A Theory of Value for Information

Patrick Beißner*

May 11, 2016

Abstract

This paper suggests a formal framework to establish a convex preference relation for information. We introduce a new commodity–price duality, when information is the commodity. The fineness and precision of payoff–relevant signals quantifies information. All candidate prior beliefs that become inconsistent with additional consumption of information can be excluded and lead to an increase of utility. A heterogeneous agent economy is introduced, in which information can be traded, based on individual demand for information. In this context, once information is sold, it remains the property of the seller. In a final step, we show that this information–trade equilibria exist.

Key words and phrases: Information as a commodity, reduction of ambiguity, price for information, concavity of utility from information, equilibrium

JEL subject classification: D08

1 Introduction

"The most valuable commodity I know of is information."
Gordan Gekko in the movie Wallstreet (1987)

Rather than a movie quotation, we might instead consider Arrow (1996): *"Information is an economic good, in the sense that it is considered as valuable and costly."* In other words, scarcity and the heterogeneous distribution of information can result in the willingness to pay for it. The present paper aims

*Research School of Economics– The Australian National University, Canberra, Australia. Email: `patrick.beissner@anu.edu.au`. I thank Simon Grant, Ali Khan and Rabee Tourky for many fruitful discussions.

to present a formal model of an economy where information, as a tradable commodity, receives its value from an equilibrium price.

Following Radner (1968) and Aumann (1974), we define information as a special system of events, a σ-algebra. A larger or finer σ-algebra gives the decision maker a more precise conditional expectation of any evaluated signal or payoff. For instance, the coarsest σ-algebra reveals only the expected value of the signal. Given this uncommon type of commodity, the traditional microeconomic–modeling principle suggests beginning with the introduction of an appropriate commodity space for information. The usual consumption set, a set of σ-algebras, consists of a set of sets of sets. We term this a *3-stage* commodity space. A standard commodity space for contingent claims would then be of *1-stage* type. This distinction becomes immediately relevant in light of the basic task of finding a meaningful notion of closeness between two information commodities.

The standard approach to modeling information as a commodity considers a 3-stage consumption set. Similar attempts can be found in Allen (1986), Cotter (1986), Stinchcombe (1990), Van Zandt (1993) and Khan, Sun, Tourky, and Zhang (2008). We have nothing to contribute concerning such a modeling approach given that each notion of closeness possesses several advantages.[1] In the same vein, Allen (1990) presents a detailed account of information as a commodity. In contrast, our analysis relies on the following assumption: *Every σ–algebra is generated by a set of random variables.*

This assumption is rather weak but has several advantages. First, from a technical stance, we reduce our analysis of commodity spaces from *stage 3*–type to *stage 2*–type, since only the index set of generating random variables matters. Second, each random variable can now be interpreted as knowledge about the distribution of a signal. The resulting information is then the awareness about all events that stands in relation to these signals. A natural example of a stage 2–type commodity is land (see Berliant (1986)). It turns out that in many cases the above stage 2 assumption for information holds automatically. For instance, as explained in Example 2, if uncertainty is based on a Brownian motion, this assumption is already satisfied.

In the model presented in this paper, additional (and payoff relevant) information is used entirely to exclude possible priors about the uncertain future. To establish a utility functional on the consumption set of information,

[1] Alternative approaches are Gilboa and Lehrer (1991); Grant, Kajii, and Polak (1998).

we must impose an ambiguity attitude for the agent. We focus here on a worst–case expected utility as axiomatized by Gilboa and Schmeidler (1989). The new twist lies in modeling information as a device to discipline ambiguity. Formally, this is quantified by the *information–ambiguity* (IA) correspondence. The way imprecise information generates a set of possible beliefs is to some extent related to work by Chateauneuf and Vergnaud (2000), Gajdos, Hayashi, Tallon, and Vergnaud (2008) and Gul and Pesendorfer (2015).

In contrast to Radner and Stiglitz (1984) and Chade and Schlee (2002) we model the quantification of utility from information differently. Consequently, utility from additional information is indeed *concave*. Moreover, an increase in available information results in a decrease in the degree of ambiguity about the true probability law that describes the possible distributions of signals. The updated set is the value of the IA-correspondence. Under the assumption of ambiguity aversion, we thus obtain a utility improvement. This remains in line with Blackwell (1953), where under standard assumptions, a more precise information system is always preferred to one that is less so (see also Hervés-Beloso and Monteiro (2013)).

Based on such a well-behaved commodity space and utility for information, we follow some classical steps from demand theory. Budget and demand correspondences for information have similar analytic properties to their classical counterparts. We take into account that doubling the same information yields no improvement. This fact also changes the nature of information allocations when compared with the usual concept of feasibility.

In a final step, we establish a general equilibrium existence result, where information is traded the same as contingent claims, i.e. contracts are closed today and delivery takes place at a later point in time. However, the nature of information allows selling them to more than one agent. The value of information is usually defined as the increase of utility with respect to the best action.[2] Here, we again follow the tradition of general equilibrium theory, so that the *value of information* receives its foundation through the endogenous equilibrium price system. Apart from existence, we show that equilibrium prices allow for a representation in terms of an information–price density.

The paper is organized as follows. In Section 2 we present a leading and re-

[2] In the present model, such an approach is odd as additional information comes at a monetary cost from buying and selling information in the market for information. Consequently, this affects the best action, due to wealth changes.

peatedly emerging example. Section 3 begins with the formal description of the model and introduces the basic consumption set. Section 4 formulates the relation between information and falsifiability of priors via the IA–correspondence. Then the resulting utility from information is specified. Before we come to the notion and existence of an information–trade equilibrium in Section 6, we discuss standard concepts of price systems, budget and demand correspondences for information. Section 7 concludes, and the appendices presents the proofs.

2 Unveil the Ellsberg Urn

This section presents a thought experiment based on an Ellsberg urn. It serves as the leading example and helps to clarify the meaning of (*i.*) information as a commodity, (*ii.*) the information–ambiguity correspondence and (*iii.*) utility from information. The sequence of examples 3, 4, 5, 7 and 8 are built on each other, and continue the discussion of the present section.

Suppose an agent is confronted with a gamble based on a two–color urn of 12 balls with an equal number of red and green balls. The distribution is perfectly known to the agent and she is free to choose and accept one or none of the following gambles: If green is drawn, the agent receives \$10 and \$0 otherwise, or the same payoffs with changed colors.[3] Suppose the risk averse agent agrees to play the gamble for not more than \$4.

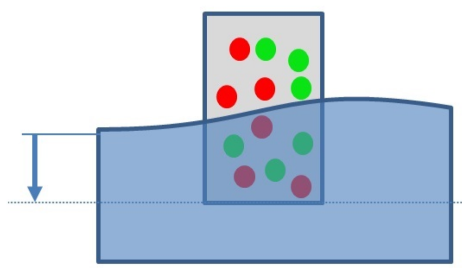

Figure 1: The content behind the veil of ignorance describes a specified gamble. *How much would you pay for a partial revelation?*

Now assume the agent is in a different situation. She knows the precise color of every second ball and that only red or green balls are in the urn. Moreover, she is aware of the fact that information about one half of the urn yields no

[3]At the present state of the thought experiment, the payoff on the respective color seems and is obsolete. In the second part, this issue is of some importance.

advantage to estimating the distribution of colors in the unknown part of the urn. In this new situation, the agent owns obviously *less* information about the gamble than before. Consequently, she agrees to play the gamble for not more than \$3, because of (moderate) ambiguity aversion or pessimism, and selects the color on which she can win 10\$ through a random device, i.e. a fair coin. Ambiguity aversion is a behavioral attitude related to the presence of multiple priors describing the gamble.

Based on this new situation and as illustrated in Figure 1, suppose there is an opportunity to reveal additional components of the hidden urn. The agent would pay to receive additional information about the distribution of balls. From this thought experiment we see that information about the unknown components of the urn has a positive value. The certainty equivalent for the gamble, when the contents of the urn is perfectly known, delivers an upper bound for the reservation price of full information. For the given gamble, \$4–\$3 is then the indifference price for knowledge about the contents behind the veil.

The formal consideration for the following section relies on the idea of information, where the belief of the agent is contingent on what she knows. Several possible distributions of the 12 balls can only be excluded if additional information about the urn is accessible.

Remark 1 *Information about a priori observed balls is regarded as a collection of random variables and not as the realization of signals. For simplicity it is assumed that the probability of each observed ball color is zero or one. The balls behind the veil could be non–trivial random variables. The additional information about these balls is then the exact probability of their color.*

3 Information as a Commodity

This section introduces the basic framework. For the rest of the paper we make the following

Standing Assumptions:

1. Information is modeled in terms of σ-algebras.
2. Every σ-algebra is generated by a set of random variables.[4]

[4]Recall, a random variable $X : \Omega \to \mathbb{R}$ generates a σ-algebra, denoted by $\sigma(X)$, through the collection of inverse images: $\sigma(X) = \{X^{-1}(A) \subset \Omega : A \in \mathcal{B}(\mathbb{R})\}$. Equivalently, $\sigma(X)$ is the smallest σ-algebra where X is measurable.

Any agent in the model can only consider those information structures which are generated by an (arbitrary) collection of observable distributions of signals. On the modeling side, this simplifies many technical difficulties, since the magnitude of complexity is reduced by moving from systems of set-systems to systems of sets (as mentioned in the Introduction from *stage 3* to *stage 2*). As indicated in Remark 1, information about the realization of a signal is not part of the model.

3.1 Information and Uncertainty

Fix a probability space $(\Omega, \mathcal{H}, \mathbb{P})$. $\mathbb{P} : \mathcal{H} \to [0,1]$ denotes the objective probability measure about an uncertain future. Let $\Delta(\Omega; \mathcal{H})$ be the set of all probability measures on the measurable space $(\Omega, \mathcal{H})$. We assume that the finest σ-algebra $\mathcal{H}$ is generated by a set of real–valued random variables $X_{\mathtt{h}} : \Omega \to \mathbb{R}$, indexed by a metric space H. The Borel σ-algebra of H is denoted $\mathcal{B}(H)$. We then have $\mathcal{H} = \sigma(X_{\mathtt{h}} : \mathtt{h} \in H)$.

Nevertheless, we may also take a discrete space $H = \{0, \ldots, N\} \subset \Omega = \mathbb{R}^N$, with $N \in \mathbb{N}$, as discussed in the following example.

Example 1 *Let $\mathbb{P}$ be the model of n uncorrelated real–valued random variables. Suppose a normal distribution with zero mean that is $\mathbb{P} = N(0, \Sigma)$, where the $n \times n$ covariance matrix Σ is the identity matrix. For instance $X_1 \sim N(0,1)$ refers to a one–dimensional random variable with a standard normal distribution. $\mathcal{F}_1 = \sigma(X_1)$ is then the knowledge about the distribution of the first component. The finest information structure is $\mathcal{H} = \sigma(X_{\mathtt{h}} : \mathtt{h} \in H)$. Analogously, every subset $\{X_{\mathtt{h}}\}_{\mathtt{h} \in F}$ of random variables, with $F \subset H$, describes a partial knowledge $\mathcal{F} = \sigma(F)$ of $\mathcal{H}$.*

A further example is based on a Brownian motion and the strong connection between closed subspaces, sub σ-algebras and conditional expectations.

