Antonio Cabrales, Olivier Gossner, Roberto Serrano,

A normalized value for information purchases

Article (Accepted version) (Refereed)

Original citation: Cabrales, Antonio and Gossner, Olivier and Serrano, Roberto (2017) *A normalized value for information purchases.* Journal of Economic Theory, 170. pp. 266-288. ISSN 0022-0531

DOI: 10.1016/j.jet.2017.05.007

A Normalized Value for Information Purchases

Antonio Cabrales, Olivier Gossner, and Roberto Serrano*

This version: April 2017

Abstract: Consider agents who are heterogeneous in their preferences and wealth levels. These agents may acquire information prior to choosing an investment that has 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 who is indifferent between accepting and rejecting the purchase, and it is characterized by a "duality" principle that states that agents with a 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 a competitive economy with no-arbitrage investments, our model therefore comes close to describing an inverse demand for information, by predicting what pieces of information are acquired by agents and which ones are not. 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, decision under uncertainty, 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 are grateful to Mihaïlo Backović, Héctor Calvo, Andrew Caplin, Mark Dean, Eddie Dekel, Rich Kihlstrom, Romain Rancière, Phil Reny, Larry Samuelson, Marco Scarsini, Eytan Sheshinski, Ran Shorrer, Laura Veldkamp, and anonymous referees 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. The current paper aims at answering the following questions. First, given an information purchase, can its normalized value, which captures the information-price tradeoff, be uniquely characterized?[1] Second, in a competitive economy, who are the agents willing to go ahead with a given information purchase?

We answer these questions by analyzing information purchases made by risk-averse agents (or investors) prior to choosing among risky investments. Key to our analysis is the set of investments available which we call *no-arbitrage investments*. In order to study information acquisition based on investment motives, we assume that no-arbitrage investments are not profitable under the agents' prior; the only investors who find investments profitable are the ones who acquire some information. Second, we note that for the value of information to be meaningful and comparable across heterogeneous agents, one needs the set of available investments to be rich enough; in particular, we assume the existence of complete markets.[2] We follow the literature on this subject (such as Kelly, 1956; Arrow, 1971) and take as *no-arbitrage* investments the set of all those with a nonpositive expected return under the prior.[3]

We begin by showing that an agent's demand, or her preference, for information

[1] We add the "normalized" qualification because the index will also rely on the price of the purchase, and not be based only on the information structure. The term "normalized informativeness" could also be used to refer to our index.

[2] For instance, some agents find it beneficial to have access to a certain futures market, while others do not value information about this market as much.

[3] No-arbitrage investments, in this sense of not offering any profitable or "arbitrage" opportunity, were also used in Cabrales, Gossner, and Serrano (2013). "No-speculative" investments could be an alternative name to express the same idea.

is characterized by her degree of risk aversion. Less risk-averse agents have a stronger preference for information than do more risk-averse agents, in the following sense. We show that an agent u_1 is uniformly less risk-averse than agent u_2 *if and only if* the fact that agent u_2 acquires some information is enough to guarantee that agent u_1 also acquires that information, independently of the wealth levels considered.[4] Therefore, agents' demand for information in our model is entirely captured by their uniform ranking of risk aversion.

We seek an objective underpinning of normalized values. That is, paralleling the approach of Aumann and Serrano (2008) for ordering riskiness, we apply the following duality principle to define the normalized value of an information purchase:[5] For an information purchase to be considered as objectively 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, when we say that "u_1 likes information better than u_2," we 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, then 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 who is 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

[4] To be precise, u_1 is uniformly less risk-averse than u_2 when, 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.

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

is increasing with respect to the former and decreasing with respect to the latter.[6]

The fact that a purchase has a higher normalized value than another one does not mean that more agents will accept it. Rather, a purchase having a higher normalized value is equivalent to that purchase having a larger set of agents who accept it regardless of their initial wealth. Thus, for CARA agents, the more valuable a purchase is, 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 less than the normalized value for all wealth levels, will accept it. The remaining agents, whose risk aversion for some initial wealth level is below the normalized value but above it for other wealth levels, may accept or reject the purchase; in other words, their decision is "subject to wealth effects." Were it not for this latter set of agents, the entire set of agents who accept an information purchase would be monotonically increasing in the normalized value of the purchase. Having said that, if we assume that agents are sufficiently heterogeneous in their wealth, and that we can know the degree of risk aversion of agents, then the normalized value of information purchases is a useful tool that delivers a clear prediction of what pieces of information will be acquired in a competitive economy (i.e., in which each investor views herself as so small that her information purchase decision does not signal any relevant information to the rest of the economy).[7]

[6] In the appendix, we provide two alternative definitions of preference for information: (i) "If u_2 accepts a purchase at some wealth level, then u_1 accepts it at some wealth level" (minimal comparison); and (ii) "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 based on the duality principle. Strikingly, all three orderings of information purchases coincide: all three are represented by our normalized-value formula. As we also show 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.

[7] Even nonatomic agents are often able to avoid detection of informed trading in financial markets by, for example, hiding the order size. See, e.g. Bessembinder, Panayides, and Venkataraman

Our normalized-value measure provides some interesting insights on the demand for information. For instance, a decrease by a certain percentage in the cost of information translates into an increase in the same proportion in the normalized value of the corresponding information purchase. This means that, whenever the price of information drops by half, agents who are twice more risk-averse than those who initially bought a piece of information, become now willing to acquire that 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 those with null informational content. Perhaps surprisingly, the most valuable purchases include not only those that always allow the purchaser to learn the true fundamental state, but also all those that always allow the purchaser to exclude one fundamental state from the set of all possible states that will be realized. More generally, our ordering of information purchases is monotonic with respect to Blackwell (1951)'s ordering of information structures.

