Entropy and the Value of Information for Investors: The Prior-Free Implications

Ran I. Shorrer*

December 21, 2015

Abstract

Blackwell dominance is a partial order over information structures that is based on the uniform preference of all decision makers for one structure over another in all decision problems. I restrict attention to investment decision problems with ruin-averse investors (Cabrales, Gossner and Serrano, 2013), and compare substantive rejections of information transactions (uniform over all utilities and prior beliefs). This yields the partial order of *prior-independent investment dominance*, which strictly extends Blackwell's order. Thus, many pairs of information structures that are not ranked by the Blackwell criterion may still be ranked in investment problems independently of the prior.

Understanding the demand for information is crucial for understanding many important economic environments. Yet, comparing the desirability of different information structures in a sensible way is an elusive undertaking. The reason is that in some settings certain pieces of information may be vital for some agents, while in other settings, or for different agents, other pieces of information will be more important.[1] The implication is that it is not possible to rank all information structures so that higher-ranked structures are preferred to lower-ranked ones by all agents at every decision-making problem and for all prior beliefs. Some pairs of information structures may, however, be compared in this manner. In his seminal paper, Blackwell (1953) showed that one information structure is preferred to another by all agents in all settings if and only if the latter is a garbling of the former, that is, if one is a noisy version of the other.[2] But this order is partial and cannot be used to compare many pairs of information structures.[3]

*Harvard University, Department of Economics, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138, rshorrer@gmail.com.

[1]The difficulty in generating a ranking that is independent of agents' preferences is discussed in Willinger (1989), which studies the relation between risk aversion and the value of information. Willinger (1989) discusses his choice of using the *expected value of information* (EVI) or *"asking price,"* which was defined by LaValle (1968). The EVI measures a certain decision maker's willingness to pay for certain information, thereby *"... the difficulty of defining a controversial continuous variable representing the 'amount of information' can be avoided."*

[2]A simple proof is provided in Leshno and Spector (1992).

[3]Lehmann (1988), Persico (2000), Athey and Levin (2001), Jewitt (2007), and others have extended this partial order by restricting the class of decision problems and agents under consideration.

Cabrales, Gossner and Serrano (2013) use an approach of total rejections in the spirit of Hart (2011) to make such comparisons.[4] In a model à la Arrow (1972), they restrict attention to investment decision problems and define a relation that they term *uniform investment dominance*. This turns out to be a complete order over all *information structures*, which extends the order suggested by Blackwell (1953). Their order, however, depends on the (unique, fixed, common) prior of the decision makers, and thus they get, rather, a continuum of orders, one for each prior. These orders are indeed different from each other; there exists pairs of information structures that are ranked differently depending on the prior selected. This means that *prior-independent investment dominance* is a partial order.

This paper treats a question that was left unanswered in Cabrales, Gossner and Serrano (2013); namely, is prior-independent investment dominance the same as Blackwell's partial order, or does it provide further insights for *prior-free* comparisons of information structures in *investment settings*? This question is important since an analyst cannot always observe the priors of agents in the market. My answer is that the latter alternative is correct: I prove that (many) pairs of information structures that *cannot* be compared by Blackwell's order *can* be compared by the partial order of prior-independent investment dominance. I also provide a complete characterization of these pairs of structures restricting attention to information structures with two states of the world and two signals. Two such structures chosen uniformly at random are comparable by prior-independent investment dominance with probability approximately .94, compared with probability 2/3 by Blackwell's criterion.

1 Preliminaries

I use the model and notation of Cabrales, Gossner and Serrano (2013). I provide a brief review, as a complete discussion can be found in their paper.[5]

I consider agents with a concave, twice continuously differentiable utility function for money, who have some initial wealth, w, and face uncertainty about the state of nature. There are $K \in \mathbb{N}$ states of nature,[6] $\{1, ..., K\}$, over which the agents have the prior $p \in \Delta(K)$, which is assumed to have a full support.

