealth.

In a similar vein, we have so far made comparisons for agents with the same level of wealth. This can be avoided provided the pricing is done as a proportion of wealth.

Definition 4 *Let $\lambda \in \mathbb{R}^k$, $\lambda < 1$. Information structure α $\lambda-$investment* dominates *the information structure β if for any $\mu = \lambda w$ whenever α is rejected by all agents with utility $u \in \mathcal{U}^*$ for every opportunity set $B \in \mathcal{B}^*$, so is β.*

Lemma 3 *Let $\lambda \in_+$, $\lambda < 1$. Given an information structure α, price $\mu = \lambda w$, α is rejected by all agents with utility $u \in \mathcal{U}^*$ given every opportunity set $B \in \mathcal{B}^*$ if and only if α is rejected by an agent with $\ln$ utility for the opportunity set B^*.*

Once Lemma 3 is established it is then straightforward to show

Theorem 4 *Information structure α $\lambda-$investment dominates the information structure β if and only if $I(\alpha) \geq I(\beta)$.*

5.3   Examples

We now compute some examples. The first one illustrates how our framework serves to complete Blackwell's ordering.

Example 1 *Let $K = \{1, 2, 3\}$ and fix a uniform prior. Consider two information structures that are not ordered in the Blackwell sense. For instance, let each of the two information structures have two signals:*

$$\alpha = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \beta = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \\ 0 & 1 \end{bmatrix}$$

To see that they are not ranked according to Blackwell, we exhibit two decision problems where a decision maker would rank them differently. For instance, in Problem 1 the agent must choose one of two actions, and action 1 gives a utility of 1 only in the first two states, 0 otherwise, while action 2 gives a utility of 1 only in the third state, 0 otherwise. Problem 2, in contrast,

has action 1 pay a utility of 1 only in the first state, 0 otherwise, while action 2 gives a utility of 1 only in states 2 or 3, 0 otherwise. Facing Problem 1, the decision maker would value α_1 more than α_2: following the first signal in α_1 he would choose the first action and following the second signal in α_1 he would choose the second action, securing a utility of 1. This would be strictly greater than his utility after α_2. On the other hand, facing Problem 2, under α_2 he would choose action 1 after the first signal and action 2 after the second, yielding a utility of 29/30, which is greater than his optimal utility after α_1.

But by calculating their entropy reduction from the uniform prior, we know that $I(\alpha_1) > I(\alpha_2)$. Thus, for every investment problem we consider and every utility function in our allowable class, the first information structure is more valuable –more informative– than the second when starting from a uniform prior.

The second example illustrates the optimization over investment sets.

Example 2 *Let prior belief p and investment opportunity b^i be such that for a unit of investment the return is:*

$$b_k^i = \begin{cases} \frac{1}{p(k)} - 1 & \text{if } k = i \\ -1 & \text{if } k \neq i \end{cases}$$

then one can easily see that

$$\sum_{k=1}^{K} p(k) b_k^i = 0$$

23

and hence b^i is belief supported *given initial belief p. Suppose that B is composed of a set of K perfectly scalable investment opportunities b^1 through b^K, one for each state of nature. Then suppose w_i is the amount an investor invests in b^i. We have that*

$$v(\ln, w, B, q) = \sup_{b \in B, \ b \ feasible} \sum_{k=1}^{K} q(k) \ln\left(\frac{w_k}{p(k)}\right)$$

so that the optimal $w_k^* = q(k)w$. Hence

$$v(\ln, w, B, q) = \sum_{k=1}^{K} q(k) \ln\left(\frac{q(k)}{p(k)}\right) + \ln w$$

For other CRRA utility functions with parameter ϕ

$$v(\phi, w, B, q) = \sup_{b \in B, \ b \ feasible} \sum_{k=1}^{K} \frac{q(k)}{1-\phi} \left(\frac{w_k}{p(k)}\right)^{1-\phi}$$

so that the optimal $w_k^* = \lambda q(k)^{1/\phi} p(k)^{-(1-\phi)/\phi} w$ with $\lambda = 1/\sum_{k=1}^{K}\left(q(k)^{1/\phi}p(k)^{-(1-\phi)/\phi}\right)$ Hence

$$v(\phi, w, B, q) = \sum_{k=1}^{K} \frac{q(k)}{1-\phi} \left(\frac{q(k)}{p(k)}\lambda w\right)^{\frac{1-\phi}{\phi}}$$

The next example illustrates how entropy ranks lotteries over information structures compared to their "expectation."