As with any index, it is important to understand what our formula is telling us. Ours is the first index to directly capture the information-price tradeoff in an information purchase. It does so by separating that tradeoff from potential wealth effects. We assume a competitive economy in which only no-arbitrage investments are available before the arrival of new information. In such an economy, the higher the index of an information purchase, the larger set of risk-averse investors who would accept the purchase regardless of their wealth. Also, a smaller set of risk-averse investors would reject it regardless of their wealth. Since it is a numeric index, our normalized-value formula orders *all* information purchases. And hence, for a fixed price, it orders *all* information structures.

It is important to compare our results here with those of Cabrales, Gossner, and Serrano (2013), where we have provided another information index. In that paper, the informativeness of an information structure is characterized by the reduction of

(2009).

entropy from the prior to the posteriors. That entropy informativeness index is silent regarding the information-price tradeoff in an information purchase. Nevertheless, 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.

The methodology we use here is very different, and it complements the one we used in Cabrales, Gossner, and Serrano (2013). In that paper, one piece of information is more valuable than another one if the maximal price any agent in the economy is willing to pay for the former is larger than the maximal willingness to pay for the latter. The comparison is made while all agents have the same wealth level. Also, the class of utility functions considered there excluded the possibility of ruin. Therefore, that index is more relevant when agents are homogeneous in wealth and wish to end up with nonnegative wealth with probability 1. In contrast, in this paper, we allow wealth levels to differ across agents, who consider the contingency of negative wealth. This is either because of a high price in the information purchase or because of losses in the investment chosen. Importantly, our index is a vehicle to reveal wealth-independent demand for information, defined as the set of agents who accept the purchase regardless of their wealth, rather than on the maximal price an agent in the economy is willing to pay for it. Thus, we view both indices as providing different useful ways to evaluate a given information structure. As an additional insight, we have shown that while our "maximum price dominance" led to a well-known object –entropy reduction– in our previous paper, thereby endowing entropy with new meaning, our more axiomatic approach in the current study, based on duality, leads to a formula that is new to information theory.

The paper is organized as follows. Section 2 describes the model. Section 3 relates the value of information and 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 levels of risk aversion in the economy. Section 6 examines 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.

2 The Model

We consider an agent who, prior to making an investment decision, may acquire some information at a cost. In this section we define the conditions under which this agent acquires a 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 nondecreasing, strictly concave,[8] and twice differentiable. We let $\mathcal{U}$ be the set of such monetary utility functions. We identify agents by their monetary utility functions; thus, the term "agent u" refers 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_2 is uniformly less risk-averse than agent u_1 whenever $\overline{R}(u_2) \leq \underline{R}(u_1)$. It is sometimes necessary to assume that u has decreasing absolute risk aversion in wealth. We thus let $\mathcal{U}_{DA}$ be the subset of agents in $\mathcal{U}$ for whom ρ_u is nonincreasing.

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

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 nonpositive expected return: $B^* = \{x \in \mathbb{R}^K, \sum_k p_k x_k \leq 0\}$.[9] When (i) the agent's initial wealth is w, (ii) $x \in B^*$ is chosen, and (iii) state k is realized, the agent's final wealth is $w + x_k$.

Some features of the set B^* of available investments are worth emphasizing. First, $0 \in B^*$, that is, not investing is feasible. Second, no agent in our class prefers to invest in the absence of new information, although this may change if some new information arrives. With "no-arbitrage" assets, the only motive for investing is the arrival of new information, and for this reason "no-arbitrage" assets provide a useful framework to measure the value of information. Finally, B^* is unbounded, an assumption which, although common in the literature on investment under limited information (see, e.g., Kelly, 1956; Arrow, 1971), may appear incompatible with real-life markets. Note, however, that all agents whose posteriors are in the interior of $\Delta(K)$ choose investments that are bounded. This implies in particular that replacing B^* by a sufficiently large compact subset spanning all dimensions of uncertainty (complete markets) would not change the analysis when posteriors are interior. See also Section 6.6 for more on this point.

2.3 Information Purchases

Before choosing an investment, the agent has the opportunity to engage in an information purchase $a = (\mu, \alpha)$. Here, $\mu > 0$ represents the cost of the information

[9] The vector $p = (p_1, \ldots, p_k, \ldots, p_K)$ 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 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 E.3 in the appendix.

purchase, paid up front, 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.[10] When the state of nature is k, $\alpha_k(s)$ is the probability that the signal observed by the agent is s. Signal s has an ex-ante probability $p_\alpha(s) = \sum_k p_k \alpha_k(s)$ of being realized, and we assume, without loss of generality, that $p_\alpha(s) > 0$ for every s. For each signal $s \in S_\alpha$, we let $q^s_\alpha = (q^s_k)_k \in \Delta(K)$ be the posterior probability distribution conditional on s, and derived from Bayes' rule.

We say that a is *excluding*, if for every signal s, there exists k such that $q^s_k = 0$. 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 her 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:

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

[10] In Cabrales, Gossner, and Serrano (2012), an earlier working paper version of the current work, we show (in Section 6.6) that this finiteness assumption is not crucial for our results.

In particular, the agent is small enough so that her acceptance/rejection decision as well as her investments do not affect prices, or the information available in the economy. This assumption rules out situations in which agents need to take into account the strategic consequences of their own actions. In this sense, our framework fits that of a competitive economy. Also, the right-hand side in the above inequality implies that the risk-free rate is zero, an assumption that could be easily modified.

3 Risk Aversion and Preference for Information

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

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

Our concept of *uniform preference for information* requires agent u_2 to accept the information purchase at all wealth levels whenever u_1 accepts it at some wealth level. That is, "agent u_2 uniformly likes information better than agent u_1" means that, whenever agent u_1 is interested in purchasing information, it is certain that agent u_2 is also interested.