Example 2 *Let $(C_0[0,1], \mathcal{H}, \mathbb{P}_0)$ be the Wiener space so that the canonical process is the Brownian Motion (W_t). $\mathcal{H}$ denotes the Borel σ-algebra that is induced by the usual* sup*-norm of $C_0[0,1]$ – the space of continuous paths on $[0,1]$ starting in zero. In this case, we have $\mathcal{H} = \sigma(W_{\mathtt{h}} : \mathtt{h} \in H)$, where $W_{\mathtt{h}} = \int_0^1 \mathtt{h}_s \mathrm{d}W_s$ is the stochastic Itô integral with deterministic integrand $\mathtt{h} \in H$ and is normally distributed, i.e. $W_{\mathtt{h}} \sim N(0, \|\mathtt{h}\|_H)$. Here, H denotes the space of square integrable functions on $[0,1]$, that is $\|\mathtt{h}\|_H = \left(\int_0^1 \mathtt{h}_s^2 \mathrm{d}s\right)^{\frac{1}{2}} < \infty$.*

Any subset $F \in \mathcal{B}(H)$ generates a coarser σ–field than $\mathcal{H}$. For instance, choosing for every $s \in (0,1)$, the closed sub vector–space $F_s = L^2([0,s])$ results *in* $\mathcal{F}_s = \sigma(W_r : r \in [0,s])$ *the information filtration generated by observing the Brownian motion itself up to time* s*, where* $W_r = W_{\mathtt{h}}$ *with* $\mathtt{h} = 1_{[0,r]}$.

From Example 2 we see that the present approach to model information is consistent with continuous-time modeling.

Given $(\Omega, \mathcal{H}, \mathbb{P})$, we turn to the formal description of some agent being equipped with some initial information endowment $\mathcal{F} = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F)$, or in abuse of notation $\mathcal{F} = \sigma(F)$, such that $F \in \mathcal{B}(H)$. As illustrated especially in Example 2, the index set of random variables allows us to identify every relevant sub-information structure $\mathcal{F}' \subset \mathcal{H}$, by a Borel-subset $F' \in \mathcal{B}(H)$.

3.2 Commodity Space of Information

A *pure* information commodity is an element $F \in \mathcal{B}(H)$, that is the perfect knowledge $\sigma(X_{\mathtt{h}} : \mathtt{h} \in F)$ about the distributions of all random variables indexed by F. Clearly, we can write this set in terms of an indicator function $1_F : H \to \{0,1\}$. The boundedness of such functions delivers an adequate commodity space $L^\infty_H := L^\infty(H, \mathcal{B}(H), \mu)$ of bounded measurable elements, where μ is a given positive measure on $(H, \mathcal{B}(H))$. Whenever $H \subset \mathbb{R}^M$, for some $M \in \mathbb{N}$, we consider the Lebesgue measure $\mu = \lambda$ on H. In abuse of notation, we have[5]

$$2^{\mathcal{H}} \subset [0, 1_H] \subset L^\infty_H.$$

As we motivate and explain in Section 4, the closed, bounded and convex set $[0, 1_H] = [0,1]_{L^\infty_H} = \{f \in L^\infty_H : 0 \le f(\mathtt{h}) \le 1 \ \mu\text{–}a.e.\}$ defines the *consumption set.* $A \in \mathcal{B}(H)$ holds $\mu - a.e.$ if $\mu(A^c) = 0$, where $A^c = H \setminus A$.

A function $f \in [0, 1_H]$ allows for several representations as limits of simple functions, that is $f(\mathtt{h}) = \lim_N \sum_{0 \le k \le N} a_k 1_{A_k}(\mathtt{h})$. We focus here on the unique representation of pairwise disjoint collections $\{A_k\}_{k \le N} \subset \mathcal{B}(H)$ and corresponding factors $a_k \in [0,1]$.

As a primitive of the economy the given information $\mathcal{F}$ of some agent is coarser than $\mathcal{H}$, where $\mathcal{F} = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F)$ and $F \in \mathcal{B}(H)$. The set of pure and desirable information commodities is $\mathcal{B}(F^c)$. Given the agent's information F, she considers the commodity $F' \in \mathcal{B}(H)$ only as relevant if $F' \nsubseteq F$. In this case the set F' can be purified (relatively to F) by considering only nontrivial sets $F' \setminus F$ in $\mathcal{B}(F^c)$. The next section discusses other (non pure) functions in the consumption set $[0, 1_H]$ that model a notion of non–perfect information.

[5]The first inclusion stems from $\mathcal{F} \subset \mathcal{H}$ and the assumption to consider only sub σ-algebras of the form $\mathcal{F} = \sigma(F)$. The defining index set F of generating random variables is identified by the associated function $1_F \in [0, 1_H]$.

4 Utility from Information

Before utility from information can be quantified, we need to clarify how additional information reduces the degree of ambiguity about the true prior $\mathbb{P} : \mathcal{H} \to [0,1]$. The agents' belief $\mathbb{P}_F : \sigma(F) \to [0,1]$ depends on her initial information σ-algebra $\mathcal{F} = \sigma(F)$.[6] Those events $E \notin \sigma(F) = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F)$, which she is not aware of, are not assigned with a probability. F is the index set of random variables, whose probability laws is known to the agent. Hence, her belief $\mathbb{P}_F : \mathcal{F} \to [0,1]$ is contingent on her σ-algebra, and can only capture probabilities about events that are in her information set. We assume that the restriction of $\mathbb{P}$ to $\mathcal{F}$ coincides with $\mathbb{P}_F$. In view of Section 2, this corresponds to the knowledge of the colors of visible balls.

4.1 Information–Ambiguity Correspondence

Information is a device to exclude alternative priors. These alternatives are, without additional information, reasonable probability measures on $\mathcal{H}$. Such a perspective parallels the thought experiment of Section 2: additional information about the color of the balls allows the agent to exclude prior that were plausible without some additional information commodity.

Suppose an agent starts her consideration with an a priori given information $F \in \mathcal{B}(H)$ (the index set of observable random variables). Given the information endowment $\mathcal{F} = \sigma(F)$, this results in the set of possible priors[7] on $\mathcal{H}$

$$\mathscr{P}_F \equiv \{P \in \Delta(\Omega; \mathcal{H}) : \; P = \mathbb{P}_F \text{ on } \sigma(F)\}\,. \tag{1}$$

We have $\mathscr{P}_\emptyset = \Delta(\Omega; \mathcal{H})$. The set $\mathscr{P}_F$ contains all priors on the finest σ-algebra $\mathcal{H}$ that are consistent with the given coarser information $\mathcal{F} = \sigma(F)$ and the given belief $\mathbb{P}_F$ on $\mathcal{F}$.

The following definition clarifies how additional information updates the set of possible priors. Apart from the intuitive appeal to use information as a device to exclude alternatives, the advantage of the approach relies on the quantification of consequences from receiving information.

[6] Here, we deviate from Aumann (1974) and assume that the subjective belief of the agent is only defined on her private σ-algebra.

[7] "$P = \mathbb{P}$ on $\mathcal{F}$" means $P(E) = \mathbb{P}(E)$ for all events $E \in \mathcal{F}$. For technical reasons, we also assume $P \ll \mathbb{P}$ and $\frac{dP}{d\mathbb{P}} \in L^2$. $P \ll \mathbb{P}$ means that P is absolutely continuous with respect to $\mathbb{P}$, i.e. if $\mathbb{P}(A) = 0$ then $P(A) = 0$. Note that, $A \in \mathcal{H} \setminus \mathcal{F} = \sigma(F^c)$ implies $\mathbb{P}_F(A) = 0$. The square integrability of the Radon Nykodym derivative $\frac{dP}{d\mathbb{P}}$ is a simplifying technical condition.

Definition 1 *The* Information–Ambiguity (IA) correspondence $\mathcal{P}_F : [0, 1_H] \Rightarrow \mathscr{P}_F$ *is defined by two updating rules:*

1. Reduction of consistent extensions: *for every* $A \in \mathcal{B}(H)$ *we have*

$$\mathcal{P}_F(1_A) \equiv \big\{P \in \mathscr{P}_F : P = \mathbb{P} \text{ on } \sigma(F \cup A)\big\}.$$

2. Precision of information for A*: For every* $a \in [0, 1]$ *we have*[8]

$$\mathcal{P}_F(a1_A) \equiv a\mathcal{P}_F(1_A) + (1 - a)\mathcal{P}_F.$$

The factor a in the second part denotes the *precision of information* (PI) for A. A small PI yields a large value of the IA–correspondence. From Definition 1 it directly follows that full information means knowledge about the true prior:

$$\mathbb{P} = \mathcal{P}_F(1_{H\setminus F}). \tag{2}$$

With notation $\mathcal{P}_F(a1_A) = \mathcal{P}_{FaA}$, the two rules of Definition 1 suffice to deduce *any* evaluation of the IA–correspondence. Let us begin with a result for the reduction of extensions for (imprecise) information bundles A, B with $A \cap B = \emptyset$. Set $a^- = 1 - a$.

Lemma 1 *For all* $f = a1_A + b1_B \in [0, 1_H]$*, the IA–correspondence is given by*

$$\mathcal{P}_F(f) = ba^-\mathcal{P}_{FB} + b^-a\mathcal{P}_{FA} + ba\mathcal{P}_{FBA} + b^-a^-\mathcal{P}_F. \tag{3}$$

Note that (ba^-, b^-a, ba, b^-a^-) is a probability weight on the power set of $\{A, B\}$. If $A \cap B \neq \emptyset$, then (3) incorporates a weighting on the power set of $\{A, A \cap B, B\}$.

Definition 1 and Lemma 1 focus on rather simple information structures. A straightforward extension considers the case where $f = \sum_{k \in N} a_k 1_{A_k}$ is defined by a finite sum of indicator functions.[9] In this case, (3) results in a sum over the power set of the mutually disjoint collection $\{A_1, \ldots, A_N\}$. As presented in Appendix A, every $f \in [0, 1_H]$ yields a well–defined set $\mathcal{P}_F(f) \subset \Delta(\Omega; \mathcal{H})$.

Several remarks on the meaning of the IA–correspondence are worth making. Some simple cases clarify the idea and intuition behind the required algebraic manipulations in Definition 1.

1. Simple reduction: In the case of $f = 1_A$, the IA–correspondence means full information about the set of random variables indexed by A. The representation

[8] Minkowski sums of sets are defined by $C + D = \{c + d : c \in C, d \in D\}$.