Example 3 *Assume a prior p of K equiprobables states. Then*

$$H(p) = -\sum_{k=1}^{K} \frac{1}{K} \log_2(K) = -\log_2(K)$$

Consider the following information structure:

$$\alpha_1 = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \\ 0 & 1 \end{bmatrix}$$

Given the uniform prior, each signal is equally likely. And the corresponding posteriors are $(2/3, 1/3, 0)$ and $(0, 1/3, 2/3)$. The expected entropy of the posteriors is $\frac{-(2/3)\ln 2 + \ln 3}{\ln 2}$, which is approximately $0.6365/\ln 2$, approx. 0.9182.

On the other hand, consider a situation in which the agent is offered two information structures α_1 and α_2, each with one half probability, as follows:

$$\alpha_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \alpha_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Note how the "average" of these two information structures is the original information structure α.

In α_1, the first signal has probability of 2/3 and the second signal has probability of 1/3. The corresponding posteriors are $(1/2, 1/2, 0)$ and $(0, 0, 1)$. Following α_2, the probability of the first signal is 1/3 and of the second is 2/3, and the corresponding posteriors are $(1, 0, 0)$ and $(0, 1/2, 1/2)$. The corresponding expected entropy is $\frac{(2/3)\ln 2}{\ln 2}$, which is 2/3.

Thus, the experiment consisting of purchasing the lottery of the two information structures is more informative than its expected information structure. More generally, this follows from the convexity of the $\log_2$ function.

Finally, although we have worked with finitely many states to avoid measure theoretic technicalities, the results in this paper are easy to extend to distributions with a continuum of states. This is useful because, many applications assume such a continuum.

25

Example 4 *For example, the normal distribution has a particularly simple entropy. Suppose s is an $n \times 1$ vector of normally distributed variables, where $s \sim N\left(\mu, \Sigma\right).$ Then, letting $|\Sigma|$ be the determinant of the variance-covariance matrix*

$$H\left(s\right) = -\frac{1}{2}\log_2\left(\left(2\pi e\right)^n |\Sigma|\right).$$

When s follows a univariate uniform distribution in the interval $[A, B]$

$$H\left(s\right) = -\int_A^B \frac{1}{B-A}\log_2\left(\frac{1}{B-A}\right) du = \log_2\left(B-A\right)$$

It is clear that information structures can be ordered by the expected reduction on $|\Sigma|$ in the first case and by the expected reduction of $B - A$ in the second.

References

ARROW, K. (1971): "The Value of and Demand for Information," in *Decision and Organization*, ed. by C. McGuire, and R. Radner, pp. 131–139, Amsterdam. North-Holland.

ATHEY, S., AND J. LEVIN (2001): "The value of information in monotone decision problems," Stanford University.

AUMANN, R. J., AND R. SERRANO (2008): "An Economic Index of Riskiness," *Journal of Political Economy*, 116, 810–836.

AZRIELI, Y., AND E. LEHRER (2008): "The value of a stochastic information structure," *Games and Economic Behavior*, 63, 679–693.

BLACKWELL, D. (1953): "Equivalent comparison of experiments," *Annals of Mathematical Statistics*, 24, 265–272.

DUFFIE, D. (1996): *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, N.J.

FOSTER, D. P., AND S. HART (2009): "An Operational Measure of Riskiness," *Journal of Political Economy*, 117, 785–814.

GANUZA, J.-J., AND J. S. PENALVA (2010): "Signal orderings based on dispersion and the supply of private information in auctions," *Econometrica*, 78(3), 1007–1030.

GOSSNER, O., P. HERNÁNDEZ, AND A. NEYMAN (2006): "Oprimal use of communication resources," *Econometrica*, 74(6), 1603–1636.

HART, S. (2010): "Comparing Risks by Acceptance and Rejection," Discussion Paper Series 531, Center for Rationality and Interactive Decision Theory, Hebrew University of Jerusalem.

HERSHEY, D. (2009): *Entropy theory of aging systems: humans, corporations and the universe*. World Scientific, New York, NJ.

JEWITT, I. (2007): "Information order in decision and agency problems," Nuffield College.