Definition 1 *Agent u_2 uniformly likes information better than agent u_1 if, for every pair of wealth levels w_1, w_2, and for every information purchase a, it is the case that if u_1 accepts a at wealth w_1, then u_2 accepts a at wealth w_2.*

Alternative concepts of preferences for information and their consequences are studied in the appendix. Our first result follows:

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

1. *u_2 is uniformly less risk-averse than u_1,*

2. *u_2 uniformly likes information better than u_1.*

The formal proof of the result is in the appendix. Here we provide a verbal and elementary proof showing that (1) implies (2). As is apparent from this proof, this implication does not rely on the fact that *all* no-arbitrage investments are available, but only on the fact that *all* available investments have the no-arbitrage property. Let us assume here that the set of possible investments is an arbitrary closed set B.

Assume that u_1 accepts a at wealth w_1. For every possible signal s, let x_s be an optimal investment for u_1, given signal s and posterior beliefs q^s_α. Starting from wealth level w_1, making the information purchase a, and then making corresponding investments, is a risky bet that yields an expected utility to agent u_1 at least equal to $u_1(w_1)$. Therefore, the same risky bet, starting from wealth w_2, yields an expected utility to agent u_2 which is at least equal to $u_2(w_2)$, since this agent is uniformly less risk-averse than u_1. This does not mean, of course, that investment q^s_α after signal s is optimal for agent u_2, but it does imply that, when agent u_2 is allowed to choose investments in B, she obtains an expected utility at least as large as $u_2(w_2)$. Therefore, u_2 accepts a at w_2 as well.

The proof of the converse part is somewhat more involved than that for the direct part, as one needs to derive a conclusion about the risk-aversion levels of the agents at all wealth levels, and risk aversion is a local property. In order to prove the result in the appendix, we rely on information structures that provide "little" information, on agents who take "small" investments, and thus on situations where only local properties of utility functions matter.

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 logarithmic utility agrees to 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 must be necessarily less risk-averse.

It is appropriate to underscore the limitations of Theorem 1. Although many pairs of agents cannot be ranked uniformly by their risk aversion, many other pairs of agents still can. The result shows that, in our settings, those pairs of agents are exactly those who can be ranked according to their uniform preference for information. Indeed, the result applies in a nonvacuous way only to pairs of utility functions whose risk aversion levels can be separated by a constant: the risk aversion of u_1 for any wealth level always exceeds a constant, while for u_2, it always falls below that constant. Nonetheless, the economic environments described by our no-arbitrage condition provide a general framework in which willingness to pay for information decreases with risk aversion (see Eeckhoudt and Godfroid, 2000, and our literature review section below).

The following example illustrates the relationship between risk aversion and preference for information. We do this in a set-up in which all available investments have the no-arbitrage property. At the same time, not all investments with that property are available. The direct proof in the example that willingness to pay for information decreases with risk aversion is already involved. This is true even though the case featured by the example we consider is the simplest one that is not entirely trivial. There are only two states of nature and the information structure fully reveals the state. The difficulty in proving something so simple shows the usefulness of Theorem 1, which, of course, applies to general no-arbitrage investment sets and general information structures.

Example 1 *Suppose that a businessperson with CARA preferences is deciding whether or not to invest an amount to support a technology startup in order to do research on a prototype electric car. There are two states of the world: high (the research is successful) and low (the research fails). The payoff from not investing is* 0 *in both states. The net payoff from investing is* G *in the high state and* $-L$ *in the low state, where* $G, L > 0$ *are parameters. We consider an information structure* α *that perfectly reveals the state. Assume that the prior probability of the state being low, denoted by* p*, is such that:*

$$G(1-p) - Lp < 0,$$

or, equivalently,

$$p > \frac{G}{G+L}$$

so that even a risk-neutral agent will not invest in the absence of new information. The maximal amount that the businessperson with initial wealth w *and coefficient of risk-aversion* r *is willing to pay for* α*, denoted* I*, is independent of* w *and is given by the relationship:*

$$\exp(-rI) = (1-p)\exp(-rG) + p.$$

Hence, we have:

$$I = -\frac{1}{r}\left(\ln(1-p)\exp(-rG) + p\right).$$

If we let

$$f(r) = -\ln[(1-p)\exp(-rG) + p],$$

then

$$I = \frac{f(r)}{r},$$

but I *is then a decreasing function of* r*. To see this, note first that*

$$\frac{\partial\left(\frac{f(r)}{r}\right)}{\partial r} = \frac{f'(r)\,r - f(r)}{r^2}.$$

According to the mean value theorem, since $f(0) = 0$, there is some $y \in (0, r)$ such that

$$f(r) = f'(y)\, r;$$

and since $f(r)$ is strictly concave, $f'(y)\, r > f'(r)\, r$, and thus

$$\frac{\partial \left(\frac{f(r)}{r}\right)}{\partial r} = \frac{f'(r)\, r - f'(y)\, r}{r^2} < 0.$$

In closing the section, we observe that willingness to pay for information may not be decreasing with respect to risk aversion if some available investments do not satisfy the no-arbitrage property.

Example 2 *Consider a CARA agent with a coefficient of risk aversion r. The agent has the option not to invest, or to invest \$1 in an asset that pays \$10 with probability 0.9 (the state is good) or \$0 with probability 0.1 (the state is bad) . The investor could purchase information at a price $\mu = \$0.50$ and learn the state for sure before investing.*

The payoff to the agent who does not purchase the information and always invests is:

$$1 - 0.9 \exp(-9r) - 0.1 \exp(r);$$

the payoff to an agent who purchases the information and invests in the good state is:

$$1 - 0.9 \exp(-8.5r) - 0.1 \exp(0.5r);$$

and the payoff to an agent who does not invest is 0.