I identify agents with utility functions, and denote the *Arrow–Pratt coefficient of relative risk aversion* of agent u at wealth w by[7]

$$\varrho_u(w) := -w\frac{u''(w)}{u'(w)}.$$

[4] In fact, they follow Hart's (2011) *utility uniform rejections*, which lead in his setting to the index of riskiness suggested in Foster and Hart (2009). In their 2014 paper they follow Hart's (2011) *wealth uniform rejections*, which lead in his setting to the index of riskiness suggested in Aumann and Serrano (2008). This second approach leads to their index of appeal of information transactions.

[5] I present a simplified version of Cabrales, Gossner and Serrano (2013). Simplifications are for ease of exposition, and have no effect on any of my results.

[6] With a slight abuse of notation, I also denote $\{1, ..., K\}$ by K. The meaning of K should be clear from the context.

[7] See Pratt (1964) and Arrow (1965, 1971).

I restrict attention to agents with relative risk aversion that is increasing in their wealth (IRRA). This means that $\varrho_u(\cdot)$ is non decreasing for all agents considered. Justifications for this assumption include theoretical considerations (Arrow, 1971) and observed behavior in the field (Binswanger, 1981; Post et al., 2008). IRRA utility functions include *constant absolute risk aversion* (CARA) utilities, as well as *constant relative risk aversion* (CRRA) utilities. I further focus on agents that are *ruin averse*, namely, agents that satisfy $\lim_{w \to 0^+} u(w) = -\infty$. I denote by $\mathcal{U}^*$ the class of these utility functions.

The set of *investment opportunities* $B^* = \left\{ b \in \mathbb{R}^K \;\middle|\; \sum_{k \in K} p_k b_k \leq 0 \right\}$ consists of all no-arbitrage assets. When an agent with initial wealth w chooses investment $b \in B^*$ and state k is realized, his wealth becomes $w + b_k$. Hence, B^* includes in particular the option of inaction. Referring to the members of B^* as no-arbitrage investment opportunities attributes to the prior, p, an additional interpretation as the price of an Arrow–Debreu security that pays 1 if the state k is realized and nothing otherwise. Hence, p plays a dual role in this setting.

Agents may choose their investment freely, but bankruptcy (the possibility of negative wealth) is not allowed. I say that an investment b is *feasible* at wealth w when $w + b_k \geq 0$ in every state $k \in K$, and denote by $B_F^w := \{b \in B^* \mid \; w + b_k \geq 0 \;\; \forall k \in K\}$ the set of investment opportunities that are feasible at wealth w. Before choosing a feasible investment, the agent has an opportunity to engage in an *information transaction* (μ, α), where $\mu > 0$ is the cost of the transactions, and α is the information structure representing the information that it entails. To be more precise, α is given by a finite set of signals S_α and probability distributions $\alpha_k \in \Delta(S_\alpha)$ for every $k \in K$. When the state of nature is k, the probability that the signal s is observed equals $\alpha_k(s)$. Thus, the information structure may be represented by a stochastic matrix M_α, with K rows and $|S_\alpha|$ columns, and the total probability of the signals is given by the vector $p_\alpha := p \cdot M_\alpha$. For simplicity, assume that $p_\alpha(s) > 0$ for all s, so that each signal is observed with positive probability. Further, denote by q_k^s the probability that the agent assigns to state k conditional on observing the signal s, using Bayes' law. Note that although the notation does not indicate it, $(q_k^s)_{k=1}^K = q^s \in \Delta(K)$ depends on α and the prior p.

Agents choose the optimal feasible investment opportunity in B^* given their belief and their wealth. Therefore, the expected utility of an agent with utility u, initial wealth w, and belief q is[8]

$$V(u, w, q) = \sup_{b \in B_F^w} \sum_k q_k u(w + b_k).$$

In the case where the agent acquires no information, his belief is given by the prior p. Since the agent is risk averse, in such a case his optimal choice is inaction. Hence,

$$V(u, w, p) = u(w).$$

[8]Throughout, I use the convention that $(-\infty) \cdot 0 = 0$.

Accordingly, an agent u with wealth w *accepts* an information transaction (μ, α) if

$$\sum_s p_\alpha(s) V(u, w - \mu, q^s) > V(u, w, p) = u(w)$$

and *rejects* it otherwise.

