A Normalized Value for Information Purchases

Antonio Cabrales, Olivier Gossner and Roberto Serrano*

This version: January 2016

Abstract: Consider agents that are heterogeneous in their preferences and wealth levels. These agents may acquire information prior to choosing an investments holding a property of no-arbitrage, and each piece of information bears a corresponding cost. We associate a numeric index to each information purchase (information-cost pair). This index describes the normalized value of the information purchase: it is the risk aversion level of the unique CARA agent indifferent between accepting and rejecting the purchase, and it is characterized by a "duality" principle that states that agents with stronger preference for information should engage more often in information purchases. No agent more risk averse than the index finds it profitable to acquire the information, whereas all agents less risk averse than the index do. Given an empirically measured range of degrees of risk aversion in an economy with no-arbitrage investments, our model therefore comes close to describing an inverse demand for information, predicting what pieces of information are explored by agents and which ones are left unexplored. Among several desirable properties, the normalized value formula induces a complete ranking of information structures that extends Blackwell's classic ordering.
JEL classification numbers: C00, C43, D00, D80, D81, G00, G11.
Keywords: informativeness, information purchases, free energy, Kullback-Leibler divergence, relative entropy, no-arbitrage investment, Blackwell ordering.

---

*Cabrales: Department of Economics, University College London, a.cabrales@ucl.ac.uk. Gossner: CNRS – École Polytechnique and Department of Mathematics, London School of Economics, olivier.gossner@polytechnique.edu. Serrano: Department of Economics, Brown University, roberto_serrano@brown.edu.
This is a major rewrite of a paper previously circulated under the title "The Appeal of Information Transactions." We thank the editor and three anonymous referees for their extremely helpful feedback. We are grateful to Mihaïlo Backović, Héctor Calvo, Andrew Caplin, Mark Dean, Eddie Dekel, Romain Rancière, Phil Reny, Larry Samuelson, Marco Scarsini, Eytan Sheshinski, Ran Shorrer, and Laura Veldkamp for comments and encouragement. We thank Judith Levi for her excellent editing job.

1 Introduction

We refer to any pair consisting of an information structure and a price for it as an *information purchase*. Such purchases, if they happen, are the manifestation of the demand for information. How many people purchase a piece of information necessarily depends on three components, the quality of that information, the cost of acquiring it, and the agents' primitives given by their wealth and preferences. In settings in which agents can make no-arbitrage investments, the current paper aims at answering the following questions.[1] First, given an information purchase, can its normalized value, capturing the information-price tradeoff, be uniquely characterized?[2] Second, who are the agents willing to go ahead with a given information purchase? We will be able to provide clean answers to such questions, provided that wealth effects are factored out, as explained in the sequel.

In our no-arbitrage settings, we begin by showing that an agent's demand or preference for information is characterized by its degree of risk aversion. Less risk averse agents have a stronger preference for information than do more risk averse agents, in the following sense. If agent u_1 is uniformly less risk averse than agent u_2 , and agent u_2 acquires information, then so does agent u_1, independently of the wealth levels considered.[3] Therefore, agents'

[1]No-arbitrage investments were also used in Cabrales, Gossner, and Serrano (2013). The name refers to the fact that it is a set of investments such that no risk neutral or risk averse agent would be interested in investing in the absence of new information. This is the sense in which no-arbitrage assets constitute a nice framework to measure the value of information.

[2]We add the "normalized" qualification because the index will rely also on the price of the purchase, and not be based only on the information structure.

[3]To be precise, u_1 is uniformly less risk averse than u_2 if, for all wealth levels w_1 and w_2, the coefficient of absolute risk aversion of u_1 at w_1 is not greater than the coefficient of risk aversion of u_2 at w_2. We also say that agent u is uniformly less risk averse than a constant ρ if, independently of w, u's coefficient of absolute risk aversion at w is not greater than ρ. Note that Ross (1981) proposes a related definition: u_1 is strongly less

demand for information is entirely captured by their uniform ranking of risk aversion in settings with no-arbitrage investments.

We seek an "objective" underpinning of normalized values. That is, paralleling the approach of Aumann and Serrano (2008) for ordering riskiness, we pursue a duality logic to define the normalized value of an information purchase.[4] For an information purchase to be considered as more valuable than another one, it must be the case that, whenever an agent is willing to accept the latter, every agent with a stronger preference for information must, a fortiori, accept the former. To be more precise, "u_1 likes information better than u_2" will mean "If u_2 accepts a purchase at some wealth level, then u_1 accepts it at any wealth level" (uniform comparison). For this ordering, we introduce a suitable corresponding ordering of information purchases according to the duality principle described above: "if u_1 likes information better than u_2 and if a_1 is more valuable than a_2, u_1 should accept a_1 if u_2 accepts a_2". We show that this yields a complete ordering of information purchases, which is characterized by our normalized value formula. The normalized value of the purchase turns out to be equal to the risk aversion of the unique CARA (constant absolute risk aversion) agent indifferent between accepting and rejecting it. Such a critical level of risk aversion is expressed as a specific function of relative entropies and the price of the purchase, where the function is increasing with respect to the former and decreasing with respect to the latter.[5]

---

risk averse than u_2 if, $-\frac{u_1''(w_a)}{u_1'(w_b)} \leq -\frac{u_2''(w_a)}{u_2'(w_b)}$ for all w_a, w_b. While both are not unrelated, they are distinct notions.

[4]In Aumann and Serrano (2008), riskiness was conceived as "dual" to risk aversion, while here the value of information is "dual" to preference for information.

[5]In the appendix, we provide two alternative definitions of preference for information: "If u_2 accepts a purchase at some wealth level, then u_1 accepts it at some wealth level" (minimal comparison); and "If u_2 accepts at some wealth level w, then u_1 also accepts at w for every w" (wealth wise comparison). We formulate the corresponding orderings

The fact that a purchase has a higher normalized value does not mean that more agents will accept it. Rather, it is equivalent to the set of agents who unmistakably (i.e., regardless of their initial wealth) would accept it is larger. Thus, for CARA agents, the more valuable a purchase the more of them accept it. But more importantly, given our results connecting preferences for information and risk aversion in our settings, any agent (CARA or not) whose risk aversion always exceeds the normalized value will reject the purchase, while any agent whose risk aversion is everywhere less than the normalized value will accept it. The rest of agents, for whom for some initial wealth their risk aversion is below the normalized value, while for others it is above it, may accept or reject the purchase, i.e., their decision is "subject to wealth effects." The latter set is the only impediment that prevents the set of agents in the economy who accept a purchase from being monotonically increasing in its normalized value. Having said that, if we assume that agents are sufficiently heterogeneous in their wealth, and that the degree of risk aversion of agents in the economy is known, the normalized value of information purchases is a useful tool that delivers a stark prediction of what pieces of information will be explored or remain unexplored in the economy, provided that no-arbitrage assets comprise the available set of investments for the agents.

Our normalized value measure provides some interesting insights on the demand for information. For instance, a decrease by a certain percentage of the cost of information translates into an increase by the same amount of the normalized value of the corresponding information purchase. This means that, whenever the price of information drops by half, agents who are

based on the duality principle. Strikingly, all three orderings of information purchases coincide: all three are represented by our normalized value formula. As also shown in the appendix, another characterization of the normalized value is expressed in terms of the group of agents who are willing to accept a given information purchase. This parallels the work of Hart (2011) who provides this comparison of orderings for indices of riskiness.

twice more risk averse become willing to acquire some piece of information. Another insight is obtained by examining the least and most valuable purchases. Quite intuitively, the least valuable purchases are the ones associated to the lesser informational content. Due to our no-arbitrage assumption, the most valuable purchases do not consist only of the ones always allowing to learn the true fundamental state, but more generally all those of which that always allow to exclude one fundamental state from the possible ones that will realize. More generally, our ordering of information purchases induces a completion of Blackwell's classic ordering of information structures.

Although we include a section on related literature, it is useful to close the introduction by comparing our index to a couple of closely related contributions. Our earlier work in Cabrales, Gossner, and Serrano (2013) provides a complete ordering of information structures. In that paper, the informativeness of an information structure is characterized by the reduction of entropy from the prior. With a uniform prior, and for small amounts of information, that index is close to the index proposed here when the purchase price is kept constant, but they differ significantly when the amount of information in the signals is larger; see Subsection 6.5 for details. Our two papers have different methodologies and results. In the current paper, we are interested in the demand for information, so we allow wealth levels to differ across agents, and we compare the sets of agents who are willing to accept a purchase with the sets of agents who are willing to accept another one. In Cabrales, Gossner, and Serrano (2013), we were asking whether the maximal price any agent in the economy is willing to pay for one piece of information is larger than the maximal for the other piece of information, while all agents are held at the same wealth level.

Another related paper is Kelly (1956), which studies an environment where an individual receives information on a set of betting opportunities about events that happen in exclusive and independent states of nature, and that are priced with ex-ante fair odds. In other words, the bettor faces a decision environment that is composed of no-arbitrage assets, like our decision-makers. Kelly (1956) proves that the betting strategy that maximizes long-run wealth is the one that maximizes instantaneous expected logarithmic utility. As a result, the value of information in that environment (which he calls *Gmax*, p. 922) is equivalent to the one we propose in Cabrales, Gossner, and Serrano (2013), which is unsurprising given the important role the agent with logarithmic utility plays in that paper. Notice, though, that the motivation and nature of the results are starkly different between those two papers and even more different from the current one.[6]

The paper is organized as follows. Section 2 describes the model. Section 3 relates the value of information and uniform risk aversion. Section 4 introduces the "uniformly more valuable" ordering, the normalized value formula, and establishes our main result. Section 5 presents our results connecting the normalized value of an information purchase to the economy's levels of risk aversion. Section 6 goes over a number of properties of the normalized value and presents several examples. Section 7 is devoted to related literature, and Section 8 concludes. Some of the more technical proofs and additional justifications and properties of the index are collected in an Appendix.

[6]Samuelson (1969) already discusses Kelly (1956) and the properties of logarithmic utility investing, and Blume and Easley (2002) the potential for long-run dominance in a market of Kelly investors. A good summary on "Kelly investing" is MacLean, Thorp, and Ziemba (2011). Cesa-Bianchi and Lugosi (2006) note the formal equivalence between sequential gambling and forecasting under the logarithmic loss function.

2 The Model

We consider an investor, who, prior to making an investment decision, may acquire some information at a cost. In this section we define the conditions under which the agent accepts acquiring this piece of information.

2.1 Utility for wealth

We consider an investor with initial wealth w and a monetary utility function u defined on $\mathbb{R}$. We assume that u is non-decreasing, strictly concave[7], and twice differentiable. We let $\mathcal{U}$ be the set of such monetary utility functions. We identify agents by their monetary utility functions, thus speaking of agent u to refer to an agent with utility function u.

Given $u \in \mathcal{U}$ and $w \in \mathbb{R}$, let $\rho_u(w) = -\frac{u''(w)}{u'(w)}$ be the Arrow-Pratt coefficient of absolute risk aversion of agent u at wealth w. We also let $\overline{R}(u) = \sup_w \rho_u(w)$, and $\underline{R}(u) = \inf_w \rho_u(w)$. We say that agent u_1 is uniformly more risk averse than agent u_2 whenever $\underline{R}(u_1) \geq \overline{R}(u_2)$. It is often necessary to assume that u has decreasing absolute risk aversion. We thus let $\mathcal{U}_{DA}$ be the subset of $\mathcal{U}$ such that ρ_u is non-increasing.

2.2 Investments

There is a finite set K of states of nature, about which the agent is uncertain. The agent's prior on K is $p \in \Delta(K)$, assumed to have full support. The set of investment opportunities consists of all no-arbitrage assets given p, that is, assets with a non-positive expected return: $B^* = \{x \in \mathbb{R}^K, \sum_k p_k x_k \leq 0\}$.[8]

[7] In our framework, risk neutral agents would sometimes make unbounded optimal investments, which creates technical problems while adding little to content. We hence exclude them from the analysis.

[8] The vector $p = (p_1, \dots, p_k, \dots, p_K)$ in the definition of no-arbitrage assets corresponds to the price vector of Arrow securities, where p_k can be interpreted as the price of an asset

When the agent's initial wealth is w, $x \in B^*$ is chosen and state k is realized, the agent's final wealth is $w + x_k$.

Two features of the set B^* of available investments are worth emphasizing. First, $0 \in B^*$, that is, not investing at all is feasible. Second, no agent in our class would prefer to invest over the zero investment in the absence of new information, while this may change after such an arrival. In this sense, these "no-arbitrage" assets provide a useful framework to measure the value of new information. We remark that, although investments are unbounded in each state, in many cases, particularly when the information does not lead to major variations in the prior, one could restrict attention to a bounded set of investments. In order to obtain our characterizations, however, it will be convenient to allow the largest set of no-arbitrage assets, as written.