[9] Note that the collection of such simple functions is a dense subset of $[0, 1_H]$ with respect to the norm $\|f\|_\infty = \inf\{M : |f| \leq 0 \; \mu - a.e\}$ on L^∞_H.

of the IA-correspondence is consistent with the extension that is considered in De Castro and Yannelis (2010).

2. Precision for information: The IA–correspondence incorporates all values between the extreme case $a = 0$ –no precision $\mathcal{P}_F(0 \cdot 1_A) = \mathscr{P}_F$ and $a = 1$ –full precision $\mathcal{P}_F(1 \cdot 1_A) = \mathscr{P}_{F \cup A}$. In view of Section 2, the meaning of the PI refers to the additional possibility of increasing the transparency of certain parts of the veil that hides the urn. Knowledge about the color of a ball is then revealed with precision a. Specifically, the agent may receive information about a so-far veiled ball K. But she only infers that K is red with probability $\mathbb{P}(K = \text{“}red\text{”}) \geq a$.[10]

3. Reduction for information bundles: In the case of full precision, we get $\mathcal{P}_F(1_A + 1_B) = \mathcal{P}_F(1_{A \cup B})$. In view of the bundle $a1_A + b1_B$, we see in (3) that a high PI for both sets results in a high weight ba of the smallest set of priors $\mathcal{P}_{FAB} = \mathcal{P}_F(1_{A \cup B})$.

4. Unique representation by pairwise disjoint sets: If $A \cap B \neq \emptyset$ and $a = b = \frac{1}{2}$, then the extension of Lemma 1 applies only to the unique representation $f = \frac{1}{2}(1_{A \setminus B} + 1_{B \setminus A}) + 1_{A \cap B}$. Every $f \in [0, 1_H]$ allows for a μ–a.e. unique representation in terms of pairwise disjoint sets. The informational content of f is then exclusively displayed by this decomposition.

We continue with an example that applies the IA–correspondence $\mathcal{P}_F$ to the setting of Section 2.

Example 3 *Coming back to the Ellsberg urn thought experiment in Section 2, we see that every probability p_R of the event to draw a red ball lies between $[\frac{3}{12}, \frac{12-3}{12}] = [\frac{1}{4}, \frac{3}{4}]$. This set of probabilities corresponds to*

$$\mathscr{P}_F = \prod_{1 \leq k \leq 6} \{\delta_{c_k}\} \times [0,1]^6 \subset \Delta\left(\{R, G\}^{12}\right),$$

where δ_{c_k} denotes the Dirac measure for the k-th ball to have the color $c_k \in \{R, G\}$. If a 7th and 8th ball is revealed, say one is red and the other green, then p_R can only lie in the strictly smaller range $[\frac{4}{12}, \frac{8}{12}] \subset [\frac{1}{4}, \frac{3}{4}]$, which now corresponds to $\mathcal{P}_F(1_{\{7,8\}})$, see Definition 1.1.

Example 4 is continuous with Example 3 and extends it to the case of imprecise information a. In this case, only the transparency (parametrized by a)

[10] However, there are two possible interpretations for a. On the one hand the source itself announces the imprecision of the signal. An alternative viewpoint can be that the source of information does not reveal any additional information about its precision. In that case, the agent has a belief about the precision that is again captured by a. Remark 4 continues with this discussion.

of the veil is increased.

Example 4 *Now suppose that the 7th and 8th balls are revealed only partially. That is with transparency or reliability* $a = 0.6$*. Again, one ball is red and the other is green, then* $\mathbb{P}(c_7 = \text{"red"}), \mathbb{P}(c_8 = \text{"green"}) \geq 0.6$*. The probability* p_R *of drawing a red ball lies then in the range* $[0.3, 0.7] = [\frac{3.6}{12}, \frac{8.4}{12}]$*, which now corresponds to* $\mathcal{P}_F(0.6 \cdot 1_{\{7,8\}})$*. The information about the 7th ball yields an increase from* $\frac{3}{12}$ *to* $\frac{3,6}{12}$ *for the lower bound of the interval. In view of Definition 1.2, the heuristic about the role of the PI* $a \in (0,1)$ *is consistent, since* $[\frac{3.6}{12}, \frac{8.4}{12}] = 0.6 \cdot [\frac{4}{12}, \frac{8}{12}] + (1 - 0.6)[\frac{3}{12}, \frac{9}{12}]$.

The IA–correspondence is a basic tool to define a preference relation for information. In preparation, the following proposition lists properties that will be essential to the resulting utility representation.

Proposition 1 *The IA–correspondence* $\mathcal{P}_F : [0, 1_H] \Rightarrow \mathscr{P}_F$ *is*[11]

1. compact– and convex–valued.
2. upper hemi-continuous.
3. monotone shrinking: *If* $f \leq g$ *then* $\mathcal{P}_F(f) \supseteq \mathcal{P}_F(g)$.
4. convex: *For every* $\alpha \in [0,1]$ *and* $f, g \in [0, 1_H]$ *we have* $\mathcal{P}_F(\alpha f + (1-\alpha)g) \subseteq \alpha\mathcal{P}_F(f) + (1-\alpha)\mathcal{P}_F(g)$.

From the proposition, we obtain that the true prior $\mathbb{P}$ is always contained in the IA–correspondence.

Corollary 1 *For all* $f \in [0, 1_H]$*, we have* $\mathbb{P} \in \mathcal{P}_F(f)$.

4.2 The Functional Form of Utility from Information

Based on the IA–correspondence, we are now in a position to define a utility functional on $[0, 1_H] = [0, 1]_{L^\infty_H}$. The worth of additional information is quantified by the ability to exclude priors that are ex-ante consistent but ex-post (with additional information) inconsistent.

Utility from information relies on the idea that the reduction of ambiguity increases expected utility, when evaluated at some random utility $\mathtt{U}$. For some

[11] Compactness refers to the weak topology on the set of square-integrable Radon-Nykodym densities in $L^2 = L^2(\Omega, \mathcal{H}, \mathbb{P})$. The continuity of $\mathcal{P}_F$ is with respect to the norm topology $\|\cdot\|_{L^\infty_H}$ on $[0, 1_H]$ and the weak* topology $\sigma(L^{2^*}, L^2)$ on $\mathscr{P}_F$.

utility index $u : \mathbb{R} \to \mathbb{R}$ and a given and $\mathcal{H}$-measurable[12] endowment $E : \Omega \to \mathbb{R}$, we may have $\mathtt{U} = u(E)$.

Definition 2 *Fix the agent's given information $\sigma(F) = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F) = \mathcal{F}$ and let $\mathtt{U} \in L^2(\Omega, \mathcal{H}, \mathbb{P})$ be a given random utility. The* utility from additional information *$f \in [0, 1_H]$, $f : H \to [0, 1]$, is given by*

$$U(f) = \underbrace{\min_{P \in \mathcal{P}_F(f)} E^P[\mathtt{U}]}_{\text{with new information}} - \underbrace{\min_{P \in \mathscr{P}_F} E^P[\mathtt{U}]}_{\text{only old information}} \geq 0. \tag{4}$$

The IA–correspondence $f \mapsto \mathcal{P}_F(f)$ is that of Definition 1, for $\mathscr{P}_F$ see (1). The second summand of $U(f)$ in (4) is a normalization; If some information commodity f contains no additional information, i.e. $f \leq 1_F$, this implies $U(f) = 0$ by the definition of the IA–correspondence. In this case, we have $\mathcal{P}_F(f) = \mathscr{P}_F$.

The properties of the IA–correspondence, stated in Proposition 1, deliver in turn a list of important properties for the utility from additional information. The following result lists clear counterparts of the usual utility specifications for standard commodities.

Theorem 1 *Let $f, g, f_n \in [0, 1_H]$. Utility from information $U : [0, 1_H] \to \mathbb{R}$ is*

1. *monotone:* *If $f \leq g$ then $U(f) \leq U(g)$.*
2. *continuous:* *If $f_n \to f$ in $\|\cdot\|_\infty$ then $\lim_n U(f_n) = U(f)$.*
3. *concave:* *For every $\alpha \in [0, 1]$ we have*

$$\alpha\, U(f) + (1 - \alpha)\, U(g) \;\leq\; U\big(\alpha f + (1 - \alpha) g\big).$$

In view of the usual non–concavity for the value of information, such as considered in Radner and Stiglitz (1984), the third part of the theorem is most remarkable. It means in particular that ambiguity aversion implies a preference for diversification of information.

Example 5 applies Theorem 1.1 to the Ellsberg urn from Section 2.

Example 5 *Let there be two gambles X^g, X^r on the urn.* \$16 *is paid if a specified color (r = red or g = green) is drawn and* \$0 *else. The color is chosen after receiving additional information. A priori the agent only knows that there are*

[12]Note that the underlying commodity space for contingent claims $X : \Omega \to \mathbb{R}$ is L^2. Since utility is complete on this space by incorporating all possible extensions from $\mathcal{F}$ to $\mathcal{H}$, we may assume that endowment is measurable with respect to the finest information σ-algebra $\mathcal{H}$. By (1), $\frac{dP}{d\mathbb{P}}$ is square integrable. We have $E^P\mathtt{U} < \infty$ for all $P \in \mathscr{P}_F$ by the Minkowski inequality.

three red and green balls. According to Example 4, the probability p^r to draw a red ball lies in $[\frac{1}{4}, \frac{3}{4}] \simeq \mathcal{P}_F(0)$, where $F = \{1, \ldots, 6\}$. Now suppose the agent receives information that the 7th and 8th ball are red, then $\mathrm{p}^r \in [\frac{5}{12}, \frac{3}{4}] \simeq \mathcal{P}_F(1_{\{7,8\}})$. The revealing of a 9th and 10th ball, when both are green, yields $\mathrm{p}^r \in [\frac{5}{12}, \frac{7}{12}] \simeq \mathcal{P}_F(1_{\{7,\ldots,10\}})$. With a utility index $u(x) = \sqrt{x}$ and $\mathtt{U}^r = \sqrt{X^r}$, we derive

$$\begin{aligned}\max\big(U^g(1_{\{7,8\}}), U^r(1_{\{7,8\}})\big) &= U^r(1_{\{7,8\}}) \\ &= \min_{P\in\mathcal{P}_F(1_{\{7,8\}})} E^P \mathtt{U}^r - \min_{P\in\mathcal{P}_F(0)} E^P \mathtt{U}^r \\ &= \frac{5}{12}u(16) - 1 = \frac{2}{3}.\end{aligned}$$

The same calculation yields $U^g(1_{\{7,\ldots,10\}}) = U^r(1_{\{7,\ldots,10\}}) = \frac{2}{3}$. Full information $1_{F^c} = 1 - 1_F$ let the ambiguity vanish, thus $U^r(1_{\{7,\ldots,12\}}) = U^g(1_{\{7,\ldots,12\}}) = E^{\mathbb{P}}\mathtt{U}^g = 1$. The indifference price $C^{\mathtt{U}}(f) = u^{-1}\left(\max\left(U^g(f), U^r(f)\right)\right)$ of the gamble $X = \mathtt{U}^2$ for the respective information commodities, reveals the reservation price for additional information: $\frac{4}{9} = C^{\mathtt{U}}(1_{\{7,8\}}) \;=\; C^{\mathtt{U}}(1_{\{7,\ldots,10\}}) \;<\; C^{\mathtt{U}}(1_{F^c}) = 1$.