KONDOR, P. (2010): "The more we know, the less we agree: Public announcements, higher-order expectations, and rational inattention," European Central University.

LEHMANN, E. L. (1988): "Comparing location experiments," *The Annals of Statistics*, 16(2), 521–533.

MARSCHAK, J. (1959): "Remarks on the Economics of Information," in *Contributions to Scientific Research in Management*, pp. 79–98, University of California, Los Angeles. Western Data Processing Center.

MAS-COLELL, A., M. D. WHINSTON, AND J. GREEN (1995): *Microeconomic Theory*. Oxford University Press.

PENG, L. (2005): "Learning with Information Capacity Constraints," *The Journal of Financial and Quantitative Analysis*, 40(2), 307–329.

SHAKED, M., AND J. G. SHANTHIKUMAR (2007): *Stochastic orders*. Springer.

SHANNON, C. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379–423 ; 623–656.

SIMS, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50(3), 665–690.

———— (2007): "Rational inattention: a research agenda," Princeton University.

VELDKAMP, L. (2011): *Information choice in macroeconomics and finance.* Princeton University Press, Princeton, NJ.

A   Proofs

A.1   Proof of Lemma 1

Let k, s be such that $q_\alpha^s(k) = 1$, and $b \in B$ be such that $b_k > 0$. If b_k is feasible, we have $v(u, w, B, q_\alpha^s) \geq u(w + b_k) > u(w)$. Hence

$$V(\alpha, u, w, B) \geq p_\alpha(s)(u(w + b_k) - u(w)) > 0.$$

A.2   Proof of Theorem 1

Given a class $\mathcal{U} \subseteq U_0$ of utility functions, we say that an investment b individually satisfies NINI if for every $u \in \mathcal{U}$, $w \in \mathbb{R}_+$ such that b is feasible,

$$\sum_k p(k)u(w + b_k) \leq u(w).$$

Thus, b individually satisfies NINI when, under no information, the agent does not prefer b to opting out. We call the set $\tilde{B}$ of investments that individually satisfy the NINI investment set. Since 0_K satisfies NINI, the NINI investment set is a non-empty investment set.

Lemma 4 *Given $\mathcal{U}$, $\mathcal{B}$ satisfies NINI if and only if $\mathcal{B}$ is the class of investment sets contained in $\tilde{B}$.*

Proof. $\mathcal{B}$ satisfies NINI if and only if it contains all the investment sets B such that for every $w > 0$ and $u \in \mathcal{U}$,

$$V(\underline{\alpha}, u, w, B) = 0.$$

I.e., if B is such that for every $w > 0$, $u \in \mathcal{U}$,

$$\sup_{b \in B,\ b \text{ feasible}} \sum_k p(k)u(w + b_k) = 0,$$

which is equivalent to: for every $u \in \mathcal{U}, w > 0$ and $b \in B$ feasible,

$$\sum_k p(k)u(w + b_k) \leq u(w),$$

which is finally equivalent to $B \subseteq \tilde{B}$. $\blacksquare$

From this point on, we thus assume that $\tilde{B}$ is the NINI investment set corresponding to a set of utility functions $\mathcal{U}$, and that $\mathcal{B}$ is the class of investment sets contained in $\tilde{B}$.

We say that a set $A \subseteq \mathbb{R}^K$ is *comprehensive* if, for every feasible b' and $b \in A$ such that $b'_k \leq b_k$ for every k, we also have $b' \in A$.

Lemma 5 *$\tilde{B}$ is comprehensive.*

Proof. Assume that $b \in \tilde{B}$ and b' is such that $b'_k \leq b_k$ for every k. Then, b is feasible at wealth w whenever b' is, and for every $u \in \mathcal{U}, w \in \mathbb{R}_+$,

$$\sum_k p(k)u(w + b'_k) \leq \sum_k p(k)u(w + b_k) \leq u(w).$$

30

Hence, $b' \in \tilde{B}$. ∎

We remark that, if $\tilde{B}$ is not investment prone, neither is any element of $\mathcal{B}$, subset of $\tilde{B}$. In this case, SCAI becomes trivially equivalent to $\mathcal{U} = \emptyset$. The following proposition characterizes SCAI when $\tilde{B}$ is investment prone.

Proposition 2 *If $\tilde{B}$ is investment prone, then $\mathcal{U}$ satisfies SCAI if and only if $\mathcal{U} = \mathcal{U}^*$.*

We divide the proof in a series of Lemmata.