2.3 Information Purchases

Before choosing an investment, the agent has the opportunity to engage in a (costly) information purchase $a = (\mu, \alpha)$. Here, $\mu > 0$ represents the cost of the information purchase, and α is the information structure representing the information obtained from a. That is, α is given by a finite set of signals S_α, together with 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 practice 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 $(\alpha_k(s))_{s \in S_\alpha}$. Signal s has a total probability $p_\alpha(s) = \sum_k p_k \alpha_k(s)$, and we assume, without loss of generality, that $p_\alpha(s) > 0$ for every s. For each signal $s \in S_\alpha$, we let $q_k^s \in \Delta(K)$ be the

that pays 1 in state k and 0 in all other states. The fact that this vector coincides with the agent's prior p means that no-arbitrage assets cannot yield a positive expected return. We disentangle the two roles of p, price and priors, in Subsection D.3 in the appendix.

probability of state k conditional on s computed using Bayes' rule.

We say that a is *excluding* if for every signal s, there exists k such that $q_k^s = 0$. It is *nonexcluding* otherwise. Excluding information purchases are such that, for every received signal, there exists a state of nature that the agent can exclude.

2.4 Optimal Investment after Receiving Information

Given a belief q, an agent with wealth w and utility u chooses $x \in B^*$ in order to maximize his expected utility over all states $k \in K$. The maximum expected utility is then $V(u, w, q)$, given by:

$$V(u, w, q) = \sup_{x \in B^*} \sum_k q_k u(w + x_k).$$

2.5 Acceptance of Information Purchases

The agent with utility function u and wealth w *accepts* an information purchase $a = (\mu, \alpha)$ if and only if paying μ upfront to receive information according to α generates an expected utility greater than or equal to staying with wealth w. This is the case if and only if:

$$\sum_s p_\alpha(s) V(u, w - \mu, q_k^s) \geq u(w).$$

3 Risk Aversion and Preference for Information

In order to arrive at the concept of normalized value of information purchases, it is useful to first understand what characteristics of an agent's utility function make his demand for information increase or decrease. As it turns out, in our setting, an agent's preference for information is determined by his uniform risk aversion.

Theorem 1 *Given $u_1, u_2 \in \mathcal{U}$, the following two conditions are equivalent:*

1. *u_1 is uniformly more risk averse than u_2*
2. *u_2* uniformly likes information better *than u_1, i.e., for every pair of wealth levels w_1, w_2, and information purchase a, whenever agent u_1 accepts a at wealth w_1, then so does agent u_2 at wealth w_2.*

Proof. For the first direction, assume that $\overline{R}(u_2) \leq \underline{R}(u_1)$. For every z, w_1, and w_2, we have

$$\frac{u_1''(w_1+z)}{u_1'(w_1+z)} \leq \frac{u_2''(w_2+z)}{u_2'(w_2+z)}.$$

By integration on z, we have:

$$\begin{cases} \ln u_1'(w_1+z) - \ln u_1'(w_1) & \leq & \ln u_2'(w_2+z) - \ln u_2'(w_2) & \text{if } z \geq 0; \\ \ln u_1'(w_1+z) - \ln u_1'(w_1) & \geq & \ln u_2'(w_2+z) - \ln u_2'(w_2) & \text{if } z \leq 0; \end{cases}$$

which is the same as:

$$\begin{cases} \frac{u_1'(w_1+z)}{u_1'(w_1)} & \leq & \frac{u_2'(w_2+z)}{u_2'(w_2)} & \text{if } z \geq 0; \\ \frac{u_1'(w_1+z)}{u_1'(w_1)} & \geq & \frac{u_2'(w_2+z)}{u_2'(w_2)} & \text{if } z \leq 0. \end{cases}$$

By a second integration on z, for every z:

$$\frac{u_1(w_1+z) - u_1(w_1)}{u_1'(w_1)} \leq \frac{u_2(w_2+z) - u_2(w_2)}{u_2'(w_2)}. \tag{1}$$

Thus, for every $q \in \Delta(K)$ and $\mu \geq 0$:

$$\frac{V(u_1, w_1-\mu, q) - u_1(w_1)}{u_1'(w_1)} \leq \frac{V(u_2, w_2-\mu, q) - u_2(w_2)}{u_2'(w_2)}.$$

And finally, for every information structure α,

$$\frac{\sum_s p_\alpha(s) V(u_1, w_1-\mu, q_k^s) - u_1(w_1)}{u_1'(w_1)} \leq \frac{\sum_s p_\alpha(s) V(u_2, w_2-\mu, q_k^s) - u_2(w_2)}{u_2'(w_2)}.$$

This implies that for every w_1, w_2, if u_1 accepts $a = (\mu, \alpha)$ at wealth w_1, then u_2 also accepts it at wealth w_2.

The converse relies on Lemma 7, which itself relies on Lemmas 4, 5, and 6, all found in the appendix.

Assume that u_2 uniformly likes information better than u_1, that is, for any two wealth levels w_1, w_2, if u_1 accepts an information purchase at w_1, then u_2 accepts this information purchase at w_2. To prove that u_1 is more risk averse at w_1 than u_2 is at w_2, which is a local property at w_1, w_2, the proof, provided in the appendix relies on information structures $\alpha(\varepsilon)$, which are "little informative", hence induce small investments. Lemma 7 in the appendix characterizes the amount that an agent is willing to pay for "small information", and we obtain in our case that for every w_1, w_2 and for a small enough $\varepsilon > 0$,

$$\frac{p_k + p_l}{2\rho_{u_2}(w_2)p_k p_l}\varepsilon^2 \geq \frac{p_k + p_l}{2\rho_{u_1}(w_1)p_k p_l}\varepsilon^2.$$

Hence, $\rho_{u_2}(w_2) \leq \rho_{u_1}(w_1)$, which implies $\overline{R}(u_2) \leq \underline{R}(u_1)$. ■

Theorem 1 establishes the connection between preference for information and risk aversion. Lemma 2 in Cabrales, Gossner, and Serrano (2013) shows that an agent with ln utility accepts an information purchase whenever a more risk-averse agent does. Theorem 1 both extends this result to general pairs of utility functions, and shows that a converse result holds, namely, that an agent whose preference for information is higher than another is necessarily less risk-averse. The proof of this converse part is somewhat more involved than for the direct part, as one needs to derive a conclusion about the risk aversion levels of the agents at all wealth levels.

To understand intuitively this result, think of the following: let $u_1(w) = u_2(w) = 0$ and let $u_1'(w) = u_2'(w) = 1$ (a normalization that is consistent with VNM utilities). Then, observe from equation 1 that the fact that u_1 is uniformly less risk averse than u_2 (as $\overline{R}(u_1) \leq \underline{R}(u_2)$ entails) is equivalent to the nestedness of the two utility functions: that is, one function is uniformly above the other for any value of final wealth higher or lower than w. Or more formally, $u_1(w + x_k) > u_2(w + x_k)$ for any $x_k \neq 0$. That makes it clear that, for a given prospect of an investment portfolio, the agent with the less concave utility function values information more, as he gains more from the added utility that information entails.

In closing the section, we remark that the identification of preference for information and risk aversion demonstrated in Theorem 1 is specific to our no-arbitrage settings. Indeed, such an identification may not hold were investments violate this assumption. For example, suppose an agent can invest at most \$1 in an asset that pays \$10 with probability 0.9 and \$0 with probability 0.1. The investor could purchase information at a price $\mu = 0.5$ and learn the state for sure before investing. Note how a risk neutral (or approximately risk neutral) agent would not purchase the information, while some risk averse agents would (e.g., $u(x) = \sqrt{x+1}$). Thus, an agent that is uniformly more risk averse exhibits in this case a stronger preference for information, but notice how the proposed investment violates no-arbitrage, as the prior-evaluated expected payoff is positive.[9] The reader can check how, in this example, our conclusion would hold again if one restores the no-arbitrage assumption, for instance by increasing the loss in the bad state.

4 Preference for Information and a Value for Information Purchases

This section proposes an "objective" way to define the normalized value of information purchases. The approach is based on ordering preferences for

[9]We thank an anonymous referee for this example.

information. We offer three variants of the same idea, all of them leading to the same normalized value index. We present one here and relegate the other two to the appendix, which also includes an additional approach based on total rejections/acceptances.

Our first task is to define what it means for an agent to like information better than another one. In general terms, an agent u_1 likes information better than another agent u_2 when u_1 accepts information more often than u_2. In order to make the concept precise, we need to be careful about the wealth levels at which we compare u_1 and u_2 acceptance and rejection of information purchases.

Our first version of this concept is the uniform preference for information, already used in Theorem 1. It requires agent u_1 to accept the information purchase at all wealth levels whenever u_2 accepts it at some wealth level. That is, agent u_1 uniformly likes information more than agent u_2 means that, whenever agent u_2 is interested in purchasing information, for sure so is agent u_1. Alternative definitions of the concept are provided in the appendix.

We move now to define the comparative normalized value of two information purchases based on the rankings over preferences for information. The definition formalizes the natural idea that if an information purchase is accepted by a first agent, then any purchase that is deemed more valuable should a fortiori be accepted by an agent who likes information better than the first.[10]

Definition 1 *Let a_1, a_2 be two information purchases. We say that a_1 is* uniformly more valuable *than a_2 if, given two agents $u_1, u_2 \in \mathcal{U}$ such that u_1 uniformly likes information better than u_2 and given two wealth levels w_1, w_2, whenever agent u_2 accepts a_2 at wealth level w_2, then agent u_1 accepts a_1 at*

[10] This principle was referred to as "duality" in Aumann and Serrano (2008).

wealth level w_1.

4.1 Normalized Value of Information Purchases

For two probability distributions p and q, the *relative entropy* from p to q, also called their *Kulback-Leibler divergence*, has been proposed as a non-symmetric measure of their discrepancy. It is defined as follows:

$$d(p\|q) = \sum_k p_k \ln \frac{p_k}{q_k}.$$

It is always non-negative, and equals zero if and only if $p = q$. It is finite provided the support of q contains that of p, and we let it take the value $+\infty$ otherwise. Thus, p and q are "maximally different" when q rules out one possibility that p does not.[11]

Based on the relative entropy, we define the *normalized value* of an information purchase a as this quantity:

$$\mathcal{NV}(a) = -\frac{1}{\mu} \ln \left(\sum_s p_\alpha(s) \exp(-d(p\|q_\alpha^s)) \right). \tag{2}$$

In the above formula, and throughout the paper, we use the convention $\exp(-d(p\|q_\alpha^s)) = 0$ by continuity if $d(p\|q_\alpha^s) = +\infty$. The normalized value $\mathcal{NV}(a)$ of a is thus well-defined and finite if and only if there exists s such that $-d(p\|q_\alpha^s)$ is finite, which is the case if a is nonexcluding. We let $\mathcal{NV}(a) = +\infty$ by continuity if a is excluding.

While at this point the normalized value formula may appear somewhat mysterious, we solve the "mystery" right away, in the following subsection. Suffice it for now to say that the normalized value is also equivalent to the

[11] If p were the true distribution and q an approximate hypothesis, information theory views the relative entropy from p to q as giving the expected number of extra bits required to code the information if one were to use q instead of p. See, e.g., Kraft (1949) and McMillan (1956), or Kelly (1956) for a betting market interpretation.

level of risk aversion of the unique CARA individual indifferent between accepting and rejecting the purchase. Once one figures out that agent's optimal investment, the written formula is found.

The normalized value of an information purchase decreases with its price and increases with the relative entropy of the prior to the posteriors. Specifically, the normalized value of an information purchase is measured by the inverse of its price multiplied by the natural logarithm of the expected exponentials of the negative of relative entropy from the prior to each of the generated posteriors. We remark, though, that simply taking an average of relative entropies divided by price yields a different ordering, as detailed in Subsection 6.4.[12]

4.2 The Main Result

Theorem 2 *Let a_1, a_2 be two information purchases. The following two statements are equivalent:*

1. *a_1 is uniformly more valuable than a_2.*

2. $$\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2).$$

Proof. Recall the class of CARA (constant absolute risk aversion) utility functions. Given $r > 0$, let u_C^r be the CARA utility function with parameter r given by $u_C^r(w) = -\exp(-rw)$ for every w. For a CARA agent with coefficient r and wealth level w, we consider the problem of optimal portfolio

[12] If we ignore the price μ, the rest of the expression in the normalized value formula is, remarkably, referred to as "free energy" in theoretical physics (see, e.g., Landau and Lifshitz, 1980), where relative entropy plays the role of the Hamiltonian of the system. A similar formula appears under the term "stochastic complexity" in machine learning (Hinton and Zemel, 1994).

choice when the agent's belief is q. The next lemma shows that the solution is interior when q has full support.