The *entropy reduction* is defined by:

$$I(\alpha, p) = H(p) - \sum_s p_\alpha(s) \cdot H(q^s),$$

where

$$H(q) = -\sum_{k \in K} q_k \ln(q_k),$$

and $0 \ln 0 = 0$ by continuity.[9]

Definition. For a fixed prior p, information structure α *uniformly investment-dominates* (or *investment-dominates*, for short) information structure β whenever, for every wealth w and price $\mu < w$ such that (μ, α) is rejected by all agents with utility $u \in \mathcal{U}^*$ at wealth w, β is also rejected by all those agents.

Theorem. *[Cabrales, Gossner and Serrano] For a fixed prior p, information structure α investment-dominates information structure β if and only if $I(\alpha, p) \geq I(\beta, p)$.*

Corollary. *If α Blackwell-dominates β, then $I(\alpha, p) \geq I(\beta, p)$ for all p.*

Definition. An information structure α *investment-dominates β independently of the prior* (or *prior-independently investment-dominates*), whenever α investment-dominates β for any prior p.

Theorem. *[Cabrales, Gossner and Serrano] There exists no linear ordering that orders information structures according to the ordering of investment dominance independently of the prior.*

Example 1. Let $K = \{1, 2, 3\}$ and let $p^1 = (.5 - \epsilon, .5 - \epsilon, 2\epsilon)$ and $p^2 = (2\epsilon, .5 - \epsilon, .5 - \epsilon)$. Consider the information structures

$$\alpha^1 = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \\ .5 & .5 \end{bmatrix}, \ \alpha^2 = \begin{bmatrix} .5 & .5 \\ 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}.$$

It is easy to verify that for $\epsilon > 0$ sufficiently small $I(\alpha^1, p^1) > I(\alpha^2, p^1)$, but $I(\alpha^1, p^2) < I(\alpha^2, p^2)$. This is intuitive since, given p^i, α^i contains (almost) all the information that an investor could hope for, while α^{-i} could be improved upon significantly.

[9] Recall that q^s is not independent of the prior p, even though the dependence is not made explicit by the notation I use.

2 Results

While information structures cannot be linearly ordered according to investment dominance independently of the prior, the corollary above suggests that this relation is not vacuous. There are some cases where one information structure investment-dominates another for any prior, for example, when the former Blackwell-dominates the latter. A natural question that is left unanswered in Cabrales, Gossner and Serrano (2013) is whether these are the only cases. In other words, are prior-independent investment dominance and Blackwell dominance the same? I answer this question in the negative.

Theorem 1. *There exist α and β such that α investment-dominates β independently of the prior, but α does not dominate β according to Blackwell's order.*[10]

Proof. Follows from Example 2. □

Example 2. Let $K = \{1, 2\}$ and consider the information structures

$$\alpha_1 = \begin{bmatrix} .3 & .7 \\ .7 & .3 \end{bmatrix}, \alpha_2 = \begin{bmatrix} .3 & .7 \\ .1 & .9 \end{bmatrix}.$$

I claim that $I(\alpha_1, p) \geq I(\alpha_2, p)$ for all p. I identify p with p_1, the probability of state 1 that lies in $[0, 1]$. Fixing the two information structures, I define a function $\phi_{\alpha_1,\alpha_2} : [0, 1] \longrightarrow \mathbb{R}$ as follows:

$$\phi_{\alpha_1,\alpha_2}(\cdot) := I(\alpha_2, \cdot) - I(\alpha_1, \cdot).$$

For $p \in \{0, 1\}$, $I(\cdot, p) \equiv 0$, and hence $\phi_{\alpha_1,\alpha_2}(0)=\phi_{\alpha_1,\alpha_2}(1) = 0$. It follows from the properties of I that $\phi_{\alpha_1,\alpha_2}(\cdot)$ is continuous and twice continuously differentiable. Furthermore,

$$\phi''_{\alpha_1,\alpha_2}(p) = \frac{0.0252 - 0.0192p}{(-0.3 - 0.4p)(0.7 - 0.4p)(0.3(-1 + p) - 0.1p)(1 + 0.3(-1 + p) - 0.1p)}. \tag{2.0.1}$$