In view of Theorem 1.3, we discuss situations where information yields a strict utility improvement and situations where this is not the case.

Example 6 *In view of Example 5, the utility of additional information can be $U(1_{\{7,8\}}) = 0$, if the gamble X^g is fixed at the beginning. In that case, the alternative information $1_{\{9,10\}}$ yields a strict utility improvement. However, a priori the agent is unaware about the colors of the two additional balls and hence indifferent between $1_{\{7,8\}}$ and $1_{\{9,10\}}$. In the interim step, she can only choose to accept the gamble or not. Before making this choice, she will hedge the ambiguity about colors and strictly prefer $f = \frac{1}{2}1_{\{7,8\}} + \frac{1}{2}1_{\{9,10\}}$ with $U(f) \geq \frac{1}{2}(\frac{2}{3} + 0) > 0 = \min(U(1_{\{7,8\}}), U(1_{\{9,10\}}))$, as she cannot choose the color to bet on.*

The illustrated effect of Example 6, can be considered as a discontinuity of pure information commodities. We mention several alternative functional forms of utility under ambiguity.

Remark 2 *In Definition 2 and Theorem 1, we analyze the case of extreme ambiguity aversion found in Gilboa and Schmeidler (1989). This representation for ambiguity aversion serves as a device to receive a utility representation for information that quantifies the reduction of consistent priors. Alternative construction through other functional forms that quantify ambiguity attitudes include smooth ambiguity of Klibanoff, Marinacci, and Mukerji (2005) and variational preferences of Maccheroni, Marinacci, and Rustichini (2006).*

5 Prices and Demand for Information

To quantify optimal behavior under budget constraints with a given preference relation for information in the sense of Section 4, we specify what we mean by a (consistent) price system for information. Then we consider the resulting budget set and the single-agent optimization problem. However, the concavity of U provides early insights into the presence of a supporting hyperplane and of following, in some situations, the classical approach to those notions needed to define a meaningful equilibrium concept.

5.1 Commodity–Price Duality

We now return to the commodity space $L_H^\infty = L^\infty(H, \mathcal{B}(H), \mu)$ introduced in Subsection 3.2. The triple $(L_H^\infty, \|\cdot\|, \leq)$ is a classic Banach lattice, such that $(L_H^\infty)_+$ has a nonempty norm interior. It is straightforward to define a linear and positive price system $p^\nu(\cdot) = \int \cdot \mathrm{d}\nu$. The (topological) dual space $ba(H, \mathcal{B}(H), \mu)$, the space of bounded finitely additive set functions on $(H, \mathcal{B}(H))$ being absolutely continuous with respect to μ seems to be rather large. As such we restrict our attention to price functionals within the subspace $L_H^1 = L^1(H, \mathcal{B}(H), \mu)$. In this situation, a positive price system for information $p^\nu : L_H^\infty \to \mathbb{R}$ can be represented by an *information–price density* $\psi : H \to \mathbb{R}$ such that $\psi = \frac{d\nu}{d\mu} \in L_H^1$ and $\psi \geq 0$ yields

$$p^\nu(f) = \langle \psi, f \rangle = \int_H \psi(\mathtt{h}) f(\mathtt{h}) \mathrm{d}\mu(\mathtt{h}), \tag{5}$$

where $\langle \cdot, \cdot \rangle$ is the natural bilinear form of the pairing $\langle L_H^1, L_H^\infty \rangle$. Let $f = 1_{F'}$ be simple. An immediate consequence of (5) is $p^\nu(1_{F'}) = \int_{F'} \psi \mathrm{d}\mu = \nu(F')$. In general, the price system $p^\nu : [0, 1_H] \to \mathbb{R}$ satisfies three natural basic properties:

1. *additivity of disjoint information*: Let $F, G \in \mathcal{B}(H)$ be disjoint then $\nu(F \cup G) = \nu(F) + \nu(G)$. This condition refers to the linearity of disjoint fractions of information, which holds for countable disjoint sets as well.
2. *homogeneity of degree one for precision*: Let $F \in \mathcal{B}(H)$, $a \in [0, 1]$ then
$$p^\nu(a1_F) = \int_H a1_F \mathrm{d}\nu = a \int_F \mathrm{d}\nu = a\nu(F) = ap^\nu(1_F).$$
For the PI a, homogeneity quantifies the reliability of information F.
3. *monotonicity*: Let $F \subset G$ in $\mathcal{B}(H)$ then $p^\nu(1_F) \leq p^\nu(1_G)$. This simply means that noisier information is cheaper. The same form of monotonicity holds for $f \leq g$ in $[0, 1_H]$.

Property 1. can be extended if $g = \sum a_i 1_{A_i} \in [0, 1_H]$ and (A_i) are pairwise disjoint then $p^\nu(g) = \sum a_i \nu(A_i)$. If the collection of sets (A_i) fails to be disjoint then we have $p^\nu(g) \geq \sum a_i \nu(A_i)$, since $A_i \cap A_k \neq \emptyset$, for some $i \neq k$, implies that some information in g appears with a higher PI. The sub-additivity then follows from the unique representation in terms of mutually disjoint indicatior functions.

Summing up, the commodity price duality is given by $\langle L_H^\infty, L_H^1 \rangle$.

Remark 3 *The present notion allows us, in principle, to model a negative value of information, as examined in Hirshleifer (1971). In that case, agent's utility and the price system are no longer monotone increasing. By the Yoshida–Hewitt decomposition, the representing measure is the sum of a negative and positive measure.*

5.2 Walrasian Budget Set for Information

Motivated by the last subsection, we may consider a budget set based on a given price system $p^\nu : L_H^\infty \to \mathbb{R}_+$, amount of wealth $w \geq 0$ and information $\mathcal{F} = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F)$ for some $F \in \mathcal{B}(H)$. The budget set is given by

$$\mathbb{B}(p^\nu, w, F) = \Big\{ g \in [0, 1_{F^c}] : p^\nu(g) \leq w \Big\} \subset [0, 1_H], \tag{6}$$

where the positive and σ–additive measure $\nu : \mathcal{H} \to \mathbb{R}_+$ represents the given linear and positive price system p^ν via $p^\nu(g) = \int g\psi \mathrm{d}\mu$. Formally, the resulting budget correspondence is given by $\mathbb{B} : L_H^1 \times \mathbb{R}_+ \times [0, 1_H] \rightrightarrows [0, 1_H]$. Some properties, such as convex–, compact–valuedness, follow directly from the definition in (6).

Lemma 2

1. *The budget sets $\mathbb{B}(p^\nu, w, F)$ in (6) with $\frac{d\nu}{d\mu} = \psi \geq 0$ are non–empty, convex and weakly $\sigma(L_H^1, L_H^\infty)$–compact.*

2. *For any fixed $F \in \mathcal{B}(H)$, the budget correspondence $\mathbb{B}(\cdot, \cdot, F)$ is homogeneous of degree zero in price–wealth pairs.*

3. *Let X be a finite dimensional subvector space of L_H^∞, p be the restriction of p^ν to X and $w > \min_{x \in [0,1]} px$ (cheaper point). Then $\mathbb{B}^X(p, w, F) = \mathbb{B}(p, w, F) \cap X$ is a continuous correspondence at (p, w) in X.*

The budget set in (6) allows for several modifications. As preparation for the information–trade economy, we mention here one type of extension that also incorporates *information sells*.

We could allow the agent to sell parts of her information through the price system p^ν, as long as the agent owns this information, that is $f^{sell} \in [0, 1_F]$. The positive number $p^\nu(f^{sell})$ then denotes the realized proceeds. The resulting budget correspondence with information sales $\mathbb{B}^\circ : L^1_H \times [0, 1_H] \times \mathbb{R}_+ \times [0, 1_F] \Rightarrow [0, 1_H]$ is then given by

$$\mathbb{B}^\circ(p^\nu, F, w|f^{sell}) = \Big\{ g \in [0, 1_{F^c}] : p^\nu(g) \leq w + p^\nu \left(f^{sell}\right) \Big\}. \tag{7}$$

In (7), we only consider the case of single information sales. If there is more than one agent demanding some information in $[0, 1_F]$ then information can be sold simultaneously to several agents, and extends $\mathbb{B}^\circ$ in the appropriate way. In Section 6, the budget set $\mathbb{B}^*$ in (10) and the information–trade protocol takes this into account.

The following remark departs from an alternative perspective and discusses the agent's ability to sell information that she does not possess.

Remark 4 *Consider an economy with two agents, J and K. If K buys an information commodity from J, then K may believe that J is not reporting truthfully or he does not own the information that is reported. For the latter case, this does not imply that the sold information is not true.*[13] *Of course we could assume that the agents care about their reputation or follow a gentleman's agreement, and for this reason such complications may not occur. However, the present setting also allows us to incorporate dishonest sellers. We mention two ways this information creates friction that can enter the budget set.*

The first way requires collateral to sell information at a given price system p^ν. If J sells to K information $G \in \mathcal{B}(H)$, then he must hold some wealth w^J_G. A restriction of collateral is then $p^\nu(1_G) \leq w^J_G$. This collateral is transferred to agent K, if it turns out that J delivered false information. As in Geanakoplos (2010), this reliability problem lets contracts become pairs of promises and collateral $(1_G, w^J_G)$.

An alternative approach refers to buying with awareness of imprecision. *Suppose seller J offers the pure bundle 1_G. In the case that J enjoys a questionable reputation about telling the truth, then agent K may interpret the offered bundle as $a^K_J 1_G$. Here, $a^K_J \in [0, 1]$ captures the precision or the likelihood of K that J reports true information. For agent K, the personalized value of 1_G, to be delivered by J, reduces to $a^K_J p^\nu(1_G)$.*

[13] However, this thought experiment also applies to Arrow-Debreu modeling where contingent claims with future maturity are traded and promised today.

5.3 Demand for Information

Fix a price system $p^\nu : [0, 1_H] \to \mathbb{R}$ in L^1_H, initial wealth $w \in \mathbb{R}_+$ and initial information $F \in \mathcal{B}(H)$. The next step points to a solvable formulation of the agent's optimization problem. Combining the utility specification of Section 4 and the budget set of Section 5.2, the individual utility maximization problem reads as follows,

$$\max_{h \in \mathbb{B}(p^\nu, w, F)} U(h). \tag{8}$$

Proposition 2 justifies many properties from classical demand theory.