Lemma 1 *Let $q \in \Delta(K)$ have full support. The optimal portfolio for the CARA agent with risk-aversion coefficient r and belief q is independent of w, and is given by*

$$x_k = -\frac{1}{r}(-d(p\|q) + \ln \frac{p_k}{q_k}).$$

Proof. The agent's objective is to maximize

$$\sum_k q_k \exp(-r(w + x_k)),$$

subject to the constraint $\sum_k p_k x_k = 0$. The first-order condition shows that

$$q_k \exp(-rx_k) = \lambda p_k,$$

where λ is independent of k. We then have, for every k,

$$-rx_k = \ln \lambda + \ln \frac{p_k}{q_k}.$$

Summing over these expressions, after we multiply each of them by p_k, gives

$$0 = \ln(\lambda) + d(p\|q),$$

and hence, the result. ■

We continue the proof with Lemma 2, which shows that $\mathcal{NV}(a)$ can be equivalently defined as the level of risk aversion of a CARA agent that is indifferent between accepting and rejecting the purchase.

Lemma 2 *Let a be an information purchase and w be any wealth level.*

1. *If $r > \mathcal{NV}(a)$, then an agent with utility u_C^r rejects a at wealth w.*

2. If $r \leq \mathcal{NV}(a)$, then an agent with utility u_C^r accepts a at wealth w.

Proof. The agent accepts a if and only if

$$\sum_s p_\alpha(s) V(u_C^r, w-\mu, q_\alpha^s) \geq u_C^r(w).$$

If a is excluding, then the left-hand side of the inequality equals 1, and the inequality is satisfied for all r and w. If a is nonexcluding, then the agent accepts a if and only if

$$\exp(-rw) \geq \exp(-r(w-\mu)) \sum_s p_\alpha(s) \exp(-d(p \| q_\alpha^s)).$$

This is equivalent to

$$\exp(-r\mu) \geq \sum_s p_\alpha(s) \exp(-d(p \| q_\alpha^s)),$$

which in turn is equivalent to $r \leq \mathcal{NV}(a)$. Thus, for $r \leq \mathcal{NV}(a)$, the agent accepts a at every wealth level, whereas for $r > \mathcal{NV}(a)$, the agent rejects a at every wealth level. ■

Equipped with Theorem 1 and Lemma 2, we can now proceed to prove Theorem 2.

First assume that a_1 is uniformly more valuable than a_2, and that $\mathcal{NV}(a_2)$ is finite. By Lemma 2, a CARA agent with a coefficient of risk aversion $\mathcal{NV}(a_2)$ accepts a_2 at every wealth level. This agent uniformly likes information better than itself since, by Lemma 2, acceptance or rejection for CARA agents is independent of wealth. Since a_1 is more valuable than a_2, this CARA agent also accepts a_1 at every wealth level, which implies (also by Lemma 2) that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$.

The case in which $\mathcal{NV}(a_2) = \infty$ is dealt with similarly: by Lemma 2 every CARA agent accepts a_2 at every wealth level, which implies that the same agent also accepts a_1 at every wealth level. By Lemma 2 again, this implies that we also have $\mathcal{NV}(a_1) = \infty$.

Now assume that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$. Consider two agents u_1 and u_2 such that u_1 uniformly likes information better than u_2. Given wealth levels w_1 and w_2, and assuming that u_2 accepts a_2 at w_2, we need to prove that u_1 accepts a_1 at w_1. By Theorem 1 we have $\overline{R}(u_1) \leq \underline{R}(u_2)$. Since $\overline{R}(u_1) > 0$ and $\underline{R}(u_2) < \infty$, $\underline{R}(u_2)$ is positive and finite. Let $r = \underline{R}(u_2)$. Since $\overline{R}(u_r^C) = r$, the agent u_r^C likes information better than agent u_2 does, by Theorem 1; hence the former accepts a_2 at any wealth level. By Lemma 2 this means that $r \leq \mathcal{NV}(a_2)$, and hence also $r \leq \mathcal{NV}(a_1)$, so that u_r^C also accepts a_1 at any wealth level. Since $\overline{R}(u_1) \leq r = \underline{R}(u_r^C)$ and u_1 likes information better than u_r^C (also by Theorem 1), it follows that u_1 accepts a_1 at wealth level w_1. $\blacksquare$

5 Results on Demand for Information

Our next results show that, in our settings, the normalized value of an information purchase adequately characterizes the demand for information, namely the set of agents who are willing to go ahead with any given information purchase. The first result in this section shows that if the minimum coefficient of absolute risk aversion of an agent over all levels of wealth is bigger than the normalized value of information, he rejects a purchase independently of his wealth. If on the other hand the maximum coefficient of absolute risk aversion of an agent over all levels of wealth is smaller than the normalized value of information, he accepts a purchase independently of his wealth.

Theorem 3 *Consider an information purchase a and $u \in \mathcal{U}$.*

1. If $\underline{R}(u) > \mathcal{NV}(a)$, then agent u rejects a at all wealth levels w.

2. If $\overline{R}(u) \leq \mathcal{NV}(a)$, then agent u accepts a at all wealth levels w.

Proof. We begin by characterizing the function $V(u_C^r, w, q)$ for CARA agents. This is done in the lemma below. We recall the convention that $\exp(\ -d(p\|q) - rw) = 0$ by continuity if $d(p\|q) = \infty$, and state:

Lemma 3 *For every r, w, q:*

$$V(u_C^r, w, q) = -\exp(\ -d(p\|q) - rw).$$

Proof. First, assume that q has full support; hence, $d(p\|q)$ is finite. Using the optimal-portfolio characterization of Lemma 1, we obtain:

$$\begin{aligned} V(u_C^r, w, q) &= \sum_k q_k \exp(-r(w + x_k)) \\ &= \exp(-rw) \sum_k q_k \exp(-d(p\|q) + \ln \frac{p_k}{q_k}) \\ &= \exp(-rw - d(p\|q)) \sum_k q_k \frac{p_k}{q_k} \\ &= \exp(-rw - d(p\|q)). \end{aligned}$$

Now assume that $q_{k_0} = 0$ for some k_0; hence, $d(p\|q) = +\infty$. The investment x^0 given by :

$$\begin{cases} x_{k_0}^0 = -\frac{1-p_{k_0}}{p_{k_0}}; & \\ x_k = 1 & \text{if } k \neq k_0 \end{cases}$$

is such that $\lambda x^0 \in B^*$ for every $\lambda \geq 0$. For every such λ, we have

$$\begin{aligned} V(u_C^r, w, q) &\geq \sum_k q_k u_C^r(w + \lambda x_k^0) \\ &= u_C^r(w + \lambda x_k^0) \\ &= \exp(-r(w + \lambda)). \end{aligned}$$

Since $\lim_{\lambda\to\infty} \exp(-r(w+\lambda)) = 0$, we have $V(u_C^r, w, q) \geq 0$. On the other hand, $V(u_C^r, w, q) \leq \sup_z u_C^r(z) = 0$. The desired conclusion is therefore that $V(u_C^r, w, q) = 0$. ■

We now proceed to complete the proof of Theorem 3.

A CARA agent with risk aversion r accepts a if and only if

$$\sum_s p_\alpha(s) V(u_C^r, w-\mu, q_\alpha^s) \geq u_C^r(w).$$

If a is excluding, then the left-hand side of the inequality equals 0, and the inequality is satisfied for all r and w. If a is nonexcluding, then the agent accepts a if and only if

$$\exp(-rw) \geq \exp(-r(w-\mu)) \sum_s p_\alpha(s) \exp(-d(p\|q_\alpha^s)).$$

This is equivalent to

$$\exp(-r\mu) \geq \sum_s p_\alpha(s) \exp(-d(p\|q_\alpha^s)),$$

which in turn is equivalent to $r \leq \mathcal{NV}(a)$. Thus, for $r \leq \mathcal{NV}(a)$, the agent accepts a at every wealth level, whereas for $r > \mathcal{NV}(a)$, the agent rejects a at every wealth level. The theorem now follows immediately from part 1 of Theorem 1 by noting that an agent with $\overline{R}(u) \leq r$ is uniformly less risk averse than a CARA agent with risk aversion r, and an agent with $\underline{R}(u) \geq r$ is uniformly more risk averse than a CARA agent with risk aversion r. ■

Parts (1) and (2) of Theorem 3 characterize situations in our settings in which one can unequivocally say whether u accepts a or not independently of what one knows about the agent's wealth level. Whenever $\overline{R}(u) > \mathcal{NV}(a) \geq \underline{R}(u)$, it may be the case that agent u accepts a for some wealth levels, and rejects it for other wealth levels. This observation makes it clear why the normalized value index $\mathcal{NV}(a)$ is not a universal representation of preferences for information purchases. But importantly, the reason why that is true is the presence of such wealth effects.

Another way to look at this result is the following. Imagine that it has been estimated econometrically that the agents in this economy have $\rho_u(w) \in [\gamma_1, \gamma_2]$ for all relevant w. Then, given an information structure α one can identify two prices, call them μ_1 and μ_2 as follows:

$$\gamma_1 = \mathcal{NV}(\mu_{1,}\alpha), \gamma_2 = \mathcal{NV}(\mu_{2,}\alpha)$$

where μ_1 and μ_2 offer the following interpretation: under our assumption of no-arbitrage investments, for prices $\mu > \mu_1$, the information purchase (μ, α) will be unanimously rejected, where as for prices $\mu < \mu_2$ the purchase will be unanimously accepted. This is the sense in which for all information structures, we can identify, thanks to the index of normalized value, the minimum and maximum prices for individuals within a group whose $\rho_u(w)$ are known or have been estimated, provided that new information can potentially be used in investments that have the no-arbitrage property.

More can be said when u is DARA, providing a characterization of the utility functions that exhibit unanimous acceptance and unanimous rejection of a purchase, as we show next:

Theorem 4 *Consider an information purchase a and the class of utility functions $\mathcal{U}_{DA}$.*

1. *An agent $u \in \mathcal{U}_{DA}$ rejects a at all wealth levels if and only if $\underline{R}(u) > \mathcal{NV}(a)$.*

2. *An agent $u \in \mathcal{U}_{DA}$ accepts a at all wealth levels if and only if $\overline{R}(u) \leq \mathcal{NV}(a)$.*

The proof of this result is in the appendix.

5.1 Examples and Calibrations of the Model

This subsection illustrates the results derived for the demand for information. It presents some calibrations of the model in order to get a feel for the magnitudes implied by the index. Of course, this is meant to be only suggestive, far from providing a careful empirical analysis.

According to Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2011) "Lottery responses and wealth information imply a distribution of CRRA coefficients mainly between 1 and 10,". The lowest quartile of the wealth distribution in most developed countries has zero or negative net worth (Sierminska, Brandolini, and Smeeding, 2006) and the wealth of the highest decile is between 0.36 (Italy) to 1.81 (Germany) million US$ (with the US being around 0.95). This, though, includes very young people who have not had time to acquire any assets. If we use the median wealth, instead, the figures go from about 20,000 $ (Sweden) to about 120,000 $ (Italy), with the US being about 50,000 $. This means that a large fraction of the population in the developed world can be characterized with $\overline{R}(u) = 5 \times 10^{-4}$ and $\underline{R}(u) = 1.8 \times 10^{-6}$.

Example 1 *Recall our maintained assumption that agents can make potentially large investments after receiving information. Let a be an information purchase about a binary state of the world (e.g., the US will be in recession in 2020 or not) where the two states are equally likely a priori. The information structure α consists of two signals. Conditional on* 1 *being received, the probability of a recession is β, and conditional on signal* 2 *arriving, the* *probability of recession is* $1-\beta$. *Then,*

$$\begin{aligned}
\mathcal{NV}(a) &= -\frac{1}{\mu}\ln\left(\frac{1}{2}\exp\left(-\left(\frac{1}{2}\ln\frac{1}{2\beta}+\frac{1}{2}\ln\frac{1}{2\,(1-\beta)}\right)\right)\right. \\
&+ \left.\frac{1}{2}\exp\left(-\left(\frac{1}{2}\ln\frac{1}{2\,(1-\beta)}+\frac{1}{2}\ln\frac{1}{2\beta}\right)\right)\right) \\
&= -\frac{1}{\mu}\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right).
\end{aligned}$$

The information purchase $a=(\mu,\alpha)$ *is accepted by the agents considered if its price* μ *satisfies*

$$5\times 10^{-4}\le -\frac{1}{\mu}\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right)$$

which is equivalent to

$$\mu\le -\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right)\times 2\times 10^{3}.$$

The same information purchase is rejected by all agents considered if

$$1.8\times 10^{-6} > -\frac{1}{\mu}\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right),$$

which is equivalent to

$$\mu > -\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right)\times 5.5\times 10^{5}.$$

For a numerical application, let $\beta=0.8$

$$-\ln\left(2\left(\beta\,(1-\beta)\right)^{1/2}\right)=-\ln\left(0.8\right)\simeq 0.223$$

then a *will be accepted by all those agents if*

$$\mu\le 450\,\$$$

and it will be rejected by all if

$$\mu > 123,000\,\$$$

Now let $\beta = 0.55$*,*

$$-\ln\left(2\left((0.55)^{1/2}(0.45)^{1/2}\right)\right) \simeq 5\times 10^{-3}$$

then a will be accepted by all those agents if

$$\mu \leq 10\,\$$$

and it will be rejected by all if

$$\mu > 2,750\,\$$$

Of course, both ranges in the previous example are relatively large, as they separate the case in which a large portion of the world population would accept a purchase from the case in which only a few people might. Note however that these ranges are realistic figures in the sense that it is not hard to think of pieces of information with a higher price than the maximum bound, or with a lower price than the minimal bound. For example, the Australian Securities and Investments Commission says in its website about financial advice: "The cost of the advice will depend on its scope. As a guide, expect to pay between $200 and $700 for simple advice and between $2000 and $4000 for more comprehensive advice."[13] This fits nicely with the figures in our last example. As a function of risk aversion estimates and of different pieces of information, one could come up with a more precise range for information prices, always under the assumption of frictionless financial markets allowing large investments.