Lemma 6 *Let $u \in \mathcal{U}_0$. If $u(0) > -\infty$, then for every w and B that is investment prone and feasible, there exists α always uncertain such that*

$$V(u, w, \alpha, B) > 0.$$

Proof. Fix u, w, and B that is investment prone and feasible.

For $1 > \varepsilon > 0$ let α^ε be defined by $S_{\alpha^\varepsilon} = K$, $\alpha_k^\varepsilon(s) = 1 - \varepsilon$ if $k = s$, and $\alpha_k(s) = \frac{\varepsilon}{K-1}$ otherwise. It is easily checked that α^ε is *always uncertain* for every $\varepsilon > 0$, and as $\varepsilon \to 0$, $q_{\alpha^\varepsilon}^k(k) \to 1$ for every s.

Since B is investment prone, there exists k^* and $b^* \in B$ such that $b_{k^*}^* > 0$. We now have

$$
\begin{aligned}
v(u, w, B, q_{\alpha^\varepsilon}^{k^*}) &= \sup_{b \in B} \sum_k q_{\alpha^\varepsilon}^{k^*}(k) u(w + b_k) \\
&\geq \sum_k q_{\alpha^\varepsilon}^{k^*}(k) u(w + b_k^*) \\
&\geq q_{\alpha^\varepsilon}^{k^*}(k^*) u(w + b_{k^*}^*) + (1 - q_{\alpha^\varepsilon}^{k^*}(k^*)) u(0).
\end{aligned}
$$

Hence,

$$\lim_{\varepsilon \to 0} v(u, w, B, q_{\alpha^\varepsilon}^{k^*}) = u(w + b_{k^*}^*) > u(w),$$

which shows that for ε small enough, $v(u, w, B, q_{\alpha^\varepsilon}^{k^*}) > 0$ and therefore $V(u, w, \alpha^\varepsilon, B) > 0$. ∎

Lemma 7 *Let $u \in \mathcal{U}_0$ and assume that $\tilde{B}$ is investment prone. If $u(0) = -\infty$, then there exists w and an investment prone set B that is feasible at w such that*

$$V(u, w, \alpha, B) = 0 \text{ for every always uncertain } \alpha.$$

Proof. Since $\tilde{B}$ is investment prone, for every $k \in K$ there exists b'^k such that $b'^k_k > 0$. Let $b^+ = \min_k b'^k_k > 0$, and $b^- = \min(\min_{k \neq j} b'^k_j, -1) < 0$. The investment b^k given by $b^k_k = b^+$ and $b^k_j = b^-$ for every $j \neq k$ is such that for every j, $b^k_j \leq b'^k_j$. Since $\tilde{B}$ is comprehensive from Lemma 5, $b^k \in \tilde{B}$. Let B be the investment prone set $B = \{0_K\} \cup \{b^k, k \in K\}$. B is feasible at wealth level $w = -b^-$. Let α be *always uncertain* and assume $u(0) = -\infty$. For every $s \in S_\alpha$ and $b^k \in B$, the expected utility from investing in b^k conditional on s is

$$
\begin{aligned}
\sum_{k'} q_\alpha^s(k') u(w + b_{k'}^k) &= q_\alpha^s(k) u(w + b^+) + (1 - q_\alpha^s(k)) u(w + b^-) \\
&= q_\alpha^s(k) u(w + b^+) + (1 - q_\alpha^s(k)) u(0) \\
&= -\infty.
\end{aligned}
$$

Thus, for every $s \in S_\alpha$,

$$v(u, w, B, q_\alpha^s) = u(w),$$

32

which implies

$$V(u, w, \alpha, B) = 0.$$

∎

The previous two lemmata provide the proof of Proposition 2.

The next lemma expresses the class $\mathcal{U}^*$ in terms of relative risk aversion:

Lemma 8 *Let $u \in \mathcal{U}_0$, $u \in \mathcal{U}^*$ if and only if for every $z > 0$ $\rho(z) \geq 1$.*

Proof. We follow Hart (2010). Assume that for every $z > 0$,

$$\rho(z) = -\frac{u''(z)}{u'(z)} z \geq 1$$

By integrating between $z < 1$ and 1,

$$\ln u'(z) - \ln u'(0) \geq -\ln(z),$$

which can be rewritten

$$u'(z) \geq \frac{u'(0)}{z}.$$

A second integration between $z < 1$ and 1 shows that,

$$u(z) - u(1) \leq u'(0) \ln(z),$$

hence that $u(0) = \lim_{z \to 0} u(z) = -\infty$.