Proposition 2 *Let $F \in \mathcal{B}(H)$ be the given amount of initial information. The price system ν and wealth w are given as well. The problem in (8) has a solution and the* demand correspondence for information $\mathbb{D} : L^1_H \times \mathbb{R}_+ \times [0, 1_H] \rightrightarrows [0, 1_H]$ *defined by*

$$\mathbb{D}(p^\nu, w, F) = \arg \max_{h \in \mathbb{B}(p^\nu, w, F)} U(h) \tag{9}$$

is nonempty–, convex– and $\sigma(L^1_H, L^\infty_H)$–compact–valued. Moreover, $\mathbb{D}$ is homogeneous of degree zero in (p^ν, w), that is $\mathbb{D}(\alpha p^\nu, \alpha w, F) = \mathbb{D}(p^\nu, w, F)$ for all $\alpha > 0$.

So far, $\mathbb{D}$ is formulated only for pure information bundles of indicator type 1_F. For general $f \in [0, 1_H]$ the same conclusions follow analogously.

6 The Information–Trade Economy

In this section, we consider a finite set of agents $\mathbb{I} = \{1, \ldots, I\}$. Each i is characterized by a random utility $\mathtt{U}_i$ and private information $\mathcal{F}_i = \sigma(F_i)$. As in the standing assumption, the set F_i generates the σ-algebra $\mathcal{F}_i = \sigma(X_{\mathtt{h}} : \mathtt{h} \in F_i)$ on Ω. The *information–trade economy* is summarized by

$$\mathcal{E}^{\mathtt{Info}} = \{[0, 1_H], U_i, F_i\}_{i \in \mathbb{I}},$$

where each U_i is induced by $\mathtt{U}_i \in L^2(\Omega, \mathcal{H}, \mathbb{P})$ and Definition 2. Since we allow the agents to sell their given information, we do not consider initial wealth.

6.1 Feasible Information Allocations

Information (as a commodity) inherits the special property that its release to some agent leaves the previous property rights unchanged. For this reason, the concept of a feasible information allocation differs from feasibility of physical commodities.

Let $g_i(k) \in [0, 1_H]$ denote the information that agent k receives from agent i. In this notation, an individually rational bilateral trade of information, say $g_1(2), g_2(1)$, between agent 1 and 2 has to satisfy an individual rationality condition $(g_1(2), g_2(1)) \in [0, 1_{F_1}] \times [0, 1_{F_2}] \cap [0, 1_{F_2^c}] \times [0, 1_{F_1^c}]$.

Definition 3 *Fix an initial information allocation $(F_1, \ldots, F_I) \in \mathcal{H}^I$ and set $[0, 1_{F_{i,j}^{\,c}}] = [0, 1_{F_i \cap F_j^c}]$. An (individually rational)* information exchange protocol (IEP) matrix *G is given by:*

$$G = \begin{pmatrix} \emptyset & g_1(2) & \cdots & g_1(I) \\ g_2(1) & \emptyset & \cdots & g_2(I) \\ \vdots & \vdots & \ddots & \vdots \\ g_I(1) & g_I(2) & \cdots & \emptyset \end{pmatrix} \in \mathbb{F} := \begin{pmatrix} \emptyset & [0, 1_{F_{2,1}^{\,c}}] & \cdots & [0, 1_{F_{I,1}^{\,c}}] \\ [0, 1_{F_{1,2}^{\,c}}] & \emptyset & \cdots & [0, 1_{F_{I,2}^{\,c}}] \\ \vdots & \vdots & \ddots & \vdots \\ [0, 1_{F_{1,I}^{\,c}}] & [0, 1_{F_{2,I}^{\,c}}] & \cdots & \emptyset \end{pmatrix},$$

where $\mathbb{F} = \mathbb{F}(F_1, \ldots, F_I)$ only depends on $(F_1, \ldots, F_I)$.

Here $\mathbb{F}_{i,j} = [0, 1_{F_{i,j}^{\,c}}]$ contains all information commodities that agent i owns and which j is interested in. We receive for each pair (i, j) the restriction $g_i(j) \in \mathbb{F}_{i,j}$ for the feasibility of the IEP-matrix. The initial information allocation can be generalized to arbitrary allocations in $[0, 1_H]^I$.

Information releases are collected for each agent in the respective column, $\overrightarrow{g}(i) := \sum_{j \in \mathbb{I}} g_j(i)$ as the information sales of agent i. Information acquisitions are listed in the corresponding row, set $\overleftarrow{g}_i := \sum_{j \in \mathbb{I}} g_i(j)$.

Lemma 3 *The set of feasible allocations $\mathbb{F}$ is weakly compact and convex.*

Once again, let us reconsider the Ellsberg urn from Section 2.

Example 7 *Let there be three agents seeing the urn in Section 2 from three different angles. As illustrated in Figure 2, agent 1 is only able to identify the balls in part $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$, agent 2 (or 3) sees parts B and C (or C and D). In the present case, the set of individually rational and feasible information exchanges can be summarized by the following IEP matrix*

$$\mathbb{F} = \begin{pmatrix} \emptyset & [0, 1_{F_{2,1}^{\,c}}] & [0, 1_{F_{3,1}^{\,c}}] \\ [0, 1_{F_{1,2}^{\,c}}] & \emptyset & [0, 1_{F_{3,2}^{\,c}}] \\ [0, 1_{F_{1,3}^{\,c}}] & [0, 1_{F_{2,3}^{\,c}}] & \emptyset \end{pmatrix} = \begin{pmatrix} \emptyset & [0, 1_A] & [0, 1_{A \cup B}] \\ [0, 1_C] & \emptyset & [0, 1_B] \\ [0, 1_{C \cup D}] & [0, 1_D] & \emptyset \end{pmatrix}.$$

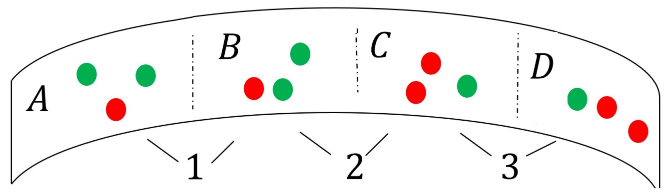

Figure 2: The unfolded lateral area of the Ellsberg urn from Section 2 – each agent captures only a share of the urn. An initial allocation of information with three agents is then given by $F_1 = A \cup B$, $F_2 = B \cup C$, $F_3 = C \cup D$.

6.2 Existence of an Information–Trade Equilibrium

Each agent can sell the same information more than once. To account for this aspect, we modify the budget set such that the received wealth from information sales can be incorporated. The general budget correspondence $\mathbb{B}^* : L^1_H \times \mathbb{R}_+ \times [0, 1_H] \times [0, I \cdot 1_H] \rightrightarrows [0, 1_H]$ with multiple information sales is given by

$$\mathbb{B}^*\left(p^\nu, F | \overrightarrow{g}(i)\right) = \left\{ \overleftarrow{g}_i \in [0, 1_{F^c}] : p^\nu\left(\overleftarrow{g}_i\right) \leq \sum_{k \neq i} p^\nu\left(g_k(i)\right) \right\}, \tag{10}$$

where $\overrightarrow{g}(i) = \sum_{k \neq i} g_k(i)$. Note that, *(i.)* the price system p^ν is defined on the entire commodity space L^∞_H, *(ii.)* the sum and p^ν in (10) commutes, *(iii.)* Lemma 2 holds by the same arguments, also for the present budget set $\mathbb{B}^*$ and *(iv.)* the strategy/budget set depends not only on prices, but also on the behavior of the other agents.

The notion of equilibrium, when information commodities are traded is introduced in the following definition.

Definition 4 *Fix an economy* $\mathcal{E}^{\texttt{Info}} = \{[0, 1_H], U_i, F_i\}_{i \in \mathbb{I}}$. *The feasible IEP–matrix* $G^* = (g^*_k(j)) \in \mathbb{F}$ *and a price system* $\nu \in ba(H, \mathcal{B}(H), \mu)$, *with* $\nu \geq 0$, *build an* Information–Trade (IT) Equilibrium, *if*

1. *for each* $i \in \mathbb{I}$, $\overleftarrow{g}^*_i$ *maximizes* U_i *in* $\mathbb{B}^*(p^\nu, F_i | \overrightarrow{g}^*(i))$,
2. $\sum_i \overleftarrow{g}^*_i \leq \sum_i \overrightarrow{g}^*(i)$ *in* L^∞_H.

In equilibrium, each agent considers the amount and multiplicity of total information sales $\overrightarrow{g}(i) \leq I - 1$ as given. However, the additional feasibility condition in the second part of Definition 4 is needed to check if in the aggregate, the received information is indeed delivered by anyone.

For the existence of equilibrium we need the following conditions for the primitives of the economy.

Assumption 1 *(Adequacy condition) There is a $\delta > 0$, arbitrarily small, such that for each $i \in \mathbb{I}$, the information of agent i is given by $(1-\delta)1_{F_i} + \delta \in \operatorname{int}\big((L_H^\infty)_+\big)$. For every i we have $\mu(F_i^c) > 0$ (no omniscient agent). Agent's utility functions $U^i : [0, 1_H] \to \mathbb{R}$ are monotone, continuous and concave.*

The first part is a variation of the cheaper point assumption. In the present setup it means that δ yields a slighter smaller set of priors. Instead of $\mathscr{P}_F$ the agent starts with $\mathcal{P}_{F\delta H} = \delta\mathbb{P} + (1-\delta)\mathscr{P}_F$. By Corollary 1, we have $\mathbb{P} \in \mathscr{P}_F$ and hence $\mathcal{P}_{F\delta H} \subset \mathscr{P}_F$.

Here is the main result of this section.

Theorem 2 *Under Assumption 1, an IT-Equilibrium exists.*

To prove the existence of equilibrium a careful look is needed when it comes to the question which group of agents participates in delivering information to some agent i. For instance, it must be clarified what happens when agent i demands information that is owned by more than one agent, say J agents own this information. In that case, we assume that the market mechanism divides the supply in J shares of equal sizes. This is an *equal–treatment property* in the trading mechanism. To accomplish this property, we use a protocol procedure that controls the feasibility of the IEP-matrix. In step 2 of the proof for Lemma 4, the procedure is spelled out in detail.