Example 2 *Assume an agent has CRRA preferences with a coefficient of 2. If that person has the American median income,* $\rho_u(w_m) = 2\times 0.2\times 10^{-4}$,

[13] Taken from: https://www.moneysmart.gov.au/investing/financial-advice/financial-advice-costs

and if she is in the highest decile $\rho_u(w_m) = 2 \times 1.1 \times 10^{-6}$. This implies that, for the same preferences, she would accept a when $\beta = 0.55$ at a price of $1,000\,\$$ if her income were in the top decile, but she would reject it if she had the median income.

6 Some Properties of the Index

We now discuss some properties of our index for the normalized value of information.[14]

6.1 Continuity

The normalized value index $\mathcal{NV}$ is jointly continuous in μ, in p_α, and in $(q_\alpha^s)_s$ on the domain of nonexcluding information purchases. Continuity is a natural and attractive property: small changes in either the price or the conditional probabilities of signals should translate into small changes in the normalized value of the purchase. By "continuity at infinity", $\mathcal{NV}(a) = +\infty$ when a is excluding.

6.2 Blackwell monotonicity

The normalized value index is Blackwell-monotonic, as expressed in the following observation:

Observation 1: If an information structure α_1 is more informative than another information structure α_2 in the sense of Blackwell, then for any price $\mu > 0$, the information purchase (μ, α_1) is more valuable than the information

[14]Some of the proofs and more properties can be found in the appendix. Additional properties and examples can be found in the working paper version available at: http://ideas.repec.org/p/cte/werepe/we1224.html.

purchase (μ, α_2). Thus we have:

$$\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \alpha_2).$$

The observation shows that the complete ordering defined by the normalized value is an extension of Blackwell's ordering over information structures evaluated at the same price. Since the normalized value is a new ordering, this is also a new result. Our simple proof of the observation does not rely on the analytical form of the normalized value function, but rather on its axiomatic underpinning.

6.3 Mixtures

A third property concerns what happens when an information structure is constructed by randomizing over two other ones. Given information structures α_1, α_2 and $1 > \lambda > 0$, we let $\lambda\alpha_1 \oplus (1-\lambda)\alpha_2$ be the information structure in which (i) a coin toss determines whether the agent's signal is chosen from α_1 (with probability λ) or α_2 (with probability $1-\lambda$), and (ii) the agent is informed of both the outcome of the coin toss and the signal drawn from the chosen information structure. Formally, the set of signals in $\lambda\alpha_1 \oplus (1-\lambda)\alpha_2$ is $S_{\alpha_1} \cup S_{\alpha_2}$ (where we assume that S_{α_1} and S_{α_2} are disjoint), and the probability in state k that the agent receives signal $s \in S_{\alpha_1}$ is $\lambda\alpha_{1,k}(s)$, whereas the probability of a signal $s \in S_{\alpha_2}$ is $(1-\lambda)\alpha_{2,k}(s)$.

Observation 2: Consider $\mu > 0$ and α_1, α_2 such that $\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \alpha_2)$. For every $1 > \lambda > 0$, we have:

$$\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2) \geq \mathcal{NV}(\mu, \alpha_2).$$

Thus, quite naturally, the normalized value of the new information structure lies between the normalized value of the most valuable one and the normalized value of the least valuable one.

6.4 The Role of ln and exp in $\mathcal{NV}$

So far we have argued that two intuitive properties of the index $\mathcal{NV}$ are that it makes the normalized value of a purchase (i) a decreasing function of its price and (ii) an increasing function of the relative entropy from the prior to each generated posterior. In this light, one could consider using the following alternative index (see Kelly (1956)):

$$\hat{A}(a) = \frac{1}{\mu} \sum_{s} p_\alpha(s) d(p \| q^s_\alpha).$$

It is apparent that the index $\hat{A}$ retains those two properties, and it also satisfies separability in the form of price homogeneity. However, the next example illustrates why it does not rank the normalized value of purchases well.

Example 3 *Let $K = \{1, 2, 3\}$ and fix a uniform prior. Consider, for instance, two information structures with each having two signals:*

$$\alpha_1 = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \ \alpha_2 = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ 1/2 & 1/2 \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

Fix an arbitrary $\mu > 0$, and define the purchases $a_1 = (\mu, \alpha_1)$ and $a_2 = (\mu, \alpha_2)$. Note that $\hat{A}(a_1) = +\infty$ because the relative entropy of the prior to the posterior generated by the first signal is infinite. On the other hand, for any $\varepsilon > 0$, $\hat{A}(a_2)$ is finite. We next argue that the normalized value of the purchases is not well measured by $\hat{A}$. Indeed, for a small enough $\varepsilon > 0$, the purchase a_2 is almost excluding, and hence, in such a case $r_1 = \mathcal{NV}(a_1) < \mathcal{NV}(a_2) = r_2$. Here, r_1 and r_2 are the risk-aversion coefficients of the two CARA individuals who define the two corresponding levels of normalized value. Let $r = (r_1 + r_2)/2$. Clearly, the CARA agent

r uniformly likes information more than the CARA r_2 agent; the CARA r_2 agent accepts a_2, which according to the index $\hat{A}$ would be less valuable than a_1; but agent r rejects a_1.

The example makes clear the role of the exponential and its compensating logarithm as a "blow up/shrink down" of relative entropies. The exponential function with negative exponents, being bounded above, avoids the problem of infinite relative entropies attached to a single signal. Only when all relative entropies are infinite does the logarithm restore an infinite normalized value. This is essential in order to satisfy the duality between uniform preferences for information and the proposed function ranking the normalized value of purchases.

6.5 Comparison with Entropy Informativeness

We next present an example, similar to one in Cabrales, Gossner, and Serrano (2013), that illustrates how our framework serves to complete Blackwell's ordering. In addition, it shows how the information index in our 2013 paper can sometimes provide a different ranking from the induced index of information structures in the current study (when the price of the purchase is kept constant), while it also shows how both can sometimes point in the same direction.[15]

Example 4 *Let $K = \{1, 2, 3\}$ and fix a uniform prior. Consider two information structures that are not ordered in the sense of Blackwell. For instance, let each of the two information structures have two signals:*

$$\alpha_1 = \begin{bmatrix} 1-\varepsilon_1 & \varepsilon_1 \\ 1-\varepsilon_1 & \varepsilon_1 \\ \varepsilon_1 & 1-\varepsilon_1 \end{bmatrix}, \ \alpha_2 = \begin{bmatrix} 1-\varepsilon_2 & \varepsilon_2 \\ 0.1 & 0.9 \\ \varepsilon_2 & 1-\varepsilon_2 \end{bmatrix}$$

[15] Remember also that our entropy informativeness coincides, because of the role of the logarithmic investor, with the value of information from Kelly (1956).

For ε_1 and ε_2 small enough, these information structures are not ranked according to Blackwell. To see this, it suffices to consider two decision problems. In Problem 1, the agent must choose one of two actions: action 1 gives a utility of 1 only in the first two states, and 0 otherwise, while action 2 gives a utility of 1 only in the third state, and 0 otherwise. In contrast, Problem 2 has action 1 pay a utility of 1 only in the first state, and 0 otherwise, while action 2 gives a utility of 1 only in states 2 or 3, and 0 otherwise. When 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, thereby securing a utility of 1. This would be strictly greater than his utility after α_2. On the other hand, when facing Problem 2, he would under α_2 choose action 1 after the first signal and action 2 after the second, yielding a utility close to 29/30, which is greater than his optimal utility after α_1.

Now let us compute $\mathcal{NV}(a_1)$ for $a_1 = (\mu, \alpha_1)$ and $\mathcal{NV}(a_2)$ for $a_2 = (\mu, \alpha_2)$.

$$\mathcal{NV}(a_i) = -\frac{1}{\mu} \ln \left(\sum_s p_{\alpha_i}(s) \exp(-d(p \| q^s_{\alpha_i})) \right),$$

and

$$\begin{aligned}
\sum_s p_{\alpha_1}(s) \exp(-d(p\|q^s_{\alpha_1})) &= \frac{2-\varepsilon_1}{3} \exp\left(-\frac{1}{3}\left(\ln\left(\frac{\frac{1}{3}}{\frac{1-\varepsilon_1}{2-\varepsilon_1}}\right) + \ln\left(\frac{\frac{1}{3}}{\frac{1-\varepsilon_1}{2-\varepsilon_1}}\right) + \ln\left(\frac{\frac{1}{3}}{\frac{\varepsilon_1}{2-\varepsilon_1}}\right)\right)\right) \\
&\quad + \frac{1+\varepsilon_1}{3} \exp\left(-\frac{1}{3}\left(\ln\left(\frac{\frac{1}{3}}{\frac{1-\varepsilon_1}{1+\varepsilon_1}}\right) + \ln\left(\frac{\frac{1}{3}}{\frac{\varepsilon_1}{1+\varepsilon_1}}\right) + \ln\left(\frac{\frac{1}{3}}{\frac{\varepsilon_1}{1+\varepsilon_1}}\right)\right)\right) \\
&\simeq \frac{2}{3} \exp\left(-\frac{1}{3}\left(\ln \frac{1}{\varepsilon_1}\right)\right) + \frac{1}{3} \exp\left(-\frac{1}{3}\left(\ln\left(\frac{1}{\varepsilon_1}\right)^2\right)\right) \\
&\simeq \frac{2}{3}\varepsilon_1^{1/3} + \frac{1}{3}\varepsilon_1^{2/3}.
\end{aligned}$$

$$
\begin{aligned}
\sum_s p_{\alpha_2}(s)\exp(-d(p\|q^s_{\alpha_2})) &= \frac{1.1}{3}\exp\left(-\frac{1}{3}\left(\ln\left(\frac{\frac{1}{3}}{\frac{1-\varepsilon_2}{1.1}}\right)+\ln\left(\frac{\frac{1}{3}}{\frac{0.1}{1.1}}\right)+\ln\left(\frac{\frac{1}{3}}{\frac{\varepsilon_2}{1.1}}\right)\right)\right) \\
&\quad +\frac{1.9}{3}\exp\left(-\frac{1}{3}\left(\ln\left(\frac{\frac{1}{3}}{\frac{1-\varepsilon_2}{1.9}}\right)+\ln\left(\frac{\frac{1}{3}}{\frac{0.9}{1.9}}\right)+\ln\left(\frac{\frac{1}{3}}{\frac{\varepsilon_2}{1.9}}\right)\right)\right) \\
&\simeq \frac{1.1}{3}\exp\left(\frac{1}{3}\ln\varepsilon_2\right)+\frac{1.9}{3}\exp\left(\frac{1}{3}\ln\varepsilon_2\right)=\varepsilon_2^{1/3}.
\end{aligned}
$$

If $\varepsilon_1=\varepsilon_2$ and both are small enough, then $\mathcal{NV}(a_2)>\mathcal{NV}(a_1)$. On the other hand, if $\varepsilon_1=\varepsilon_2^2$ and both are small, then $\mathcal{NV}(a_1)>\mathcal{NV}(a_2)$.