It follows that for low degrees of risk aversion such as $r < .231$, the agent's optimal strategy is to invest without purchasing the information. For medium degrees of risk aversion, $r \in [.232, 4.60]$, the optimal strategy is to purchase the information and invest in the good state. For large degrees of risk aversion, $r > 4.61$, it is better for the agent not to purchase the information and never invest.

Thus, for small to medium degrees of risk aversion, a more risk-averse agent exhibits a stronger preference for information. Notice how the proposed investment violates no-arbitrage, as the prior-evaluated expected payoff is positive. The reader can check how, in this example, our conclusion would still hold if one restores the no-arbitrage assumption, for instance, by increasing the loss in the bad state.

The example shows that the conclusion in Theorem 1, while applicable to many settings, is not universal. In particular, in cases in which information may be purchased in order to hedge against risk, for example to hear from experts in order to avoid large losses, if appropriate assets are not available in the market, more risk-averse agents may be more willing to pay for information.

4 Preference for Information and a Value for Information Purchases

In this section we propose an objective way to define the normalized value of information purchases. The approach is based on ordering preferences for 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.

Based on the uniform preference for information, as presented in the previous section, we move now to define a comparison over information purchases. The definition formalizes the idea that if an information purchase is accepted by a first agent, then any purchase that is deemed objectively more valuable should a fortiori be accepted by an agent who likes information better than the first.[11]

Definition 2 *Let a_1, a_2 be two information purchases. We say that a_1 is* uniformly more valuable *than a_2 when, given two agents $u_1, u_2 \in \mathcal{U}$ such that u_1 uniformly likes*

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

information better than u_2 and for every two wealth levels w_1, w_2, if agent u_2 accepts a_2 at wealth level w_2, then agent u_1 accepts a_1 at wealth level w_1.

Note that the requirement of two agents being ordered according to their uniform preference for information is extremely demanding. According to Theorem 1, it is equivalent to a situation in which absolute risk aversion –no matter where measured– is always greater for one agent than for the other. Since this strong requirement is assumed about the two agents in the definition, the definition itself actually requires very little. Moreover, what it does require seems very reasonable, namely, we would expect that agents with a stronger preference for information seek out more highly-valued pieces of information.

4.1 Normalized Value of Information Purchases

We now present the cardinal formula that characterizes our information ranking.

For two probability distributions p and q, the *relative entropy* from p to q, also called their *Kulback-Leibler divergence*, is an asymmetric measure of their discrepancy. It is defined by the formula:

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

It is always nonnegative, and equals zero if and only if $p = q$. It is finite, provided the support of q contains that of p; it takes the value $+\infty$ otherwise. Thus, p and q are "maximally different" when q rules out one possibility that p does not.[12]

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

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

[12]If p were the true distribution and q an approximate hypothesis, information theory would view the relative entropy from p to q as giving the expected number of extra bits that would be 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.

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.

The normalized value of an information purchase decreases with its price and increases with the relative entropy of the prior to the posterior probabilities. 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 posterior probabilities.[13]

As a benchmark for the normalized-value index, note that for all $t > 0$, $\mathcal{NV} = t$ is the index value for all information purchases which are such that, after every signal s, the relative entropy $d(p||q_\alpha^s)$ between the prior p and the posterior q_α^s is (i) equal to a constant, and (ii) priced exactly at $(1/t)$ times that constant. In particular, if the cost of such a purchase is exactly t, $\mathcal{NV} = 1$ includes all those information purchases that have a relative entropy of t from the prior to each posterior probability.

4.2 Main Result

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

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

[13] 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). Similar formulas appear also in models of rational inattention, where relative entropies are replaced by utilities of actions less information cost measured by entropy (see, e.g., Stevens, 2016; Matejka and McKay, 2014).

2.

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

Let us clarify briefly the usefulness of this theorem as a guide for action. Suppose first that an investor has calculated the normalized value of a purchase (hiring an information expert), which turns out to be $\mathcal{NV}_1$. She also happens to know that her neighbor, more risk-averse than her, has hired a different information expert. The neighbor's purchase has a normalized value of $\mathcal{NV}_2 < \mathcal{NV}_1$. Assume also that she views herself as a small agent, that she is not interested in investing if no new information arrives, and that she has access to developed financial markets. Then she learns from our theorem that she should hire the information expert.

In fact, we show in the next section that the expert hired by the neighbor will be turned down for sure by any person whose risk aversion is higher than $\mathcal{NV}_2$ everywhere. Also, that expert would be hired by those whose risk aversion is lower than $\mathcal{NV}_2$ everywhere. The expert with normalized-value $\mathcal{NV}_1$ that our investor is considering would be turned down by a smaller set of agents in the economy (those whose risk aversion exceeds $\mathcal{NV}_1$ everywhere). That expert would also be hired by a larger set of investors (with risk aversion lower than $\mathcal{NV}_1$ everywhere).

This is obviously far from a unanimous evaluation of one expert versus the other. That would be the case were they ordered by Blackwell and were they charging the same price. Nevertheless, our formula and results, which capture the information-price tradeoff for each expert, become useful benchmarks of analysis for the decision problem of investors in an environment satisfying our assumptions.

Theorem 2 relies on the relationship between demand for information and risk aversion, as established in Theorem 1. As we have already seen, it is necessary for this approach that only no-arbitrage assets be available and that uninformed agents do not make risky bets. On the other hand, when two information purchases are compared, not all assets play a role. This, in turn, implies that the assumption that all no-arbitrage purchases are available is not necessary when comparing a finite number of information purchases. See Section 6.6 for more details on this.