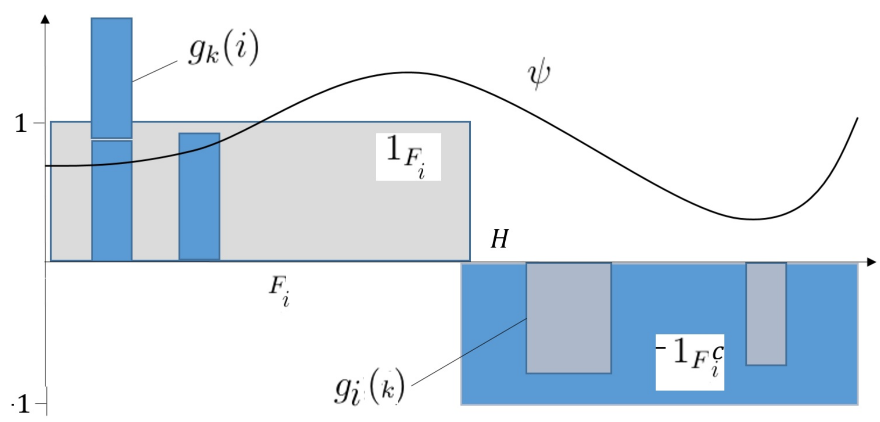

Figure 3: *Trade of information bundles.* Using the typical notation, we consider the situation for agent i. She sells a piece of information twice. One share is sold to agent k from which she buys $g_i(k) = a1_{F_i^c}$ at price $p^\nu(g_i(k)) = a\int_{F_i^c} \psi(\mathtt{h})d\mu(\mathtt{h})$.

Finally, we move back to the finite case from Section 2. The following basic and fundamental lemma is the starting point for the proof of Theorem 2.

For finite H, as in Example 1, the commodity space for information is finite dimensional and the consumption set becomes the order interval $[0,1]^{|H|}$.

Lemma 4 *Under Assumption 1, IT-equilbria exist for every finite dimensional subvector space of L^∞_H.*

To develop an intuition about Theorem 2 and Definition 4, the following example considers an IT–equilibrium in the economy of Example 7.

Example 8 *Fix the initial information endowment in Figure 2 of Example 7 (departing from the urn in Section 2) and the common random utility $\mathtt{U}_i = u_i(X)$ of Example 5 with $u_i(\cdot) = \sqrt{\cdot}$ for $i = 1,2,3$. X is a gamble, where* \$16 *is paid if red is drawn and* \$0 *if the ball is green. To simply the notation, let the urn be divided into subgroups, such that $H = \{A,B,C,D\}$. For this economy, an IT–equilibrium is given by the price system $p^\nu(f) = \sum_{\mathtt{h}\in H} \nu(\mathtt{h}) \cdot f(\mathtt{h})$, with $\nu = (2,1,1,2)$ and the IEP–matrix*

$$\begin{pmatrix} \emptyset & \frac{1}{4}1_A & 1_A + \frac{1}{2}1_B \\ \frac{1}{2}1_C & \emptyset & \frac{1}{2}1_B \\ 1_D + \frac{1}{2}1_C & \frac{1}{4}1_D & \emptyset \end{pmatrix} \in \begin{pmatrix} \emptyset & [0,1_A] & [0,1_{A\cup B}] \\ [0,1_C] & \emptyset & [0,1_B] \\ [0,1_{C\cup D}] & [0,1_D] & \emptyset \end{pmatrix}.$$

The optimal information allocation of the IT-equilibrium $(\overleftarrow{g}_1, \overleftarrow{g}_2, \overleftarrow{g}_3) = (1_{C\cup D}, \frac{1}{4}1_{A\cup D}, 1_{A\cup B})$ is fully revealing for agent 1 and 3, but not for agent 2, since she does not own a sufficient amount of exclusive information. Agent 2 diversifies her information consumption $(\frac{1}{4}1_A + \frac{1}{4}1_D)$, due to the concavity of her information utility functional.

An asymmetric distribution of initial information implies no trade IT-equilbria. To see this, assume an economy with initial wealth. Let $F_1 \subset \cdots \subset F_i \subset F_{i+1} \cdots \subset F_I = F$ and $w_1 \le \ldots \le w_i \le w_{i+1} < \ldots \le w_I$. In this situation agent $i+1$ would not buy information from i. Suppose now that $w_1 > \ldots > w_i > w_{i+1} > \ldots > w_I$ and utility comes from wealth and information, via $V_i(w,f) = u_i(w) + U_i(f)$. This specification of the primitives will lead to trade in equilibrium.

A central question points to statements of the form: *Is the equilibrium allocation of information fully revealing?* In the framework of this paper, this is an endogenous property (see Example 8 where the equilibrium is not fully revealing). In the following we give a condition that depends on the equilibrium price measure.

Corollary 2 *Let ν be an equilibrium price for information. If $\nu(F_i) < \nu(F_i^c)$, then agent i is unable to receive all the information in the economy.*

7 Conclusion

Information economics often models the agent's ability to observe the realization of some signal, such that information then relies on the observation itself. In this paper, the notion of information relies on the accuracy of the knowledge about the exact distribution of the signal. Minimal information refers to knowledge of the expected value of the probability distribution and corresponds to the trivial σ-algebra being generated by the trivial partition.

Knowledge about some payoff–relevant distribution remains in central focus for the quantification of utility. An agent is willing to pay for the reduction of imprecision that excludes possible priors. Without this additional consumption of information, these priors are considered as possible candidates for the agent's expected utility. Under the assumption of ambiguity aversion, the reduction of possible priors increases the utility. The possibility of excluding different priors has different effects on the worst case expected utility.

The approach taken by this paper relies on marginal aspects. A marginal change of utility comes from a marginal change in the size of ambiguity, which in turn is directly induced by a change of information. This reasoning allows us to assign a positive endogenous equilibrium price to any information commodity.

A Proofs for Section 4

Proof of Lemma 1 *Starting with $\mathcal{P}_F(a1_A + b1_B) = \mathcal{P}_{Fa1_A}(b1_B)$, the claim directly follows by the rules of the IA–correspondence:*

$$\begin{aligned}
\mathcal{P}_F(a1_A + b1_B) &= \mathcal{P}_{Fa1_A}(b1_B) \\
&= b\mathcal{P}_{Fa1_A}(1_B) + b^-\mathcal{P}_{Fa1_A} \\
&= b(a^-\mathcal{P}_{FB} + a\mathcal{P}_{FB}(1_A)) + b^-\mathcal{P}_F(a1_A) \\
&= b(a^-\mathcal{P}_{FB} + a\mathcal{P}_{FBA}) + b^-(a^-\mathcal{P}_F + a\mathcal{P}_F(1_A)) \\
&= ba^-\mathcal{P}_{FB} + b^-a\mathcal{P}_{FA} + ba\mathcal{P}_{FBA} + b^-a^-\mathcal{P}_F
\end{aligned}$$

Note that $\mathcal{P}_F(a1_A + b1_B) = \mathcal{P}_{Fb1_B}(a1_A)$ delivers the same result. ■

In the following, we give an extension of Lemma 1 for those information commodities that can be represented as finite sums. Let $f = \sum_{k\in N} a_k 1_{A_k} \leq 1_H$, with $N \in \mathbb{N}$ and $A_k \in \mathcal{B}(H)$, pairwise disjoint, then the IA–correspondence is given by

$$\mathcal{P}_F(f) = \sum_{M\in 2^N} \rho_M \mathcal{P}_{F\bigcup_{k\in M} A_k} \tag{11}$$

where the probability weights $(\{\rho_M\}_{M\in 2^N}) \in \Delta(2^N)$ are given by

$$\rho_M = \rho_M(\{a_k\}_{k\in M}) = \prod_{k\in M} a_n \prod_{k\in N\setminus M} (1-a_n).$$

As the consumption set $[0, 1_H]$ allows for more general $\mathcal{B}(H)$-measurable functions $f : H \to [0,1]$ as representation of information, the next step clarifies how Definition 1 extends to arbitrary element in $[0, 1_H]$.

A successive repetition of the rules of Definition 1 gives us for a general $f = \sum_{n\in\mathbb{N}} a_n 1_{A_n}$ in $[0, 1_H]$ the following set of consistent priors

$$\mathcal{P}_F(f) = \sum_{N\in 2^{\mathbb{N}}} \rho_N(\{a_m\}_{m\in N})\mathcal{P}_F\left(1_{\bigcup_{m\in N}} A_m\right) = \int_{[0,1)} \mathcal{P}_F(x)\mathrm{d}\rho(x), \tag{12}$$

where $\mathcal{P}_F(x) = \mathcal{P}_{F\bigcup_{k\in N} A_k}$, with $x = \mathtt{b}(N)$. As in (11), $\rho_N = \rho(x)$ is induced by the inductive application of Lemma 1. We apply here the existence of a bijection $\mathtt{b} : 2^{\mathbb{N}} \to [0,1)$ between the power set of natural number and the interval $[0,1)$. ρ is a probability measure on $[0,1)$.

Proof of Proposition 1 *In abuse of notation, set $\mathcal{P}_F(1_A) = \mathcal{P}(A) = \mathcal{P}_{FA}$. By construction, the correspondence maps into the convex subset of $\mathscr{P}_F$.*

1. *$\mathcal{P}_F(h)$ is a convex set, as (Minkowski) sums of convex sets are again convex. Similar arguments holds true for compactness: The Minkowski sum of compact sets is closed. Since the sum is in a compact set $\mathscr{P}_F$ the sum is therefore compact.*

2. *The correspondence $\mathcal{P}_F$ has a compact Hausdorff range. Hence by the Closed Graph Theorem it suffices to show that $\mathcal{P}_F$ has a closed graph with respect to the norm to weak topology $\|\cdot\|_{L^\infty_H} \times \sigma((L^2)^*, L^2)$. As mentioned in footnote 7, we assume that priors in $\mathscr{P}_F$ are absolutely continuous with respect to $\mathbb{P}$ and have a square integrable Radon–Nykodym derivative, for this reason, we may take the weak star topology of $L^2 = L^2(\Omega, \mathcal{H}, \mathbb{P})$.*

   Fix a general information bundle $f \in [0, 1_H]$ and note that $\mathscr{P}_F$ is weakly compact. In order to show upper hemi-continuity, it suffices to prove that $f \rightrightarrows \mathcal{P}_F(\sum_{1\le n\le\infty} a_n 1_{A_n})$ has a closed graph in the above mentioned topology. For the rest of the proof we set any PI $a_n \in [0,1]$ to 1, since they solely appear as factors. In general, the proof follows by the same convergence argument.

Set $A^k = \cup_{1\le n\le k} A_n$. Consider a sequence $f^k = \sum_{1\le n\le k} 1_{A_k}$ converging in the $\|\cdot\|_\infty$-norm to f. Let P^k converge to P^∞ in the weak star $\sigma((L^2)^*, L^2)$ topology, where $P^k \in \mathcal{P}_F(f^k)$. We have to prove that $P^\infty \in \mathcal{P}_F(f)$.