Let us now compute the entropy reduction from the uniform prior, which we denote by $I_e(\cdot)$, letting $H(q)=\sum_{k=1}^{3}-q_k\ln(q_k)$.

$$
\begin{aligned}
I_e(\alpha_1) &= H(p)-\sum_{s=1}^{2}p^s_{\alpha_1}H\left(q^s_{\alpha_1}\right) \\
&= 3\left(-\frac{1}{3}\ln\left(\frac{1}{3}\right)\right)-\frac{2-\varepsilon_1}{3}\left(-2\frac{1-\varepsilon_1}{2-\varepsilon_1}\ln\left(\frac{1-\varepsilon_1}{2-\varepsilon_1}\right)-\frac{\varepsilon_1}{2-\varepsilon_1}\ln\left(\frac{\varepsilon_1}{2-\varepsilon_1}\right)\right) \\
&\quad -\frac{1+\varepsilon_1}{3}\left(-2\frac{\varepsilon_1}{1+\varepsilon_1}\ln\left(\frac{\varepsilon_1}{1+\varepsilon_1}\right)-\frac{1-\varepsilon_1}{1+\varepsilon_1}\ln\left(\frac{1-\varepsilon_1}{1+\varepsilon_1}\right)\right) \\
&\simeq \ln 3-\frac{2}{3}(\ln 2)\simeq \ln 3-0.46210.
\end{aligned}
$$

$$
\begin{aligned}
I_e(\alpha_2) &= 3\left(-\frac{1}{3}\ln\left(\frac{1}{3}\right)\right)-\frac{1.1}{3}\left(-\frac{1-\varepsilon_2}{1.1}\ln\left(\frac{1-\varepsilon_2}{1.1}\right)-\frac{0.1}{1.1}\ln\left(\frac{0.1}{1.1}\right)-\frac{\varepsilon_2}{1.1}\ln\left(\frac{\varepsilon_2}{1.1}\right)\right) \\
&\quad -\frac{1.9}{3}\left(-\frac{1-\varepsilon_2}{1.9}\ln\left(\frac{1-\varepsilon_2}{1.9}\right)-\frac{0.9}{1.9}\ln\left(\frac{0.9}{1.9}\right)-\frac{\varepsilon_2}{1.9}\ln\left(\frac{\varepsilon_2}{1.9}\right)\right) \\
&\simeq \ln 3-\frac{1}{3}(1.1\ln 1.1-0.1\ln 0.1+1.9\ln 1.9-0.9\ln 0.9) \\
&\simeq \ln 3-0.549815518.
\end{aligned}
$$

This implies that $I_e(\alpha_1)>I_e(\alpha_2)$ whenever $\varepsilon_1,\varepsilon_2$ are sufficiently close to zero.

The reason for the difference between entropy informativeness and the approach in this paper is the larger sensitivity of the index $\mathcal{NV}$ to information concerning low-probability events. In particular, α_1 causes a larger reduction in entropy, being associated with an almost fully informative signal (s_2). In contrast, for equal prices, a purchase of α_2 is more valuable for sequences of small ε_1 and ε_2 where $\varepsilon_1 = \varepsilon_2$. To understand the latter, note that the limits, as ε_1 and ε_2 vanish, of α_1 and α_2 lead to excluding information purchases, with infinite normalized value. The large unbounded normalized value of those purchases before going to the limits is explained by the large investments made following each signal. However, in α_1, following signal s_2, two states are becoming extremely unlikely, leading the agent to an optimal investment with large losses in these two states, whereas in α_2 large losses in the optimal investment are confined to only one state. Because of this, when ε_1 and ε_2 go to zero at the same rate, the large negative utility that a CARA agent derives from large negative wealth implies that α_2 is more valuable than α_1. Convergence rates matter, though: this conclusion is overturned if ε_1 goes to zero much faster than ε_2.

To explore somewhat more systematically the difference between the index based on entropy and the one in this paper, we investigate conditions on "small information" that renders them equivalent. Let $a_i = (\mu, \alpha_i)$. We then have the following equation:

$$
\begin{aligned}
\mathcal{NV}(a_i) &= -\frac{1}{\mu} \ln \left(\sum_s p_{\alpha_i}(s) \exp(-d(p \| q^s_{\alpha_i})) \right) \\
&= -\frac{1}{\mu} \ln \left(\sum_s p_{\alpha_i}(s) \exp \left(- \sum_k p_k \left(\ln p_k - \ln q^s_{\alpha_i}(k) \right) \right) \right)
\end{aligned}
$$

$$
\begin{aligned}
I_e(\alpha_i) &= -\sum_k p_k \ln p_k - \sum_s p_{\alpha_i}(s)\left(-\sum_k q^s_{\alpha_i}(k) \ln q^s_{\alpha_i}(k)\right) \\
&= \sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right) - \sum_k \left(p_k - q^s_{\alpha_i}(k)\right)\ln q^s_{\alpha_i}(k)\right).
\end{aligned}
$$

This implies that to a first order approximation when $q^s_{\alpha_i}$ is close to p,

$$
\begin{aligned}
\mathcal{NV}(a_i) &\simeq -\frac{1}{\mu}\ln\left(1 + \sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right)\right) \\
&\simeq -\frac{1}{\mu}\sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right) \\
&= \frac{1}{\mu}\sum_s p_{\alpha_i}(s)\left(\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right)
\end{aligned}
$$

and

$$
\begin{aligned}
I_e(\alpha_i) &= \sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right) - \sum_k \left(p_k - q^s_{\alpha_i}(k)\right)\ln q^s_{\alpha_i}(k)\right) \\
&= \sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right) + \left(\sum_k \left(\frac{q^s_{\alpha_i}(k)}{p(k)} - 1\right) p(k) \ln q^s_{\alpha_i}(k)\right)\right) \\
&\simeq \sum_s p_{\alpha_i}(s)\left(-\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right) - \left(\sum_k \left(\ln p_k - \ln q^s_{\alpha_i}(k)\right) p(k) \ln q^s_{\alpha_i}(k)\right)\right) \\
&= \sum_s p_{\alpha_i}(s)\left(-\sum_k \left(1 + \ln q^s_{\alpha_i}(k)\right) p_k \left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right).
\end{aligned}
$$

As a result, it follows that:

$$
\mathcal{NV}(a_i) \simeq \frac{1}{\mu}\sum_s p_{\alpha_i}(s)\left(\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right) \tag{3}
$$

and

$$
I_e(\alpha_i) \simeq \sum_s p_{\alpha_i}(s)\left(-\sum_k \left(1 + \ln q^s_{\alpha_i}(k)\right) p_k \left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right). \tag{4}
$$

A comparison of expressions (3) and (4) makes clear that when priors and posteriors are similar, the two indices point in the same direction - as long as it is also true that the $q^s_{\alpha_i}(k)$ vectors are all parallel to the unit vector and that $\ln q^s_{\alpha_i}(k) < -1$, that is, when priors are close to uniform and there are more than two states. Otherwise, cases such as the one provided in Example (4), when posteriors are very informative, are likely to make the indices diverge.

7 Related literature

As is well known, markets for information are increasingly important for the world economy, in spite of the difficulties in creating such markets, which Arrow (1962) already pointed out. Yet, be it through reputational concerns (Ottaviani and Sørensen, 2006), contracts (Hörner and Skrzypacz, 2010), or simply because transmitting information has its advantages (see, e.g. Vives, 1990; Fishman and Hagerty, 1995; Esö and Szentes, 2007), the markets for information exist and are quite sizable (Mehran and Stulz, 2007).

One of the difficulties of a market for information is that, unlike what happens with many other goods, measuring the quantity and quality of information is not easy. The classical approach to ranking information structures is due to Blackwell (1953).[16] This approach does not provide a complete ordering of information structures. More recent research has since focused on restricting preferences to a particular class. Lehmann (1988) restricts attention to problems with monotone decision rules, and Persico (2000), Athey and Levin (2001), and Jewitt (2007) do so to some more general classes of monotone problems.[17] The main difference between this line of research and

[16] Veldkamp (2011) shows the many ways in which economists have measured informativeness and their applications

[17] Measuring information is even harder if several agents interact as shown for instance

our approach is that we provide a complete order through a duality axiom for problems with a restricted set of investment opportunities.[18]

A central role in our research is played by risk preferences of individuals, and their heterogeneity, which has been established with a variety of sources, such as using household survey data about propensity to invest (Guiso and Paiella, 2008), from deductible choice in insurance markets (Cohen and Einav, 2005) and from laboratory experiments (Holt and Laury, 2002).

Our approach to ranking information purchases is based on a ranking of preferences for information. There are relatively few papers in the literature dealing with the comparison of different agents' preferences for information. One such study is Grant, Kajii, and Polak (1998) who explore intrinsic preferences for information, that is, preferences that are unrelated to the ability of information to make more profitable decisions. This is very different from our framework, since our agents like information precisely to help them make better decisions. But, interestingly, just as we found in Theorem 1 that risk aversion is related to preferences for information, Grant, Kajii, and Polak (1998) find that their notion of Information Loving is related to the convexity of preferences. Dubra and Echenique (2001) establish the impossibility of representing monotone preferences for information about an uncountable set of states of nature by a utility function. Li (2010) studies the relationship between intrinsic preferences for information and the aversion to ambiguity in dynamic decision settings.

Relative entropy plays an important role in our index. Kullback and Leibler (1951) shows that relative entropy measures the mean information

in Gossner (2000), Gossner and Mertens (2001) or Lehrer and Rosenberg (2006).

[18]Moscarini and Smith (2002), Azrieli (2014) and Ganuza and Penalva (2010) also study in various environments partial orderings of information structures.

per sample for distinguishing between two hypotheses when one of them is true.[19] The relative-entropy measure of proximity of probability distributions appears prominently in many economic settings (see e.g. Blume and Easley (1992) and Sandroni (2000) Hansen and Sargent (2010, 2001), Maccheroni, Marinacci, and Rustichini (2006), Gossner (2011)) However, none of these previous papers in economics uses relative entropy to measure the informativeness of signals or the normalized value of information purchases. Outside economics, relative entropy is widely used to measure both informativeness and differences between distributions. In information theory, Kraft (1949) and McMillan (1956); in statistics and econometrics Soofi and Retzer (2002); in linguistics, (Kuperman, Bertram, and Baayen, 2010; Mishra and Bangalore, 2011); in optics, (Ong, Xiaoy, Tham, and Ang, 2009); in hydrology, (Singh, 1997); in genetics, (Sherwin, 2010); in zoology, (Donaldson-Matasci, Bergstrom, and Lachmann, 2010); and even in archeology, (Justeson, 1973). Although many of these applications use relative entropy to measure informativeness, none of them provides a decision-theoretic microfoundation for such use.

8 Conclusion

There are multiple ways to index information, but ours is the first index that captures the information-price tradeoff, by indexing information purchases. Our normalized value index is based on a duality principle between value and preference for information in settings where the investment opportunities are described by a no-arbitrage condition. No-arbitrage assets provide a clean way to measure the value of information, furthering their use from

[19] A maximum likelihood estimator of a parametric model is also a minimizer of the Kullback-Leibler divergence.

our previous study (Cabrales, Gossner, and Serrano (2013)). The result we offer here can be viewed as a translation of Aumann and Serrano (2008) to informational settings with no-arbitrage investments. In such settings, the new index captures an aspect of the demand for information in a market economy. The paper has characterized agents' demand for information using a simply computable number called the normalized value of the information purchase, which relates to agents' risk aversion. For practical applications, one can rely on some of the known estimates for risk aversion levels provided in the literature in order to identify prices at which every agent/no agent will accept that information purchase. This, in effect, is a way to describe a useful inverse demand curve for information in no-arbitrage investment settings.

References

ARROW, K. (1962): "Economic welfare and the allocation of resources for invention," in *The rate and direction of inventive activity: Economic and social factors*, pp. 609–626. Nber.

ATHEY, S., AND J. LEVIN (2001): "The Value of Information in Monotone Decision Problems," Stanford University, Department of Economics Working Papers 01003.

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

AZRIELI, Y. (2014): "Comment on "The law of large demand for information"," *Econometrica*, 82(1), 415–423.

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

Blume, L., and D. Easley (1992): "Evolution and Market Behavior," *Journal of Economic Theory*, 58, 9–40.

Blume, L. E., and D. Easley (2002): "Optimality and natural selection in markets," *Journal of Economic Theory*, 107(1), 95–135.

Cabrales, A., O. Gossner, and R. Serrano (2013): "Entropy and the value of information for investors," *American Economic Review*, 103, 360–377.

Cesa-Bianchi, N., and G. Lugosi (2006): *Prediction, learning, and games*. Cambridge University Press.

Cohen, A., and L. Einav (2005): "Estimating risk preferences from deductible choice," Discussion paper, National Bureau of Economic Research.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual Risk Attitudes:Measurement, Determinants and Behavioral Consequences," *Journal of the European Economic Association*, 9, 522–550.

Donaldson-Matasci, M. C., C. T. Bergstrom, and M. Lachmann (2010): "The Fitness Value of Information," *Oikos*, 119, 219–230.

Dubra, J., and F. Echenique (2001): "Monotone Preferences over Information," *Topics in Theoretical Economics*, 1, 1–16.