Now assume that there exists $z_0 > 0$ where $\rho(z_0) < 1$. Since u is IRRA, for every $z \leq z_0$, $\rho(z) \leq \rho(z_0) < 1$. Integrating shows that for every $z \leq z_0$,

$$\ln u'(z) - \ln u'(z_0) \leq -\rho(z_0)(\ln(z) - \ln(z_0))$$

which can be expressed as

$$u'(z) \leq u'(z_0) \left(\frac{z}{z_0} \right)^{-\rho(z_0)}.$$

A second integration between $z < z_0$ and z_0 shows:

$$u(z) - u(z_0) \geq \frac{z_0 u'(z_0)}{1 - \rho(z_0)} \left(\left(\frac{z}{z_0} \right)^{1-\rho(z_0)} - 1 \right).$$

Since $1 - \rho(z_0) > 0$, the limits of the right-hand side, and hence of the left-hand side, are finite. This shows that $u(0) > -\infty$. ∎

Lemma 9 *If* $\mathcal{U} = \mathcal{U}^*$ *the only class* $\tilde{B}$ *that satisfies NINI is the class* $\mathcal{B}^*$ *of belief supported investment sets.*

Proof. We need to show that the NINI investment set $\tilde{B}$ coincides with the set B^* of belief supported assets.

For any $b \in B^*$. For any $u \in \mathcal{U}^*$ and w such that b is feasible, u is concave and increasing implying

$$\sum_k p(k) u(w + b_k) \leq u\left(w + \sum_k p(k) b_k\right) \leq u(w),$$

and hence $b \in \tilde{B}$.

Now consider $b \in \tilde{B}$. Note that u given by $u(z) = \ln(z)$ for $z > 0$ is in $\mathcal{U}^*$. Hence, $b \in \tilde{B}$ implies that for every w large enough,

$$\sum_k p(k) \ln(w + b_k) \leq \ln(w),$$

34

which is equivalent to

$$\sum_k p(k) \ln(1 + \frac{b_k}{w}) \leq 0.$$

Hence, for every $\varepsilon > 0$ small enough,

$$\sum_k p(k) \ln(1 + \varepsilon b_k) \leq 0.$$

A first-order Taylor expansion shows that this implies

$$\sum_k p(k) b_k \leq 0,$$

and hence $b \in B^*$. ∎

To wrap up the proof of Theorem 1, assume first that $\mathcal{U}$ and $\mathcal{B}$ satisfy NINI and SCAI. Then from Lemma 4, $\mathcal{B}$ is the class of investment sets contained in $\tilde{B}$. If $\tilde{B}$ is not investment prone, $\mathcal{U} = \emptyset$, in which case $\tilde{B} = \mathbb{R}^K$, a contradiction. Hence $\tilde{B}$ is investment prone, and from Proposition 2, $\mathcal{U} = \mathcal{U}^*$. It finally follows from Lemma 9 that $\mathcal{B} = \mathcal{B}^*$.

Now, we show that $\mathcal{U}^*$ and $\mathcal{B}^*$ satisfy NINI and SCAI. Assuming $\mathcal{B} = \mathcal{B}^*$, $\mathcal{U}^*$ satisfies SCAI from Proposition 2. With $\mathcal{U} = \mathcal{U}^*$, $\mathcal{B}^*$ satisfies NINI from Lemma 9.

A.3 Proof of Lemma 2

First, note that the *only if* condition is satisfied since the ln utility function belongs to $\mathcal{U}^*$ and $B^* \in \mathcal{B}^*$.