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

Lemma 1 *For every w and $r > 0$,*

1. *the optimal portfolio for agent u_C^r and belief q with full support is given by*

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

2. *for every q, the maximum expected utility is:*

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

The proof is elementary and provided in the appendix.

We continue with Lemma 2, which shows that $\mathcal{NV}(a)$ can be equivalently defined as the level of risk aversion of a CARA agent who 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 agent u_C^r rejects a at wealth level w.*

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

Proof. Agent u_C^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

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

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 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 $\mathcal{NV}(a_2)$ is infinite 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 $\mathcal{NV}(a_1)$ is also infinite.

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 $\underline{R}(u_2)$ is finite, $\underline{R}(u_2)$ 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. ■

5 The Demand for Information

In this section we show that, in our settings, the normalized value of an information purchase is useful for characterizing 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 greater than the normalized value of information, she rejects a purchase independently of her wealth. On the other hand, if the maximum coefficient of absolute risk aversion of an agent over all levels of wealth is smaller than the normalized value of information, she accepts a purchase independently of her wealth. The next theorem follows:

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

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

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

This simple but important result follows immediately from Lemma 2 and from the first direction of Theorem 1. This is so because 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 or not u accepts a, independent 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. It is important to note that this is true only because 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 a coefficient of risk aversion $\rho_u(w) \in [\gamma_1, \gamma_2]$ for all relevant w. Then, given an information structure α, one can identify two prices, 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: for prices $\mu > \mu_1$, the information purchase (μ, α) will be unanimously rejected, whereas for prices $\mu < \mu_2$, the purchase will be unanimously accepted. This is the sense in which, for all information structures, the index of normalized value allows one to identify the minimum and maximum prices for individuals within a group whose coefficients of risk aversion are known or at least have been estimated.

More can be said for the case when u is DARA. For DARA utilities, the next result characterizes utility functions that exhibit unanimous acceptance and unanimous rejection of a purchase. (The proof is also in the appendix).

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 $\overline{R}(u) \leq \mathcal{NV}(a)$.*

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 gauge the magnitudes implied by the index. Of course, this is meant to be only suggestive, and is far from constituting 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." People in 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 ranges from 0.36 (Italy) to 1.81 (Germany) million US$ (with the US being around 0.95). However, these figures include very young people who have not yet had time to acquire any assets. If we use the median wealth instead, the figures go from about US$ 20,000 (Sweden) to about US$ 120,000 (Italy), with the US being at about US$ 50,000. This means that a large fraction of the risk aversion 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 3 *Let a be an information purchase about a binary state of the world (e.g., whether or not the US will be in recession in 2020) 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, computation gives:*

$$\mathcal{NV}(a) = \frac{1}{\mu} \ln \left(2 \left(\beta \left(1 - \beta \right) \right)^{1/2} \right).$$

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

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

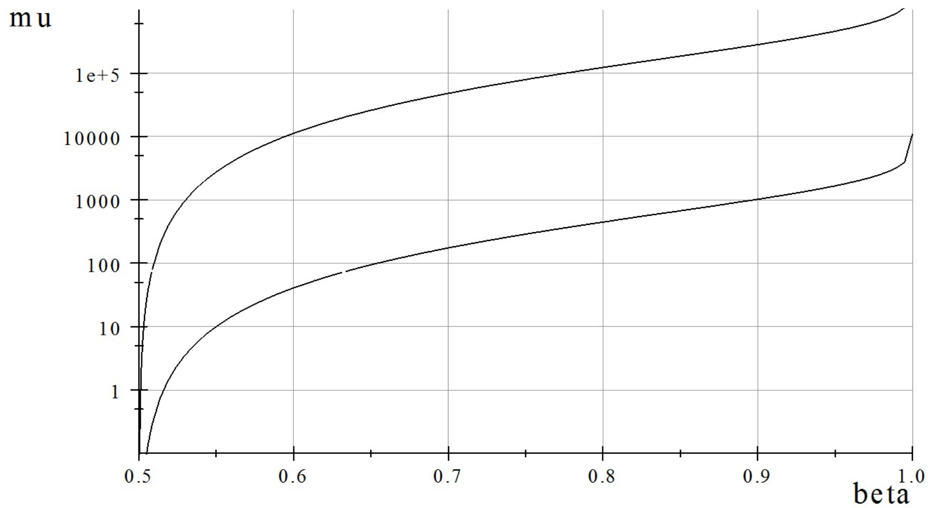

Figure 1: Maximum and minimum prices at which information is purchased by agents in the economy as a function of signal precision β

The same information purchase is rejected by all agents considered, if:

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

In the following figure, for every value of β, the information purchase is rejected by all agents considered, if μ is above the upper curve, and accepted by all agents considered, if μ is below the lower curve.

Of course, both ranges in the previous example are relatively large, since 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 provide realistic figures, given 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 minimum bound. For example, on its website, the Australian Securities and Investments Commission says this 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."[14] This fits nicely with the figures in our last example. Future research might be able to provide a more precise range for information prices as a function of risk aversion estimates and of different pieces of information, always under the assumption of frictionless financial markets that allow large investments

6 Some Properties of the Index

We now discuss some properties of our index for the normalized value of information.[15]

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," $\mathcal{NV}(a)$ is infinite if and only if 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 purchase (μ, α_2).

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

[15]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.

Thus, we have:

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

This 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 from 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 "mixed" information purchase lies between the normalized value of the most valuable one and that of the least valuable one. Here again, the comparison of the values of the normalized index stems naturally from the "more valuable than" ordering.