Esö, P., and B. Szentes (2007): "Optimal information disclosure in auctions and the handicap auction," *The Review of Economic Studies*, 74(3), 705–731.

Fishman, M. J., and K. M. Hagerty (1995): "The incentive to sell financial market information," *Journal of Financial Intermediation*, 4(2), 95–115.

Ganuza, J.-J., and J. Penalva (2010): "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica*, 78(3), 1007–1030.

Gossner, O. (2000): "Comparison of information structures," *Games and Economic Behavior*, 30(1), 44–63.

——— (2011): "Simple Bounds on the Value of a Reputation," *Econometrica*, 79.

Gossner, O., and J.-F. Mertens (2001): "The value of information in zero-sum games," *preprint*.

Grant, S., A. Kajii, and B. Polak (1998): "Intrinsic Preference for Information," *Journal of Economic Theory*, 83, 233–259.

Guiso, L., and M. Paiella (2008): "Risk aversion, wealth, and background risk," *Journal of the European Economic association*, 6(6), 1109–1150.

Hansen, L., and T. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

Hansen, L. P., and T. J. Sargent (2010): "Wanting Robustness in Macroeconomics," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3, pp. 1097 – 1157. Elsevier.

Hart, S. (2011): "Comparing Risks by Acceptance and Rejection," *Journal of Political Economy*, 119, 617–638.

Hinton, G. E., and R. S. Zemel (1994): "Autoencoders, minimum description length and free energy," in *Advances in neural information processing*, ed. by J. Cowan, G. Tesauro, and J. Alspector, vol. 6. Morgan Kaufmann.

Holt, C. A., and S. K. Laury (2002): "Risk aversion and incentive effects," *American economic review*, 92(5), 1644–1655.

Hörner, J., and A. Skrzypacz (2010): "Selling information," Cowles Foundation Discussion Paper.

Jewitt, I. (2007): "Information Order in Decision and Agency Problems," Nuffield College.

Justeson, J. S. (1973): "Limitations of Archaeological Inference: An Information-Theoretic Approach with Applications in Methodology," *American Antiquity*, 80, 131–149.

Kelly, J. L. (1956): "A new interpretation of information rate," *Information Theory, IRE Transactions on*, 2(3), 185–189.

Kraft, L. G. (1949): "A Device for Quantizing, Grouping, and Coding Amplitude Modulated Pulses," MS Thesis, Electrical Engineering Department, Massachusetts Institute of Technology.

Kullback, S., and R. A. Leibler (1951): "On Information and Sufficiency," *Annals of Mathematical Statistics*, 22, 79–86.

Kuperman, V., R. Bertram, and R. H. Baayen (2010): "Processing Trade-offs in the Reading of Dutch Derived Words," *Journal of Memory and Language*, 62, 83–97.

LANDAU, L., AND E. LIFSHITZ (1980): *Statistical Physics*, vol. 5 of *Course of Theoretical Physics*. Butterworth-Heinemann, third edn.

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

LEHRER, E., AND D. ROSENBERG (2006): "What restrictions do Bayesian games impose on the value of information?," *Journal of Mathematical Economics*, 42(3), 343–357.

LI, J. (2010): "Preferences for Information and Ambiguity," works.bepress.com.

MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

MACLEAN, L. C., E. O. THORP, AND W. T. ZIEMBA (2011): *The Kelly capital growth investment criterion: Theory and practice*, vol. 3. world scientific.

MCMILLAN, B. (1956): "Two Inequalities Implied by Unique Decipherability," *IEEE Trans. Information Theory*, 2, 115–116.

MEHRAN, H., AND R. M. STULZ (2007): "The economics of conflicts of interest in financial institutions," *Journal of Financial Economics*, 85(2), 267–296.

MISHRA, T., AND S. BANGALORE (2011): "Predicting Relative Prominence in Noun-Noun Compounds," *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: short papers*, 49, 609–613.

Moscarini, G., and L. Smith (2002): "The law of large demand for information," *Econometrica*, 70(6), 2351–2366.

Ong, L.-L., W. Xiaoy, C.-K. Tham, and M. H. Ang (2009): "Resource Constrained Particle Filtering for Real-time Multi-Target Tracking in Sensor Networks," *PERCOM '09 Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications*, 09, 1–6.

Ottaviani, M., and P. N. Sørensen (2006): "Professional advice," *Journal of Economic Theory*, 126(1), 120–142.

Persico, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68, 135–148.

Ross, S. A. (1981): "Some stronger measures of risk aversion in the small and in the large with applications," *Econometrica*, 49, 621–663.

Samuelson, P. A. (1969): "Lifetime portfolio selection by dynamic stochastic programming," *The review of economics and statistics*, pp. 239–246.

Sandroni, A. (2000): "Do Markets Favor Agents Able to Make Accurate Predictions?," *Econometrica*, 68.

Sherwin, W. B. (2010): "Entropy and Information Approaches to Genetic Diversity and its Expression: Genomic Geography," *Entropy*, 12, 1765–1798.

Sierminska, E., A. Brandolini, and T. M. Smeeding (2006): "The Luxembourg Wealth Study – A cross-country comparable database for household wealth research," *Journal of Economic Inequality*, 4, 375–383.

Singh, V. P. (1997): "The use of Entropy in Hydrology and Water Resources," *Hydrological Processes*, 11, 587–626.

Soofi, E. S., and J. J. Retzer (2002): "Information Indices: Unification and Applications," *Journal of Econometrics*, 107, 17–40.

Veldkamp, L. (2011): *Information Choice in Macroeconomics and Finance*. Princeton University Press, Princeton, NJ.

Vives, X. (1990): "Trade association disclosure rules, incentives to share information, and welfare," *the RAND Journal of Economics*, 21, 409–430.

Appendix For Online Publication

This appendix presents additional material, both completing or providing proofs of results stated in the text of the paper, as well as expansions of some of the concepts presented. We order the sections in the appendix following the logical order of appearance in the paper.

A Completing the Proof of Theorem 1

We begin the proof of the converse part of the theorem by stating and proving several auxiliary lemmas.

Lemma 4 *Fix p and consider a sequence q^n of beliefs such that $q^n \to p$. Let x^n be the optimal investment for an agent with beliefs q^n. Then, it must be true that $x^n \to 0$.*

Proof. If the property does not hold, there exists a sequence $q^n \to p$ and a corresponding sequence of optimal investments x^n together with $\varepsilon > 0$ such that, for every n, $\|x^n\|_\infty \geq \varepsilon$. Since u is strictly concave, there exists $a > 0$ such that for every z with $|z| \geq \varepsilon$,

$$u(w+z) \leq u(w) + zu'(w) - a|z|.$$

We then have for every n:

$$\begin{aligned}
V(u,w,q^n) &= \sum_k q_k^n u(w + x_k^n) \\
&\leq + \sum_{|x_k^n| < \varepsilon} q_k^n (u(w) + u'(w)x_k^n) + \sum_{|x_k^n| \geq \varepsilon} q_k^n (u(w) + u'(w)x_k^n - a|x_k^n|) \\
&= u(w) + \sum_{|x_k^n| < \varepsilon} (q_k^n - p_k^n) u'(w) x_k^n + \sum_{|x_k^n| \geq \varepsilon} (q_k^n - p_k^n) u'(w) x_k^n - a q_k^n |x_k^n|,
\end{aligned}$$

where the last equality uses $\sum_k q_k^n = 1$ and $\sum_k p_k^n x_k^n = 0$. This implies both

$$\lim_{n\to\infty} \sum_{|x_k^n|<\varepsilon} (q_k^n - p_k^n) u'(w) x_k^n = 0$$

and

$$\limsup_{n\to\infty} \sum_{|x_k^n|\geq\varepsilon} (q_k^n - p_k^n) u'(w) x_k^n - -a q_k^n |x_k^n| < 0,$$

since for every n, there exists k such that $|x_n^n| \geq \varepsilon$. This shows that

$$\limsup_{n\to\infty} V(u, w, q^n) < u(w),$$

which is in contradiction with $V(u, w, q) \geq u(w)$ for every q. We conclude that the property holds as claimed. ■

Lemma 5 *Fix p and consider q close to p. Then, the optimal investment $x(q) = (x_k(q))_{k\in K}$ for an agent with belief $q = (q_k)_{k\in K}$ is*

$$x_k(q) = \frac{1}{p_k \rho(w)} (q_k - p_k) + o(\|q - p\|).$$

Proof. The agent's problem is to maximize $\sum_k q_k u(w + x_k)$ under the constraint $\sum_k p_k x_k = 0$. The solution is uniquely given by the system of first-order conditions:

$$q_k u'(w + x_k) = \lambda p_k,$$

where λ is independent of k. Using a first order Taylor expansion of $u'(w+x_k)$, we obtain:

$$u'(w) + x_k u''(w) = \lambda \frac{p_k}{q_k} + o(x_k). \tag{5}$$

We multiply each equation by p_k and sum over k to get:

$$u'(w) = \lambda \frac{\sum_j p_j^2}{q_j} + o(x_k). \tag{6}$$

We replace the value of λ obtained using (6) into equation (5) and get:

$$x_k = \frac{u'(w)}{u''(w)}\left(\frac{p_k}{q_k \sum_j \frac{p_j^2}{q_j}} - 1\right) + o(x_k).$$

In vector form, this can be expressed as:

$$x = F(q) + \gamma(x),$$

where $(F(q))_k = \frac{u'(w)}{u''(w)}(\frac{p_k}{q_k \sum_j \frac{p_j^2}{q_j}} - 1)$ and $\gamma(x) \in \mathbb{R}^K$ is such that $\frac{\|\gamma(x)\|}{\|x\|} \to 0$ as $\|x\| \to 0$.

We now show that $\|x\| = O(\|q - p\|)$. Assume to the contrary that there exists a sequence $q^n \to p$ and a corresponding sequence x^n such that $\frac{\|x^n\|}{\|q^n - p^n\|} \to \infty$. We would then have:

$$\frac{\|x^n\|}{\|q^n - p^n\|} \leq \frac{\|F(q^n)\|}{\|q^n - p^n\|} + \frac{\gamma(x^n)}{\|x^n\|}\frac{\|x^n\|}{\|q^n - p^n\|}.$$

However, a simple computation shows that $\|F(q^n)\| = O(\|q^n - p^n\|)$, and we know from Lemma 4 that $\|x^n\| \to 0$; hence, $\frac{\gamma(x^n)}{\|x^n\|} \to 0$. This yields a contradiction, and hence the conclusion that $\|x\| = O(\|q - p\|)$.

We thus have $\frac{\gamma(x)}{\|q-p\|} \to 0$ as $\|q - p\| \to 0$. We can therefore write

$$\begin{aligned} x_k &= \frac{u'(w)}{u''(w)}\left(\frac{p_k}{q_k \sum_j \frac{p_j^2}{q_j}} - 1\right) + o(\|q - p\|) \\ &= \frac{1}{\rho(w)}\left(\frac{q_k \sum_j \frac{p_j^2}{q_j} - p_k}{q_k \sum_j \frac{p_j^2}{q_j}}\right) + o(\|q - p\|) \\ &= \frac{1}{p_k \rho(w)}(q_k - p_k) + o(\|q - p\|), \end{aligned}$$

where the last line uses the fact that $\lim_{q \to p} \sum_j \frac{p_j^2}{q_j} = 1$. ■

Lemma 6 *Fix p and consider q close to p. Then,*

$$V(u,w,q) = u(w) + \frac{1}{2}\sum_k \frac{(q_k - p_k)^2}{\rho(w)p_k}u'(w) + o(\|q-p\|^2).$$

Proof. We have

$$V(u,w,q) = \sum_k q_k u(w + x_k),$$

where $x = (x_k)_{k\in K}$ is defined as in Lemma 5. A second order Taylor expansion gives

$$\begin{aligned} V(u,w,q) &= u(w) + \sum_k q_k x_k u'(w) + \frac{1}{2}\sum_k q_k x_k^2 u''(w) + o(\|x\|^2) \\ &= u(w) + \sum_k (q_k - p_k)x_k u'(w) + \frac{1}{2}\sum_k q_k x_k^2 u''(w) + o(\|x\|^2). \end{aligned}$$

From Lemma 5 we know that $\|x\| = O(\|q-p\|)$. Hence, we can replace $o(\|x\|^2)$ by $o(\|p-q\|^2)$ in the expression above. By substituting x_k for the expression in Lemma 5 we obtain:

$$\begin{aligned} V(u,w,q) &= u(w) + \sum_k \frac{(q_k-p_k)^2}{\rho(w)p_k}u'(w) \\ &\quad + \frac{1}{2}\sum_k \frac{q_k}{\rho(w)^2 p_k^2}(q_k - p_k)^2 u''(w) + o(\|q-p\|^2) \\ &= u(w) + \frac{1}{2}\sum_k \frac{(q_k-p_k)^2}{\rho(w)p_k}u'(w) + o(\|q-p\|^2), \end{aligned}$$

which is as claimed. ■

Fix p, and two states $k, l \in K$. For $\min\{p_k, p_l\} > \varepsilon > 0$, let $q^{\varepsilon,k}$ be given by $q^{\varepsilon,k}_{k'} = p_{k'}$ for $k' \neq k, l$; $q^{\varepsilon,k}_k = p_k + \varepsilon$; and $q^{\varepsilon,k}_l = p_l - \varepsilon$. Similarly, $q^{\varepsilon,l}$ is given by $q^{\varepsilon,l}_{k'} = p_{k'}$ for $k' \neq k, l$; $q^{\varepsilon,l}_l = p_l + \varepsilon$; and $q^{\varepsilon,l}_k = p_k - \varepsilon$. Thus, the belief $q^{\varepsilon,k}$ gives slightly higher weight to state k and slightly lower weight to state l than p, whereas $q^{\varepsilon,l}$ does the opposite. Now consider an information structure $\alpha(\varepsilon)$ such that with probability $\frac{1}{2}$, the agent's posterior is $q^{\varepsilon,k}$;

and with probability $\frac{1}{2}$ it is $q^{\varepsilon,l}$.(Such an information structure exists since $\frac{1}{2}q^{\varepsilon,k} + \frac{1}{2}q^{\varepsilon,l} = p$.)