We now prove the *if* part. Assume that α is rejected by an agent with ln utility with wealth w at price μ given the investment set B^*. For $u \in \mathcal{U}^*$, Lemma 8 shows that $\rho(z) \geq 1$ for $z > 0$, hence:

$$\frac{u''(z)}{u'(z)} \leq -\frac{1}{z}.$$

By integration between w and z:

$$\begin{cases} \ln u'(z) - \ln u'(w) & \leq & -\ln(z) + \ln(w) & \text{if } z \geq w \\ \ln u'(z) - \ln u'(w) & \geq & -\ln(z) + \ln(w) & \text{if } z \leq w \end{cases}$$

Fixing w, a second integration with respect to z between w and z' shows that for every z',

$$u(z') - u(w) \leq wu'(w)(\ln(z') - \ln(w))$$

Hence, given any belief q, $B \in \mathcal{B}^*$ and $\mu < w$

$$v(u, w - \mu, B, q) - u(w) \leq wu'(w)(v(\ln, w - \mu, B, q) - \ln(w))$$

and by summation over q_α, for every $B \in \mathcal{B}^*$ and $\mu < w$

$$\pi(\alpha, u, w - \mu, B) - u(w) \leq u'(w)w(\pi(\alpha, \ln, w - \mu, B) - \ln(w))$$

Since $\pi(\alpha, u, w - \mu, B)$ is non-decreasing in B, and B^* is the maximal element of $\mathcal{B}^*$, for every $B \in \mathcal{B}^*$ and $\mu < w$ we have

$$\pi(\alpha, u, w - \mu, B) - u(w) \leq wu'(w)(\pi(\alpha, \ln, w - \mu, B^*) - \ln(w)) < 0,$$

which is the desired conclusion.

A.4 Proof of Theorem 2

Lemma 2 shows that α investment uniformly dominates β if and only if, for every w and $\mu < w$, an agent with ln (or, equivalently, $\log_2$) utility function who rejects α for the opportunity set B^* also rejects β. The following lemma characterizes the value of information for an agent with $\log_2$ utility function and opportunity set B^*.

Lemma 10 *For every $w > 0$ and belief q,*

1.

$$v(\log_2, w, B^*, q) = \log_2(w) - H(q) - \sum_k q(k) \log_2 p(k)$$

2.

$$\pi(\alpha, \log_2, w, B^*) = I(\alpha) + \log_2(w)$$

Proof. For the first point, $v(\log_2, w, B^*, q)$ is the maximum of $\sum_k q(k) \log_2(w + b_k)$ over (b_k) such that $\sum_k p(k) b_k \leq 0$. The first order condition shows that $w + b_k$ is proportional to $\frac{q_k}{p_k}$, hence equal to $w \frac{q_k}{p_k}$. We then obtain

$$
\begin{aligned}
v(\log_2, w, B^*, q) &= \log_2(w) + \sum_k q(k) \log_2 q(k) - \sum_k q(k) \log_2 p(k) \\
&= \log_2(w) - H(q) - \sum_k q(k) \log_2 p(k)
\end{aligned}
$$

For the second point, relying on the previous expression:

$$
\begin{aligned}
\pi(\alpha, \log_2, w, B^*) &= \sum_s p_\alpha(s) v(\log_2, w, B^*, q_\alpha(s)) \\
&= \log_2(w) - \sum_s p_\alpha(s) H(q_\alpha^s) - \sum_{k,s} p_\alpha(s) q_\alpha^s(k) \log_2 p(k) \\
&= I(\alpha) + \log_2(w)
\end{aligned}
$$

since $\sum_s p_\alpha(s) q_\alpha^s(k) = p(k)$. ∎

We now complete the proof of Theorem 2. Relying on Lemma 10, an agent with utility function $\log_2$ rejects α at price $\mu < w$ for the opportunity set B^* if and only if:

$$
I(\alpha) < \log_2\left(\frac{w}{w - \mu}\right)
$$

If $I(\alpha) \geq I(\beta)$, β is rejected whenever α is. If, on the contrary, $I(\alpha) < I(\beta)$, let μ be such that:

$$
I(\alpha) < \log_2\left(\frac{w}{w - \mu}\right) < I(\beta).
$$

At this price μ, α is rejected whereas β is accepted, hence α does not investment dominate β.

A.5   Proof of Lemma 3

The first steps are as in Lemma 2. Now notice that when $\mu = \lambda w$, $\mu = \lambda w$

$C(\pi(\alpha, \ln, w - \lambda w, B^*) - \ln(w))$ is in fact independent of w and since the RHS in the expression

$\pi(\alpha, u, w - \lambda w, B^*) - u(w) \leq C(\pi(\alpha, \ln, w - \lambda w, B^*) - \ln(w))$, is valid for all w the result follows.