6.4 Comparison with Average Relative Entropy

So far we have seen that two intuitive properties of the index $\mathcal{NV}$ are that it is (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, it is interesting to compare the normalized value with the following index, which simply averages out all generated relative entropies:

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

It is apparent that this index shares those two properties, and also satisfies separability in the form of price homogeneity. The next example shows that the two indices are, however, fundamentally different. Indeed, the example highlights the essential nature (for the ordering of purchases) of the combined operation of the exponential function and its compensating logarithm as a "blow up/shrink down" of relative entropies. The exponential function with negative exponents, which is 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 property is essential in order to satisfy the duality between uniform preferences for information and the proposed function ranking the normalized value of purchases.

Example 4 *Let $K = \{1, 2, 3\}$ and fix a uniform prior. Consider, for instance, two information structures, each of which has two signals:*

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

Each row represents the probability of each signal under each possible state: 1, 2, 3.

Fix an arbitrary $\mu > 0$, and define the purchases $a_1 = (\mu, \alpha_1)$ and $a_2 = (\mu, \alpha_2)$. Note that $\hat{A}(a_1)$ is infinite 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 values. Let $r = (r_1+r_2)/2$. Clearly, the CARA agent r uniformly likes information better than the CARA agent r_2: the CARA agent r_2 accepts a_2, which according to the index $\hat{A}$ would be less valuable than a_1; but agent r, who likes information more, rejects a_1. This suggests that $\hat{A}$ is not measuring value of information well.*

6.5 Comparison with Entropy Informativeness

Several authors (Kelly, 1956; Arrow, 1971; Cabrales, Gossner, and Serrano, 2013) have proposed measuring the informativeness of a statistical experiment by the reduction of uncertainty on the state of nature, quantified by its entropy. The following example compares the rankings generated by our normalized index (when the price of the purchase is kept constant) and entropy informativeness.

Example 5 *Let $K = \{1, 2, 3\}$ with a uniform prior. Consider the following two information structures, where rows correspond to states of nature, columns to signals, and cells to the probability of the signal in the corresponding state:*

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

For ε_1 and ε_2 small enough, these information structures are not ranked according to Blackwell. To see this, note that α_1 allows one to separate states 2 and 3 with greater precision than α_2 does, while α_2 allows one to statistically distinguish between states 1 and 2, which α_1 does not.

Now let us compare the entropy informativeness for a normalized price of 1: $\mathcal{NV}(a_1)$ for $a_1 = (1, \alpha_1)$ and $\mathcal{NV}(a_2)$ for $a_2 = (1, \alpha_2)$:

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

Taylor approximations give:

$$\begin{aligned} \sum_s p_{\alpha_1}(s) \exp(-d(p \| q^s_{\alpha_1})) &\simeq \frac{2}{3}\varepsilon_1^{1/3}, \\ \sum_s p_{\alpha_2}(s) \exp(-d(p \| q^s_{\alpha_2})) &\simeq \varepsilon_2^{1/3}. \end{aligned}$$

If $\varepsilon_1 = \varepsilon_2$ and both are small enough, then $\mathcal{NV}(a_2) < \mathcal{NV}(a_1)$. That is, both purchases are becoming almost excluding at the same rate, but a_1 leads to almost eliminate one more state following the second signal, allowing the investor to make very aggressive bets against more states. On the other hand, if $\varepsilon_2 = \varepsilon_1^2$ and both are small, then $\mathcal{NV}(a_1) < \mathcal{NV}(a_2)$. Namely, because of the different rates at which signal probabilities approach 0, a_2 is becoming almost excluding much faster.

Let us now estimate the entropy reduction from the uniform prior, which we denote by $I_e(\cdot)$. Straightforward computation yields the following:

$$\begin{aligned} I_e(\alpha_1) &\simeq \ln 3 - \frac{2}{3}(\ln 2) \simeq \ln 3 - 0.46; \\ I_e(\alpha_2) &\simeq \ln 3 - \frac{1}{3}(1.1 \ln 1.1 - 0.1 \ln 0.1 + 1.9 \ln 1.9 - 0.9 \ln 0.9) \\ &\simeq \ln 3 - 0.55. \end{aligned}$$

This implies that $I_e(\alpha_1) > I_e(\alpha_2)$ whenever $\varepsilon_1, \varepsilon_2$ are sufficiently close to zero, as the generated posteriors by α_1 eliminate a greater degree of uncertainty (note the difference between the two information structures in the second state).

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

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

The two previous expressions imply that, to a first-order approximation when $q^s_{\alpha_i}$ is close to p,

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

As a result, it follows that:

$$
\mathcal{NV}(a_i) \simeq \frac{1}{\mu}\sum_s p_{\alpha_i}(s)\left(\sum_k p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right) \tag{2}
$$

and

$$
I_e(\alpha_i) \simeq \sum_s p_{\alpha_i}(s)\left(-\sum_k\left(1+\ln q^s_{\alpha_i}\left(k\right)\right)p_k\left(\ln p_k - \ln q^s_{\alpha_i}(k)\right)\right). \tag{3}
$$

A comparison of expressions (2) and (3) makes it 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 $q^s_{\alpha_i}\left(k\right)$ vectors are all parallel to the unit vector and that $\ln q^s_{\alpha_i}\left(k\right) < -1$, that is, when priors are close to uniform and there are more than two states. Otherwise, cases such as the one provided in Example (5), when posteriors are very informative, are likely to make the indices diverge.

6.6 On Smaller Investment Sets

Theorem 2, which characterizes demand for information according to our normalized index, makes the assumption that B^* consists of all no-arbitrage assets. We show here that this assumption is not necessary, as long as the set of available investments is rich enough.