Lemma 7 *For ε close to 0, the maximal price $\mu(\varepsilon)$ that an agent is willing to pay for $\alpha(\varepsilon)$ is:*

$$\mu(\varepsilon) = \frac{p_k + p_l}{2\rho(w)p_k p_l}\varepsilon^2 + o(\varepsilon^2).$$

Proof. The maximal price $\mu(\varepsilon)$ is such that the informational gains exactly compensate the monetary loss. Such a price satisfies the equation:

$$\frac{1}{2}\left(V(u, w - \mu(\varepsilon), q^{\varepsilon,k}) + V(u, w - \mu(\varepsilon), q^{\varepsilon,l})\right) = u(w).$$

Relying on Lemma 6, we get:

$$u(w) - u(w - \mu(\varepsilon)) \quad = \quad \frac{u'(w - \mu(\varepsilon))}{2\rho(w - \mu(\varepsilon))}(\frac{\varepsilon^2}{p_k} + \frac{\varepsilon^2}{p_l}) + o(\varepsilon^2).$$

This shows that $\mu(\varepsilon) \to 0$ as $\varepsilon \to 0$, and therefore, by taking a first-order Taylor approximation of $u(w - \mu(\varepsilon))$, we obtain:

$$\mu(\varepsilon)u'(w) + o(\mu(\varepsilon)) \quad = \quad \frac{u'(w)}{2\rho(w)}\frac{p_k + p_l}{p_k p_l}\varepsilon^2 + o(\varepsilon^2).$$

We conclude that:

$$\mu(\varepsilon) = \frac{p_k + p_l}{2\rho(w)p_k p_l}\varepsilon^2 + o(\varepsilon^2),$$

as we wanted to show. ■

Having proved this series of lemmata, one can proceed to the rest of the proof of the converse part of Theorem 1, in the main text.

B Proof of Theorem 4

Proof.

1. Let $r = \mathcal{NV}(a)$. Assume $\underline{R}(u) \geq \mathcal{NV}(a)$. Since u is DARA, $\rho_u(w) > \mathcal{NV}(a)$ for every w. The same computation as in the proof of Theorem 1 shows that for every z,

$$\frac{u(w+z)-u(w)}{u'(w)} < \frac{u_C^r(w+z)-u_C^r(w)}{u_C^{r\prime}(w)}.$$

If q has full support, the solution to the maximization problem of $\sum_k q_k u(w+x_k)$ under the constraint $\sum_k p_k x_k \leq 0$ is interior. Let $x(q)$ achieve this maximum. We have:

$$\begin{aligned}
\frac{V(u,w-\mu,q)-u(w)}{u'(w)} &= \frac{\sum_k q_k u(w-\mu+x_k(q))-u(w)}{u'(w)} \\
&< \frac{\sum_k q_k u_C^r(w-\mu+x_k(q))-u_C^r(w)}{u_C^{r\prime}(w)} \\
&\leq \frac{V(u_C^r,w-\mu,q)-u_C^r(w)}{u_C^{r\prime}(w)}.
\end{aligned}$$

If q does not have full support, we still have:

$$\begin{aligned}
\frac{V(u,w-\mu,q)-u(w)}{u'(w)} &= \sup_{x\in B^*} \frac{\sum_k q_k u(w-\mu+x_k)-u(w)}{u'(w)} \\
&\leq \sup_{x\in B^*} \frac{\sum_k q_k u_C^r(w-\mu+x_k(q))-u_C^r(w)}{u_C^{r\prime}(w)} \\
&\leq \frac{V(u_C^r,w-\mu,q)-u_C^r(w)}{u_C^{r\prime}(w)}.
\end{aligned}$$

Note that $\mathcal{NV}(a) \leq r$ implies that $\mathcal{NV}(a)$ is finite, and hence that a is nonexcluding; therefore, there exists s such that $p_\alpha(s) > 0$ and q^s has full support. Hence:

$$\begin{aligned}
\frac{\sum_s p_\alpha(s)V(u,w-\mu,q^s)-u(w)}{u'(w)} &< \frac{\sum_s p_\alpha(s)V(u_C^r,w-\mu,q^s)-u_C^r(w)}{u_C^{r\prime}(w)} \\
&= 0,
\end{aligned}$$

where the last equality stems from the fact that the agent u_C^r is indifferent between accepting and rejecting the information purchase a. We conclude that agent u rejects a at wealth level w.

Now, assume that $\underline{R}(u) < r$ and choose r_0 such that $\underline{R}(u) < r_0 < r$. Since an agent u_C^r accepts a at any wealth level, an agent $u_C^{r_0}$ strictly prefers accepting a at wealth level 0, which can be expressed as:

$$1 - \sum_s p_\alpha(s) \sup_{b^s \in B^*} \sum_k q_k^s \exp(r_0(\mu + x_k^s)) > 0.$$

Let $(x^s)_s$ then be a family of elements in B^* such that:

$$1 - \sum_s p_\alpha(s) \sum_k q_k^s \exp(r_0(\mu + x_k^s)) > 0.$$

Let w be such that $\rho(w - \mu + \min_{s,k} q_k^s) < r_0$. We have $\rho(z) < r_0$ for every $z \geq w - \mu + \min_{s,k} q_k^s$. It follows that by the same computation as in the proof of Theorem 1, for every s, k:

$$\frac{u(w - \mu + x_k^s) - u(w)}{u'(w)} \geq \frac{u_C^{r_0}(-\mu + x_k^s) - u_C^{r_0}(0)}{u_C^{r_0\prime}(0)}.$$

Therefore:

$$\begin{aligned} \frac{\sum_s p_\alpha(s) V(u, w - \mu, q_\alpha^s) - u(w)}{u'(w)} &\geq \frac{\sum_s p_\alpha(s) \sum_k q_k^s u(w - \mu + x_k^s) - u(w)}{u'(w)} \\ &\geq \frac{\sum_s p_\alpha(s) \sum_k q_k^s u_C^{r_0}(-\mu + x_k^s) - u_C^{r_0}(0)}{u_C^{r_0\prime}(0)} \\ &> 0. \end{aligned}$$

Hence, u accepts a at wealth w.

2. Analogous.

■

C Further Justifications of the Index

C.1 Duality-Based Approaches

Since rankings of preferences for information are of interest in their own right, we examine two alternative definitions thereof, one of which being a complete ordering over agents with decreasing risk aversion. It will be apparent that the ranking introduced in Section 4 and the two alternative rankings introduced next differ significantly. Next, following a parallel approach to Definition 1, we define orderings of information purchases according to the "duality" axiom of Aumann and Serrano (2008), a monotonicity property with respect to each of the alternatives concerning preferences for information.

The second definition of a ranking for preferences for information requires agent u_1 to accept the information purchase at *some* wealth level whenever u_2 accepts it at some wealth level. This is thus weaker than the definition of uniformly liking information better that requires u_1 to accept the information at *all* wealth levels whenever u_2 accepts it at some wealth level. We restrict attention to agents who are in the class $\mathcal{U}_{DA}$ of utility functions.

Let $u_1, u_2 \in \mathcal{U}_{DA}$. We say that u_1 *minimally likes information better* than u_2 if, for every information purchase a, and for every w_2, there exists w_1 such that, if u_2 accepts a at w_2, then so does u_1 at w_1.

This definition is an extremely weak requirement and it orders a large set of agents, as will be shown shortly.

Our third definition requires the wealth levels at which u_1 and u_2 are compared to be identical. It allows utilities to be defined over any bounded or unbounded open interval. We let $\mathcal{U}_{DA}^I$ be the class of utility functions u that are defined over an open interval of $\mathbb{R}$, twice differentiable, and such that ρ_u is decreasing. The following definition is general in that it allows the wealth intervals on which the two compared utility functions are defined to differ.

Let $u_1, u_2 \in \mathcal{U}_{DA}^I$, with u_1 defined over I_1 and u_2 over I_2. We say that u_1 *wealth wise likes information better* than u_2 if $I_1 \supseteq I_2$ and, for every information purchase a and wealth level w, if u_2 accepts a at w, then so does

u_1.

For agent u_1 to wealth wise like information better than u_2, it is required that u_1 accepts information purchases more often than u_2, when the comparison holds at the same wealth level. Since u_1 cannot accept information purchases for wealth levels outside of I_1, it is necessary that that I_1 is a superset of I_2. It is therefore implicit in the last definition that the agent rejects all information purchases that would make the wealth after purchase $w - \mu$ lie outside of the domain of the utility function u.

The following Theorem characterizes the orderings of these two definitions in terms of levels of risk aversion, and it should be compared with Theorem 1.

Theorem 5 *1. Let $u_1, u_2 \in \mathcal{U}_{DA}$, u_1 minimally likes information better than u_2 if and only if:*

$$\underline{R}(u_1) \leq \underline{R}(u_2).$$

2. Let $u_1, u_2 \in \mathcal{U}^I_{DA}$, with respective domains I_1 and I_2, $I_1 \supseteq I_2$. Then, u_1 wealth wise likes information better than u_2 if and only if

$$\forall w \in I_2, \quad \rho_{u_1}(w) \leq \rho_{u_2}(w).$$

Proof. Point 1 is a direct consequence of Theorem 4. Point 2 follows from similar arguments as in the proof of Theorem 1. ■

In particular, some consequences of Theorem 5 are that the "minimally likes information" ordering is complete over the set of DARA agents, and that the uniform ordering is stronger than the wealth wise ordering which is itself stronger than the minimal ordering.

We now define orderings of information purchases according to "duality" with regards to the orderings on preferences for information. The two following definitions parallel Definition 1.

First, using the "minimally likes information better" ordering:

Definition 2 *Let a_1, a_2 be two information purchases. We say that a_1 is* minimally more valuable *than a_2 if, given two agents $u_1, u_2 \in \mathcal{U}_{DA}$ such that u_1 minimally likes information better than u_2, whenever u_2 accepts a_2 at some wealth level, u_1 also accepts a_1 at some wealth level.*

And second, relying on the "wealth wise likes information better" ordering:

Definition 3 *Let a_1, a_2 be two information purchases. We say that a_1 is* wealth wise more valuable *than a_2 if, for every $u_2 \in \mathcal{U}_{DA}^I$, and $u_2 \in \mathcal{U}_{DA}$ such that u_1 wealth wise likes information better than u_2, if u_2 accepts a_2 at some wealth level, u_1 also accepts a_1 at some wealth level.*

This definition requires the set of agents u_1 who accept a_1 at some wealth level to be neither too large, nor too small. This set is potentially smaller than the set of agents in U_{DA}^I who like information better than u_2, but it has to include all elements of U_{DA} who like information better than u_2.

Theorem 6 below states the characterization of these orderings over information purchases.

Theorem 6 *Let a_1, a_2 be two information purchases. The following three statements are equivalent:*

1. *a_1 is uniformly more valuable than a_2,*
2. *a_1 is minimally more valuable than a_2,*
3. *a_1 is wealth wise more valuable than a_2.*

And in particular, they are all equivalent to:

$$\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2).$$

Proof. The equivalence between the uniformly more valuable ordering and the normalized value is proved in Theorem 2.