Now suppose $P^\infty \notin \mathcal{P}_F(f)$. This means there is an event $E \in \sigma(A^\infty) = \sigma(X_{\mathtt{h}} : \mathtt{h} \in A^\infty)$ such that $\big|\mathbb{P}(E) - P^\infty(E)\big| > \varepsilon$. From this we infer, that for large m there is an event $E^m \in \sigma(A^m)$ such that

$$|\mathbb{P}(E^m) - P^\infty(E^m)| > \frac{\varepsilon}{2}.$$

As P^n converges to $P^\infty \in \mathscr{P}_F$, there is a $n \ge m$ such that

$$|P^n(E^m) - P^\infty(E^m)| < \frac{\varepsilon}{4}.$$

Note that the probability measure $P^n : \sigma(A^n) \to [0,1]$ can evaluate E^m since $n \ge m$. Combining the last two inequalities via the sublinearity of $|\cdot|$ gives us

$$\begin{aligned} & \big|\mathbb{P}(E^m) - P^n(E^m)\big| \\ \ge\ & \big|\mathbb{P}(E^m) - P^\infty(E^m)\big| - \big|P^n(E^m) - P^\infty(E^m)\big| \\ \ge\ & \frac{\varepsilon}{2} - \frac{\varepsilon}{4} > 0. \end{aligned} \tag{13}$$

But since $P^n \in \mathcal{P}_F(A^n)$ we have $P^n = \mathbb{P}_{\sigma(A^n)}$, where $\mathbb{P}_{\sigma(A^n)}$ denotes the restriction of $\mathbb{P}$ to $\sigma(A^n)$. This holds by the definition of the IA–correspondence and yields a contradiction with respect to (13).

3. *This follows directly from the construction. Consider $a1_A$ and $b1_A$ with $a \le b$. By definition, the precision $a \in [0,1]$ gives less weight to the smaller set in the Minkowski sum. Hence $\mathcal{P}_F(a1_A) \supset \mathcal{P}_F(b1_A)$ follows. The case $a1_A$ and $a1_B$, with $A \subset B$ results by a similar argument. The conclusion for arbitrary sums follows then by the stability of the inclusion under arbitrary intersection, i.e. if $A_n \subset B_n$ for all n then $\cap_n A_n \subset \cap_n B_n$.*

4. *First we consider the case $f = a1_A$ and $g = b1_B$. By Lemma 1, we derive*

$$\begin{aligned} & \mathcal{P}_F(\alpha a A + (1-\alpha) b B) \\ =\ & \alpha^- b (\alpha a)^- \mathcal{P}_{FB} + (\alpha^- b)^- \alpha a \mathcal{P}_{FA} + \alpha a \alpha^- b \mathcal{P}_{FBA} + (\alpha^- b)^- (\alpha a)^- \mathcal{P}_F \\ \subseteq\ & \alpha^- b \qquad \mathcal{P}_{FB} + \qquad \alpha a \mathcal{P}_{FA} + \alpha a^- \quad \mathcal{P}_F \quad + \alpha^- b^- \qquad \mathcal{P}_F \\ =\ & \alpha^- (b \mathcal{P}_{FB} + b^- \mathcal{P}_F) + \alpha (a \mathcal{P}_{FA} + a^- \mathcal{P}_F) \\ =\ & \alpha^- \mathcal{P}_F(b1_B) + \alpha \mathcal{P}_F(a1_A). \end{aligned}$$

The inclusion follows from the following list of facts: $\mathcal{P}_F \supset \mathcal{P}_{FB}, \mathcal{P}_{FA}$*,* $\mathcal{P}_{FB}, \mathcal{P}_{FA} \supset \mathcal{P}_{FBA}$*,* $\alpha^- b \geq \alpha^- b(\alpha a)^-$ *and* $\alpha a \geq (\alpha^- b)^- \alpha a$*.*

The case of $f = \sum_{k \in N} a_k 1_{A_k} \leq 1_H$*, with* N *finite, follows by the same argument, where we apply (11), the extension of Lemma 1 and apply the above arguments repeatedly.*

The general case follows by a similar argument. In view of (12), we utilize the continuity of the integral: Let $f = \sum_{n \in \mathbb{N}} a_n 1_{A_n}$*,* $g = \sum_{n \in \mathbb{N}} b_n 1_{B_n}$*, we get* $\mathcal{P}(\alpha f + (1-\alpha)g) \subset \alpha \mathcal{P}_F(f) + (1-\alpha)\mathcal{P}_F(g)$*. The inclusion follows from the same argument as in the case for* $f^N = \sum_{1 \leq k \leq N} a_k 1_{A_k}$ *and* $g = \sum_{1 \leq k \leq N} b_k 1_{B_k}$*, by applying* $\lim f^N + g^N = f + g$*.* ■

Proof of Corollary 1 *This follows directly from* $\mathcal{P}_F(1_H) = \{\mathbb{P}\}$*, as stated in (2), and an application of Proposition 1.2.* ■

Proof of Theorem 1 *1. Monotonicity follows directly from the monotonicity of the IA–correspondence. If* $f \leq g$ *then* $\mathcal{P}(f) \supseteq \mathcal{P}(g)$ *and* $U(f) \leq U(g)$ *then directly by the maxmin functional form.*

2. Continuity follows from the norm to weak upper hemi-continuity of the IA–correspondence and an application of a version of Berge's maximum theorem of Tian and Zhou (1992). Specifically, we apply their Theorem 1. For the sake of completeness let us check the conditions:

(a) Since the expectation $E^P \mathtt{U}$ *is linear in* P*, upper semi-continuity holds.*

(b) The IA–correspondence has a weak (and hence Hausdorff) compact range space and is upper hemi-continuous, by Proposition 1. By the closed graph theorem, the correspondence is closed.

(c) "feasible path transfer lower semi-continuity" holds since the primitive function $P \mapsto E^P$ *is linear in* P *and does not depend on the* f *argument in the IA-correspondence.*

With the additional upper hemi-continuity of $\mathcal{P}_F(\cdot)$ *from Proposition 1, we get the continuity of* $f \mapsto U(f)$*.*

3. *concavity: Let $\lambda \in [0,1]$ and set $\lambda^- = 1-\lambda$. We derive*

$$\begin{aligned}
\lambda U(f) + \lambda^- U(g) &= \lambda \min_{P\in\mathcal{P}(f)} E^P \mathtt{U} + \lambda^- \min_{P\in\mathcal{P}(g)} E^P \mathtt{U} \\
&= \min_{P\in\lambda\mathcal{P}(f)} E^P \mathtt{U} + \min_{P\in\lambda^-\mathcal{P}(g)} E^P \mathtt{U} \\
&= \min_{P\in\lambda\mathcal{P}(f)+\lambda^-\mathcal{P}(g)} E^P \mathtt{U} \\
&\le \min_{P\in\mathcal{P}(\lambda f+\lambda^- g)} E^P \mathtt{U} \\
&= U(\lambda f + \lambda^- g),
\end{aligned}$$

where the inequality follows from the convexity of the IA–correspondence, stated in Proposition 1.4. The third equality can be achieved by a direct computation:

$$\begin{aligned}
\lambda \min_{P\in\mathcal{P}(f)} E^P \mathtt{U} + \lambda^- \min_{P\in\mathcal{P}(g)} E^P \mathtt{U} &= \lambda E^{P_f} \mathtt{U} + \lambda E^{P_g} \mathtt{U} \\
&= E^{\lambda P_f + \lambda^- P_g} \mathtt{U} \\
&\ge \min_{P\in\lambda\mathcal{P}(f)+\lambda^-\mathcal{P}(g)} E^P \mathtt{U} \\
&= E^{\lambda P_f^* + \lambda^- P_g^*} \mathtt{U} \\
&= \lambda E^{P_f^*} \mathtt{U} + \lambda^- E^{P_g^*} \mathtt{U} \\
&\ge \lambda \min_{P\in\mathcal{P}(f)} E^P \mathtt{U} + \lambda^- \min_{P\in\mathcal{P}(g)} E^P \mathtt{U}
\end{aligned}$$

The first inequality employs the fact $\lambda P_f + \lambda^- P_g \in \lambda\mathcal{P}(f) + \lambda^-\mathcal{P}(g)$. The other derivations hold, since each $\mathcal{P}(\cdot)$ is a weakly compact–valued correspondence, by Proposition 1.1, and hence in each case the minimum is attained, by the linearity of $P \mapsto E^P \mathtt{U}$. Consequently, the inequalities of the last derivation are indeed equalities. ■

B Proofs for Section 5 and 6

Proof of Lemma 2 *1. By (5), the budget set is an intersection of a closed half spaces that contains a zero within the consumption set $[0, 1_H]$. Hence, the budget set is nonempty, closed, bounded and convex. Compactness follows by Alaoglu's theorem.*

2. *From (6), we directly see that a doubling of p^ν and w leaves the budget set unchanged.*

3. *The definition of $\mathbb{B}$ agrees with that of a standard truncated budget correspondence. Hence, upper hemi–continuity follows immediately. Due to the cheaper point condition, lower hemi–continuity follows by the standard proof.* ■

Proof of Proposition 2 *By Lemma 2, the budget set is nonempty and $\sigma(L^1_H, L^\infty_H)$–compact. By Theorem 1, $U : [0, 1_H] \to \mathbb{R}$ is continuous and concave and hence $\sigma(L^1_H, L^\infty_H)$ upper semi–continuous. A maximizer exists.*

By standard results from convex analysis, $\mathbb{D}(p^\nu, w, F)$ is convex and $\sigma(L^1_H, L^\infty_H)$-compact.

Homogeneity of degree zero follows directly from the homogeneity of degree zero of $\mathbb{B}$, stated in Lemma 2.2. ■

Proof of Lemma 3 *Clearly, $\mathbb{F}$ is a subset of $(L^\infty_H)^{I\times I}$ and consequently the product of $\sigma(L^1, L^\infty_H)$–compact and convex order intervals.* ■

Proof of Theorem 2 *We follow the proof strategy of Theorem 1 in Bewley (1972). In Lemma 4 the existence of equilibria for finite dimensional information commodity spaces, that contain $\{1_{F_i}\}_{i\in\mathbb{I}}$, is established. This serves as a base to find a converging a (sub)–net $(\overleftarrow{g}^N_1, \ldots, \overleftarrow{g}^N_I, p^N)$ of equilibra in the finite dimensional economy. The application of the Krein–Rutman version of Hahn–Banach and Alaoglu theorem are exactly the same and will not be repeated. We have a candidate equilibrium $(\overleftarrow{g}^*_1, \ldots, \overleftarrow{g}^*_I, p^*)$. Note that by Lemma 3 $\mathbb{F}$ is compact and closed, hence $(\overleftarrow{g}^*_1, \ldots, \overleftarrow{g}^*_I)$ is feasible.*

Clearly, the price system $p^* \in ba(H, \mathcal{H}, \mu)$ is positive. From the subnet convergence, we have $1 = \|p^N\|_{ba} = p^N(1_H) \to p^*(1_H)$, whence $p^* \neq 0$.

To show that the candidate is indeed an IT-equilibrium in $\mathcal{E}^{\texttt{Info}}$, the proof of Bewley (1972) for the optimality condition of a quasi-equilibrium

$$U^i(g) \geq U^i(\overleftarrow{g}^*_I) \Rightarrow p^*(g) \geq p^*(\overleftarrow{g}^*_I)$$

follows again exactly by the same lines.