Consider an information purchase $a = (\mu, \alpha)$, and assume that all posterior probabilities following α are in the interior of $\Delta(K)$. Depending on the signal $s \in S_\alpha$ received, let $x_r(s) \in B^*$ be the optimal investment for a CARA agent with parameter r given the agent's posterior belief; also let $X_r(a) = \{x_r(s), s \in S_\alpha\}$.

We now consider an arbitrary set B of no-arbitrage assets, which is not necessarily the full set B^*. All definitions, including "a_1 is more valuable than a_2" extend to the context in which acceptance of an information purchase depends on the set of available assets. The two results that follow provide counterparts to Theorem 2, while relaxing the assumptions made on the set of available assets.

Proposition 1 *Consider two information purchases a_1 and a_2, and assume that $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$. If B contains $X_r(a_1)$ for some $r \in [\mathcal{NV}(a_2), \mathcal{NV}(a_1)]$, then a_1 is more valuable than a_2 given the investment set B.*

Proof. Assume agent u_2 accepts a_2 at wealth w_2, given the available set B. Agent u_2 also accepts a_2 at w_2 if B^* is available; hence, $\underline{R}(u_2) \leq \mathcal{NV}(a_2) \leq r \leq \mathcal{NV}(a_1)$. A CARA agent with risk aversion r accepts a_1, given either B^* or B, since only investments in B are chosen by this agent. Now, any agent u_1 who likes information better than u_2 satisfies $\overline{R}(u_1) \leq r$. Hence, this agent u_1 also accepts a_1 given B. ■

The following is an extension of the converse part of our main result:

Proposition 2 *Consider two information purchases a_1 and a_2, and assume that a_1 is more valuable than a_2 given some investment set B. If B contains $B_{\mathcal{NV}(a_2)}(a_2)$, then $\mathcal{NV}(a_1) \geq \mathcal{NV}(a_2)$.*

Proof. A CARA agent with risk aversion $\mathcal{NV}(a_2)$ accepts a_2 under B. Thus, this same agent accepts a_1 under B, which implies that she also accepts a_1 under B^*, which implies that $\mathcal{NV}(a_2) \leq \mathcal{NV}(a_1)$. ■

Of course, in general, rankings of information purchases may depend on investments available. But, as shown by the previous results, as long as enough investments are available, the ranking between any two information purchases is unambiguous, and is represented by our normalized-value index.

7 Related Literature

The classical approach to ranking information structures is due to Blackwell (1953).[16] However, this approach does not provide a complete ordering of information structures. For an information structure to be more informative in the sense of Blackwell than another one, it must be the case that agents (weakly) prefer the former over the latter no matter what their preferences are. More recent research has focused on restricting preferences to a particular class. For example, Lehmann (1988) restricts attention to problems with monotone decision rules, and Persico (2000), Athey and Levin (2001), and Jewitt (2007) focus on some more general classes of monotone problems.[17] The main difference between this line of research and our approach is that we provide a complete order through a duality axiom for problems with a restricted set of investment opportunities.[18]

Working independently, several authors (Kelly, 1956; Arrow, 1971; Cabrales, Gossner, and Serrano, 2013) have studied a complete ordering of information structures, indexed by the reduction of uncertainty on the state of nature due to the signal, as measured by entropy. All three papers hinge on the fact that entropy measures the value of information for a logarithmic utility investor. Arrow (1971) simply assumes such an investor. Kelly (1956) shows that when investment oppor-

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

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

tunities are repeatedly available, the betting strategy that maximizes the growth of long-run wealth is the one that maximizes instantaneous expected logarithmic utility. Cabrales, Gossner, and Serrano (2013) measure the value of an information source by the maximal price that an agent in the economy is willing to pay to access it, and the logarithmic agent is characterized as having the maximum willingness to pay for information.[19, 20]

The present paper departs from this literature in several ways. First, we offer a way to measure not just information, but an information purchase, thus adding the price of information into the object of study. To the best of our knowledge, ours is the first index capturing this tradeoff. Second, a perhaps more important argument is that in our methodology, instead of studying the value of information to one particular agent, we obtain an index that captures *demand for* information, in the sense that larger values of the index correspond to larger sets of agents accepting the corresponding purchases.

Prior literature has investigated the relationship between risk aversion and demand for information. Freixas and Kihlstrom (1984) study an environment in which a consumer with CARA preferences decides which variety of a horizontally differentiated good to consume. The consumer can obtain a normally distributed signal that would add precision to her knowledge. The paper shows that the willingness to pay for information in that environment decreases with risk aversion.[21] Eeckhoudt and Godfroid (2000) assert at the beginning of their paper, "It is widely believed among economists and businessmen that increases in risk or in risk aversion should increase

[19]This is related to the fact that this is also the least risk-averse agent in the economy considered.

[20]Samuelson (1969) had already discussed Kelly (1956) and the properties of logarithmic utility investing, and Blume and Easley (2002) the potential for the dominance in a market in the long run 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.

[21]Alepuz and Urbano (1995) study the experimentation problem of a monopolist who is trying to learn the slope of her (linear) demand. They show that for small experimentation costs, CARA agents who are more risk-averse experiment less than do less risk-averse agents.

the value of information for decision-makers." Then they provide an example of a decision problem in which the optimal decision in the absence of information is not to invest, and in which the value of information falls with increasing risk aversion. We view our work as providing a general framework for thinking about the demand for information, and thereby extending further the logic of their example.

Finally, our approach to ranking information purchases is based on a ranking of preferences for information. Relatively few papers in the literature deal with the comparison of different agents' preferences for information. One such study is Grant, Kajii, and Polak (1998), which explores intrinsic preferences for information, that is, preferences that are unrelated to the fact that information can lead to more profitable decisions. This is very different from our framework, since our agents like information precisely because it helps them to make better decisions. But, interestingly, just as we found (see 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.