We next prove that a_1 is minimally more valuable than a_2 if and only if $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$. Assume that a_1 is minimally more valuable than a_2. Consider a CARA agent with risk aversion level $r = \mathcal{NV}(a_2)$, such an agent accepts a_2 at all (hence some) wealth levels. The same agent must also accept a_1 at some wealth level, which by Theorem 4 implies that $\mathcal{NV}(a_2) = r \leq \mathcal{NV}(a_1)$. Now assume $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$, and consider $u_1 \in U_{DA}$ who minimally likes information better than $u_2 \in U_{DA}$. If u_2 accepts a_2 at some wealth level, Theorem 5 and Theorem 4 imply $\underline{R}(u_1) \leq \underline{R}(u_2) \leq \mathcal{NV}(a_2) \leq \mathcal{NV}(a_1)$, hence u_1 accepts a_1 at some wealth level.

Now we prove that a_1 is wealth wise more valuable than a_2 if and only if $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$. Assume that a_1 is wealth wise more valuable than a_2. Again, a CARA agent with risk aversion level $\mathcal{NV}(a_2) - \varepsilon$ for any $\varepsilon > 0$ accepts a_2 at some wealth level. The same agent (being in U_{DA}) also accepts a_1 at some wealth level, which by Theorem 4 implies that $\mathcal{NV}(a_2) \leq \mathcal{NV}(a_1)$. Finally, assume that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$. Consider $u_1 \in \mathcal{U}_{DA}$ who wealth wise likes information better than $u_2 \in \mathcal{U}_{DA}$, both defined on an open interval I and assume that u_2 accepts a_2 at some wealth level $w \in I$. A CARA agent with degree of risk aversion $\inf_I \rho_{u_2}(w)$ is wealth wise less risk averse than u_2 by Theorem 5, hence accepts a_2 at all wealth levels, and hence, also at w. Since u_1 is wealth wise less risk averse than u_2, $\underline{R}(u_1) \leq \inf_I \rho_{u_1}(w) \leq \inf_I \rho_{u_2}(w)$. Hence $\underline{R}(u_1) \leq \inf_I \rho_{u_2}(w) \leq \mathcal{NV}(a_2) \leq \mathcal{NV}(a_1)$, which implies that u_1 accepts a_1 at some wealth level. ■

C.2 Total rejections or acceptances

In this subsection we provide a result in the spirit of Hart (2011), together with a similar result based on the notion of "accepting for all w" (instead of "rejecting for all w".

Following Hart (2011)'s approach (see also Cabrales, Gossner, and Serrano, 2013), we now introduce the definitions of uniform wealth rejection and acceptance:

Definition 4 *Let a_1 and a_2 be two information purchases. We say that a_2 is uniformly more rejected than a_1 if any $u \in \mathcal{U}_{DA}$ that rejects a_1 at all wealth levels also rejects a_2 at all wealth levels. We say that a_1 is uniformly more accepted than a_2 if any $u \in \mathcal{U}_{DA}$ that accepts a_2 at all wealth levels also accepts a_1 at all wealth levels.*

The first part of the definition proposes a uniform rejection of purchases within the DARA class of preferences. That is, a_2 is uniformly more rejected than a_1 because the former is rejected more often: whenever a_1 is rejected at all wealth levels, so is a_2, but not vice versa. The second part of the definition proposes a uniform acceptance of purchases within the same class of preferences. That is, a_2 is uniformly more accepted than a_1 because the former is accepted more often: whenever a_1 is accepted at all wealth levels, so is a_2, but not vice versa. The definition leads to the following result:[20]

Theorem 7 *Let a_1 and a_2 be two information purchases. The following three conditions are equivalent:*

- *a_2 is uniformly more rejected than a_1*

[20]We observe that the same theorem holds if we restrict the class of functions by imposing IRRA and ruin aversion on top of DARA. IRRA and ruin aversion are the restrictions on preferences used in Cabrales, Gossner, and Serrano (2013).

- *a_1 is uniformly more accepted than a_2*
- *$\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$.*

Proof. Assume that a_2 is uniformly more rejected than a_1. For every $\varepsilon > 0$, Theorem 4 shows that an agent $u_C^{\mathcal{NV}(a_1)+\varepsilon}$ rejects a_1 at all wealth levels. Hence such an agent also rejects a_2 at all wealth levels, which implies, again by Theorem 4, that $\mathcal{NV}(a_1) + \varepsilon \geq \mathcal{A}(a_2)$. Since this is true for every $\varepsilon > 0$, it follows that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$.

For the converse, assume that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$, and that $u \in \mathcal{U}_{DA}$ rejects a_1 at all wealth levels. Then by Theorem 4, $\underline{R}(u) \geq \mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$, and u also rejects a_2 at all wealth levels.

Assume that a_1 is uniformly more accepted than a_2. For every $\varepsilon > 0$, Theorem 4 shows that an agent $u_C^{\mathcal{NV}(a_2)-\varepsilon}$ accepts a_2 at all wealth levels. Hence such an agent also rejects a_1 at all wealth levels, which implies, again by Theorem 4, that $\mathcal{NV}(a_2) - \varepsilon \leq \mathcal{A}(a_1)$. Since this is true for every $\varepsilon > 0$, it follows that $\mathcal{NV}(a_2) \leq \mathcal{NV}(a_1)$.

For the converse, assume that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$, and that $u \in \mathcal{U}_{DA}$ accepts a_2 at all wealth levels. Then by Theorem 4, $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2) \geq \overline{R}(u)$, and u also accepts a_1 at all wealth levels. ■

D Additional Material Concerning Properties of the Index

D.1 Proof of Observation 1 (Blackwell Monotonicity)

Proof. Assuming that α_1 is more informative than α_2 in the sense of Blackwell, and fixing any arbitrary wealth level w, then any CARA agent who rejects (μ, α_1) at wealth level w also rejects (μ, α_2) at wealth level w. It follows

from the characterization of $\mathcal{NV}$ in Theorem 3 that $\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \alpha_2)$.
■

D.2 Proof of Observation 2 (Mixtures)

Proof. Fix any wealth level. From Theorem 3, a CARA agent with coefficient of risk aversion $\mathcal{NV}(\mu, \alpha_2)$ accepts both purchases (μ, α_1) and (μ, α_2) at wealth w; this agent therefore also accepts the purchase $(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2)$ at that wealth level. This shows that

$$\mathcal{NV}(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2) \geq \mathcal{NV}(\mu, \alpha_2).$$

Now consider $\varepsilon > 0$. Again from Theorem 3, a CARA agent with coefficient of risk aversion $\mathcal{NV}(\mu, \alpha_1) + \varepsilon$ rejects both purchases (μ, α_1) and (μ, α_2) at wealth w; this agent therefore also rejects the purchase $(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2)$ at that wealth level. This shows that

$$\mathcal{NV}(\mu, \alpha_1) + \varepsilon \geq \mathcal{NV}(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2)$$

for every $\varepsilon > 0$, and hence that

$$\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \lambda\alpha_1 \oplus (1-\lambda)\alpha_2).$$

■

D.3 The Role of Prices and Priors

In the model of Section 2, p plays a dual role. Indeed, p is the agent's prior before he receives any information, and it is also a vector of prices for Arrow securities that defines the set of no-arbitrage assets B^*. In order to both allow for the agent's prior to be different from the price system, and disentangle the two roles of p, we consider here agents whose prior belief $q \in \Delta(K)$ may differ from the vector p defining the set B^*.

In this more general model, an agent accepts an information purchase $a = (\mu, \alpha)$ at prior q if and only if:

$$\sum_s q_\alpha(s) V(u, w - \mu, q^s_\alpha) \geq V(u, w, q),$$

where q^s_α is the agent's posterior belief after receiving a signal s given the prior q and $q_\alpha(s) = \sum_k q_k \alpha_k(s)$. Note that if $q = p$, then $V(u, w, q)$ equals $u(w)$ so that the definition particularizes to the original one in this case.[21]

Our Definition 1 extends as follows: We say that a_1 is more valuable than a_2 at prior q if, given two agents u_1, u_2 such that u_1 uniformly likes information better than u_2 and two wealth levels w_1, w_2, whenever agent u_2 accepts a_2 at wealth level w_2 and prior q, then agent u_1 accepts a_1 at wealth level w_1 and prior q.

Then, we define the normalized value of an information purchase $a = (\mu, \alpha)$ at prior q as:

$$\begin{aligned} \mathcal{NV}(a, q) &= -\frac{1}{\mu} \ln \left(\sum_s p_\alpha(s) \exp(-d(p \| q^s_\alpha)) \right) - \frac{d(p\|q)}{\mu} \\ &= \mathcal{NV}(a) - \frac{d(p\|q)}{\mu}. \end{aligned}$$

As a word of caution, we note that in the formula above, as $(q^s_\alpha)_s$ depends on q, so does $\mathcal{NV}(a)$. Our results can be extended to this more general setting in the way one should expect (we omit details for brevity).

D.4 Sequential purchases

Another property worth mentioning concerns the normalized value of an information purchase in which the buyer receives signals sequentially from

[21] It is convenient to write the RHS of this expression this way, given our analysis of sequential purchases in the next subsection.

different information structures. Given an information structure α with a set of signals S_α and a family $\beta = (\beta_s)_{s \in S_\alpha}$ of information structures, where all the members of β share the same set of signals S_β, we let (α, β) be the information structure in which the agent first receives a signal s from α, then an independently drawn (conditional on k) signal s' from β_s. Formally, the set of signals in (α, β) is $S_\alpha \times S_\beta$, and in state k, the probability of receiving the pair of signals (s, s') is $\alpha_k(s)\beta_{s,k}(s')$. Given an information purchase $a = (\mu, \alpha)$ and a family of information purchases $b = (x_s)_s = (\nu, \beta_s)_s$, where all the members of b have the same price ν, we let $a+b$ denote the information purchase $(\mu + \nu, (\alpha, \beta))$.

Observation 3: Given information purchases a and $x = (x_1, \ldots, x_s, \ldots, x_K)$, the following hold:

1. If for every s, $\mathcal{NV}(x_s, q_s) \geq \mathcal{NV}(a)$, then $\mathcal{NV}(a + x) \geq \mathcal{NV}(a)$.
2. If for every s, $\mathcal{NV}(x_s, q_s) \leq \mathcal{NV}(a)$, then $\mathcal{NV}(a + x) \leq \mathcal{NV}(a)$.
3. In particular, if for every s, $\mathcal{NV}(x_s, q_s) = \mathcal{NV}(a)$, then $\mathcal{NV}(a + x) = \mathcal{NV}(a)$.

Proof. We prove the observation using the following auxiliary decision problem. In the first stage, the agent can either accept information purchase a or reject it. If the agent accepts a, then a signal s is drawn from α and the agent can either accept the information purchase x_s or reject it. If the agent rejects a, no other information purchase is offered to the agent. Once the agent has acquired some information (or none), any asset in B^* may be purchased; then the state k is realized, and the agent receives the corresponding payoff.

Assume that for every s, $\mathcal{NV}(x_s, q_s) \geq \mathcal{NV}(a)$, and consider an agent $u_C^{\mathcal{NV}(a)}$ at any wealth level and any prior p. In the sequential decision problem, assuming that a is accepted in the first stage by this agent, then x_s is accepted in the second stage for every s. Also, a is accepted in the first stage even if the option of acquiring x_s in the second stage is absent. Therefore, a is also accepted with the option of acquiring x_s in the second stage. Hence, an optimal strategy for the agent is to accept a, and then accept x_s no matter what s is. In particular, this strategy is better for the agent than not acquiring any information purchase. This shows that the agent accepts $a + x$, and hence that $\mathcal{NV}(a + x) \geq \mathcal{NV}(a)$.

Now assume that for every s, $\mathcal{NV}(x_s, q_s) \leq \mathcal{NV}(a)$, and consider an agent u_C^ρ with $\rho > \mathcal{NV}(a)$ at any wealth level and any prior p. In the sequential decision problem, assuming that a is accepted in the first stage, it is optimal for this agent to reject x_s after every signal s. Hence, the decision to acquire a in the sequential decision problem is equivalent to the decision to acquire a alone, and so this agent rejects a. Hence, the optimal strategy for the agent is to reject information. In particular, not acquiring any information is better than acquiring a, which is itself better than acquiring a and x_s following every s, so that no information is better than $a + x$. Therefore, the agent rejects $a + x$, which shows that $\mathcal{NV}(a + x) < \rho$ for every $\rho > \mathcal{NV}(a)$. This implies that $\mathcal{NV}(a + x) \leq \mathcal{NV}(a)$.

The third point follows immediately from the first and second points. ■

Observation 3 relates the normalized value of an information purchase involving (α, β) to the normalized value of the information purchases involving α and $(\beta_s)_s$. As a result, the normalized value of an information purchase involving α has to be measured given the prior p, as in formula (2), but the normalized value of an information purchase involving β_s has to be measured given the belief q_α^s of the agent after receiving the signal s. The observation makes intuitive sense: if the agent faces a sequence of purchases whose individual normalized value is increasing, the normalized value of the overall purchase is at least that of the normalized value of the first-stage purchase, and so on.