The adequacy condition in Assumption 1 and $p^* > 0$ implies $p^*(\overleftarrow{g}^*_i) > \inf_{g\in[0,1_H]} p^*(g)$ which in turn implies maximality of $\overleftarrow{g}^*_i$ in the i-th agent's budget set.

It remains to show that $p^* \in L^1$. By the Yosida–Hewitt theorem, we decompose $p^* = p_c + p_f$ in the countable additive part p_c and finitely additive part p_f. By contradiction we show $p_f = 0$.

Suppose there is another agent $l \notin I$. With $F_l = \emptyset$. In view of Assumption 1, she is endowed with $\delta 1_H$. Clearly, as this information commodity is owned by all other agents, no agent demands this commodity. Hence agent l, cannot affect the equilibrium outcome. Her priors are $\mathcal{P}_\delta = (1-\delta)\mathscr{P}_\emptyset + \delta\mathbb{P}$. Assume $\mathtt{U}_l$ is chosen such that

$$\{P_l\} = \arg\max_{P \in \mathcal{P}_\delta} E^P \mathtt{U}_l \qquad \text{and } \mathbb{P} \neq P_l.$$

We may find an (extremely desirable) information bundle 1_A with $A \in \mathcal{B}(H)$ such that $P_l \notin \mathcal{P}_\delta(1_A)$ and moreover for every $\varepsilon \in (0,1)$ we still have $P_l \notin \mathcal{P}_\delta(\varepsilon 1_A) = \varepsilon\mathcal{P}_\delta(1_A) + (1-\varepsilon)\mathcal{P}_\delta$, i.e. a small fraction of A already allows to exclude P_l from the new set of consistent priors.

Suppose $p_f(1_H) > 0$, then $p_c(\delta 1_H) < p^*(\delta 1_H)$, by the positive homogeneity of p^*. Now choose an ε sufficiently small such that $p_c(\delta 1_H) + p(\varepsilon 1_A) < \delta p(1_H)$ and $\varepsilon + \delta < 1$.

On the other hand, there is a decreasing sequence of sets $A_n \in \mathcal{B}(H)$ such that $\mu(A_n) \to 0$ and $p_f(H \setminus A_n) = 0$ for each n. Define $y_n = \delta 1_{H \setminus A_n} \in [0, 1_H]$ and since $y_n \to \delta 1_H$ in measure and hence also in the Mackey topology, lower semi-continuity of U^l and $U^l(\delta 1_H + \varepsilon 1_H) > U^l(\delta 1_H)$ implies $U^l(y_n + \varepsilon 1_H) > U^l(\delta 1_H)$ for n large.

However we have $p_f(y_n) = 0$ for every n and consequently $p^*(y_n + \varepsilon 1_A) < p^*(\delta 1_H)$, a contradiction, and therefore $p_f = 0$ and $p^* \in L^1(H, \mathcal{B}(\Omega), \mu)$. ■

The last part of the proof follows the proof of Theorem 8.2 of Mas-Colell and Zame (1991). The assumption that each convex consumption set coincides with the positive cone, can be weakened as only the presence of the 0 in the consumption set matters for the proof.

Proof of Lemma 4 *In abuse of notation, we formulate the information commodities as 1_F instead of $\delta + (1-\delta)1_F$, as required in Assumption 1.*

By Lemma 2, the budget correspondence is nonempty, convex and compact valued. Moreover it is continuous. A truncation of the economy is not needed.

We divide the proof in several steps.

1. Individual demand: *By Proposition 2, for every $i \in \mathbb{I}$, $\mathbb{D}_i(p, F_i | \overrightarrow{g}(i))$ is a nonempty-, convex- and compact-valued correspondence. Note that in this finite dimensional setup the weak and norm topology coincide. Lemma 2.3. and an application of Berge's maximum theorem yields the upper hemi-continuity of the demand correspondence.*

2. Recovery of the IEP matrix : *Let us fix some demand* $\overleftarrow{g}_i \in \mathbb{D}_i \subset [0, 1_{F_i^c}]$ *for some agent* i. *We identify those agents who are part of* $\overleftarrow{g}_i$. *The following 2–step procedure clarifies this.*

   (a) Decomposition via F_i^c: *There is a unique partition of* F_i^c, *indexed by the set of coalitions,*

   $$F_i^c = \bigcup_{J \subset 2^{I-1}} F_i^J,$$

   where F_i^J *denotes the information which only the agents* $j \in J$ *are aware of. If* J *is singleton, then the agent therein has an information monopoly. We can decompose* $\overleftarrow{g}_i$ *in the following way*

   $$\overleftarrow{g}_i = \sum_{J \subset 2^{I \setminus \{i\}}} 1_{F_i^J} \cdot \overleftarrow{g}_i \, .$$

   (b) Determining the protocol entries: *In order to fix the protocol entries, we have to specify what happens on* F_i^J *if* J *is non singleton. In this case, delivery of information to agent* i *is shared among the agents equally. By virtue of of (a), the protocol entry on the information that agent* $k \neq i$ *sells to agent* i *can now be formulated as*

   $$g_i(k) = \sum_{J \in 2^{I \setminus \{i\}} : k \in J} \frac{1}{|J|} 1_{F_i^J} \overleftarrow{g}_i. \tag{14}$$

   The procedure in (a) and (b) can be summarized as a function $f_i : [0, 1_{F_i^c}] \to [0, 1_{F_i^c}]^{I-1}$, *given by* $f_i(\overleftarrow{g}_i) = \{g_i(k)\}_{k \in I \setminus \{i\}}$, *that is continuous in view of (14).*[14]

3. Price player: *Let prices be normalized. Define the price player's correspondence* $\mathbb{P} : [0, 1_H]^I \Rightarrow \Delta_{|H|}$ *by*

   $$\mathbb{P}\Big(\{\overleftarrow{g}_i\}\Big) = \arg \max_{p \in \Delta_{|H|}} \sum_i p(\overleftarrow{g}_i).$$

   Clearly, $\mathbb{P}$ *is nonempty–, convex–, compact–valued and upper hemi–continuous, by Berge maximum Theorem.*

4. Fixed–point: *The correspondence* $(\mathbb{D}_i) \times \mathbb{P} : [0, 1_H]^I \times \Delta_{|H|} \Rightarrow [0, 1_H]^I \times \Delta_{|H|}$ *has a fixed–point* $(\overleftarrow{g}^*_1, \ldots, \overleftarrow{g}^*_I, p^*)$ *by the usual application of the Kakutani fixed–point theorem and step 1 and 3.*

[14]Given the initial information allocation, the procedure allows to decompose the demand for information of agent i into the entries of i–th row of the IEP matrix.

5. Feasibility: *Consider a fixed–point of step 4. By the budget set restriction in (10), we have* $p^*(\overleftarrow{g}^*_i) \leq p^*(\overrightarrow{g}^*(i))$ *for every agent* i*. This yields*

$$\begin{aligned} 0 &\geq \sum_i p^*\Big(\overleftarrow{g}^*_i - \overrightarrow{g}^*(i)\Big) \\ &= p^*\Big(\sum_i \overleftarrow{g}^*_i - \overrightarrow{g}^*(i)\Big) \\ &\geq p\Big(\sum_i \overleftarrow{g}^*_i - \overrightarrow{g}^*(i)\Big), \end{aligned}$$

for every $p \in \Delta_{|H|}$*. We have* $\sum_i \overleftarrow{g}^*_i \leq \sum_i \overrightarrow{g}^*(i)$*, i.e. the received information does not exceeds delivered information. Feasibility in the sense of the IEP matrix is accomplished by step 2.* ■

References

Allen, B. (1986): "The demand for (differentiated) information," *The Review of Eonomic Studies*, 53(3), 311–323.

——— (1990): "Information as an economic commodity," *The American Economic Review*, 80(2), 268–273.

Arrow, K. J. (1996): "The economics of information: An exposition," *Empirica*, 23(2), 119–128.

Aumann, R. (1974): "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, 1(1), 67–96.

Berliant, M. (1986): "A utility representation for a preference relation on a σ-algebra," *Econometrica*, pp. 359–362.

Bewley, T. F. (1972): "Existence of equilibria in economies with infinitely many commodities," *Journal of Economic Theory*, 4(3), 514–540.

Blackwell, D. (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, 24(2), 265–272.

Chade, H., and E. Schlee (2002): "Another look at the Radner–Stiglitz nonconcavity in the value of information," *Journal of Economic Theory*, 107(2), 421–452.

Chateauneuf, A., and J.-C. Vergnaud (2000): "Ambiguity reduction through new statistical data," *International Journal of Approximate Reasoning*, 2(24), 283–299.

Cotter, K. D. (1986): "Similarity of information and behavior with a pointwise convergence topology," *Journal of Mathematical Economics*, 15(1), 25–38.

De Castro, L. I., and N. C. Yannelis (2010): "Ambiguity aversion solves the conflict between efficiency and incentive compatibility," Discussion paper.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2008): "Attitude toward imprecise information," *Journal of Economic Theory*, 140(1), 27–65.

Geanakoplos, J. (2010): "The leverage cycle," in *NBER Macroeconomics Annual*, vol. 24, pp. 1–65. University of Chicago Press.

Gilboa, I., and E. Lehrer (1991): "The value of information-An axiomatic approach," *Journal of Mathematical Economics*, 20(5), 443–459.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

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

Gul, F., and W. Pesendorfer (2015): "Hurwicz expected utility and subjective sources," *Journal of Economic Theory*, 159, 465–488.

Hervés-Beloso, C., and P. K. Monteiro (2013): "Information and sigma-algebras," *Economic Theory*, 54(2), 405–418.

Hirshleifer, J. (1971): "The private and social value of information and the reward to inventive activity," *The American Economic Review*, 61(4), 561–574.

Khan, M. A., Y. Sun, R. Tourky, and Z. Zhang (2008): "Similarity of differential information with subjective prior beliefs," *Journal of Mathematical Economics*, 44(9), 1024–1039.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73(6), 1849–1892.

MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 1447–1498.

MAS-COLELL, A., AND W. ZAME (1991): "Equilibrium theory in infinite dimensional spaces," *Handbook of Mathematical Economics*, 4, 1835–1898.

RADNER, R. (1968): "Competitive equilibrium under uncertainty," *Econometrica*, 36(1), 31–58.

RADNER, R., AND J. STIGLITZ (1984): "A Nonconcavity in the Value of Information," *Bayesian models in economic theory*, 5(3).

STINCHCOMBE, M. B. (1990): "Bayesian information topologies," *Journal of Mathematical economics*, 19(3), 233–253.

TIAN, G., AND J. ZHOU (1992): "The maximum theorem and the existence of Nash equilibrium of (generalized) games without lower semicontinuities," *Journal of Mathematical Analysis and Applications*, 166(2), 351–364.

VAN ZANDT, T. (1993): "The Hausdorff metric of σ-fields and the value of information," *The Annals of Probability*, 21(1), 161–167.