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 (as in Aumann and Serrano (2008)) between value and preference for information in settings in which the investment opportunities are described by a no-arbitrage condition. Because no-arbitrage assets provide a clean way to measure the value of information, we have been able to extend their use from such previous studies as (Kelly, 1956; Arrow, 1971; 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. Our 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. In this way we can describe a useful inverse demand curve for information in no-arbitrage investment settings.

References

Alepuz, D., and A. Urbano (1995): "A note on risk aversion and learning behavior," *Economics Letters*, 49(1), 51–58.

Arrow, K. J. (1971): "The value of and demand for information," in *Decision and Organization*, ed. by C. B. McGuire, and R. Radner, pp. 131–139, Amsterdam. North-Holland.

Athey, S., and J. Levin (2001): "The value of information in monotone decision problems," Stanford University, 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.

Bessembinder, H., M. Panayides, and K. Venkataraman (2009): "Hidden liquidity: an analysis of order exposure strategies in electronic stock markets," *Journal of Financial Economics*, 94(3), 361–383.

Blackwell, D. (1951): "Comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102. University of California Press.

——— (1953): "Equivalent comparison of experiments," *Annals of Mathematical Statistics*, 24(2), 265–272.

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 (2012): "The appeal of information transactions," Universidad Carlos III, Department of Economics Working Paper 12-24.

——— (2013): "Entropy and the value of information for investors," *American Economic Review*, 103(1), 360–377.

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

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(3), 522–550.

Eeckhoudt, L., and P. Godfroid (2000): "Risk aversion and the value of information," *The Journal of Economic Education*, 31(4), 382–388.

Freixas, X., and R. Kihlstrom (1984): "Bayesian Models in Economic Theory," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, chap. 6, pp. 93 – 104. Elsevier, New York.

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.

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

Grant, S., A. Kajii, and B. Polak (1998): "Intrinsic preference for information," *Journal of Economic Theory*, 83(2), 233–259.

Hart, S. (2011): "Comparing risks by acceptance and rejection," *Journal of Political Economy*, 119(4), 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.

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

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.

Landau, L., and E. Lifshitz (1980): *Statistical Physics*, vol. 5 of *Course of Theoretical Physics*. Butterworth-Heinemann, Oxford.

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.

MacLean, L. C., E. O. Thorp, and W. T. Ziemba (2011): *The Kelly Capital Growth Investment Criterion: Theory and Practice*, vol. 3. World Scientific.

Matejka, F., and A. McKay (2014): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *The American Economic Review*, 105(1), 272–298.

McMillan, B. (1956): "Two inequalities implied by unique decipherability," *IEEE Trans. Information Theory*, 2(4), 115–116.

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

Persico, N. (2000): "Information acquisition in auctions," *Econometrica*, 68(1), 135–148.

Samuelson, P. A. (1969): "Lifetime portfolio selection by dynamic stochastic programming," *The Review of Economics and Statistics*, 51(3), 239–246.

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(3), 375–383.

Stevens, L. (2016): "Coarse pricing policies," University of Maryland.

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

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 Proof of Theorem 1

(1) implies (2)

We start by providing a formal proof that (1) implies (2). 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:

$$\frac{u_1(w_1+z) - u_1(w_1)}{u_1'(w_1)} \leq \frac{u_2(w_2+z) - u_2(w_2)}{u_2'(w_2)}. \tag{4}$$

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.[22]

(2) implies (1)

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

Lemma 3 *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:

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

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

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

[22]This statement follows from 4.1.2. in Aumann and Serrano (2008).

and

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

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 4 *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:

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

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

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

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 3 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

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

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

Lemma 5 *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 4. 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 4 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 4 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 6 *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 5, 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. ■

Therefore, to prove the converse statement in the theorem, 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_2 is less risk-averse at w_2 than u_1 is at w_1, which is a local property at w_1, w_2, the proof relies on information structures $\alpha(\varepsilon)$, which are "little informative", hence induce small investments. Lemma 6 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 $\overline{R}(u_2) \leq \underline{R}(u_1)$. The proof is thus complete.

B Proof of Lemma 1

Part 1: 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

$$q_k \exp(-r x_k) = \lambda p_k,$$

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

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

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

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

and hence, the result.

Part 2: First, assume that q has full support; hence, $d(p\|q)$ is finite. Using the optimal-portfolio characterization in Part 1 of the current lemma, 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

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

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 $V(u_C^r, w, q) = 0$.

C 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.

■

D Further Justifications of the Index

D.1 Other 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 2, 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 $u_1, u_2 \in \mathcal{U}^I_{DA}$, 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 2.

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

Definition 3 *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 4 *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 $u_2 \in \mathcal{U}^I_{DA}$, 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^I_{DA} 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. ■

D.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 5 *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:[23]

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*
- *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

[23] 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).

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. ■

E Additional Material Concerning Properties of the Index

E.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 2 that $\mathcal{NV}(\mu, \alpha_1) \geq \mathcal{NV}(\mu, \alpha_2)$. ■

E.2 Proof of Observation 2 (Mixtures)

Proof. Fix any wealth level. From Lemma 2, a CARA agent with coefficient of risk aversion $\mathcal{NV}(\mu, \alpha_2)$ accepts both purchases (μ, α_1) and (μ, α_2) at any 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 Lemma 2, 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

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

■

E.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 she 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:

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

where q_α^s 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.[24]

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

Our Definition 2 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:

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

As a word of caution, we note that in the formula above, as $(q_\alpha^s)_s$ 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).

E.4 Sequential Purchases

Another property worth mentioning concerns the normalized value of an information purchase in which the buyer receives signals sequentially from 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 (1), 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.