This expression is always positive for $p \in (0, 1)$, which implies that $\phi_{\alpha_1,\alpha_2}(\cdot)$ is a strictly convex and continuous function with $\phi_{\alpha_1,\alpha_2}(0) = \phi_{\alpha_1,\alpha_2}(1) = 0$. But this means that $\phi_{\alpha_1,\alpha_2}(p) < 0$ for all $p \in (0, 1)$, hence $I(\alpha_2, p) - I(\alpha_1, p) \leq 0$ for all $p \in [0, 1]$, and hence α_1 *investment-dominates* α_2 *independently of the prior.*

It remains to show that α_1 does not Blackwell-dominate α_2. Let us look at the geometry of Blackwell dominance more generally. Given a $K \times S$ information structure α, the set of all $K \times S$ information structures dominated by α is defined as $\text{Dom}(\alpha) := \{\alpha M : M \in (\Delta(S))^S\}$. Namely, α multiplied by M, where M ranges over all $S \times S$ stochastic matrices. As a linear image of the polytope $(\Delta(S))^S$, $\text{Dom}(\alpha)$ is a polytope whose vertices are images of the vertices of $(\Delta(S))^S$.

[10] Using Example 2, Shorrer (2015) shows that the same applies to the index of appeal of information transactions (Cabrales, Gossner and Serrano, 2014).

Namely,

$$\mathrm{Dom}(\alpha) = \mathrm{conv}\left\{ \alpha \begin{bmatrix} e_{i_1} \\ \vdots \\ e_{i_S} \end{bmatrix} : \ e_{i_j} \text{are vertices of } \Delta(S) \right\}.$$

The case of just two states ($K = 2$) and two signals ($S = 2$) is particularly simple. The set of all 2×2 information structures dominated by a given 2×2 information structure form a parallelogram:

$$\mathrm{Dom}\left(\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix} \right) = \mathrm{conv}\left\{ \begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}, \begin{bmatrix} y & 1-y \\ x & 1-x \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

From Figure 1, it is easily seen that the information structure α_2 from the example is not Blackwell-dominated by α_1.

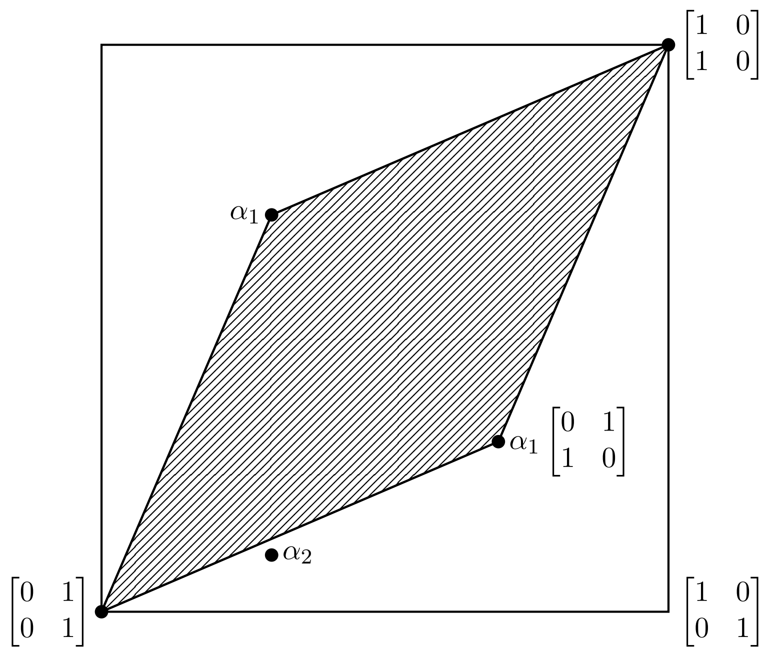

Figure 1: The figure depicts the two-dimensional space of 2×2 information structures. These matrices could be written as $\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$, where both x and y are in $[0, 1]$. In the figure, x is represented by the horizontal axis and y is represented by the vertical axis. The shaded area are the matrices that represent information structures dominated by α_1 in the Blackwell sense. The point α_2 is outside the shaded area.

The above counterexample extends to any number of states and signals. I now focus on the case of 2×2 information structures (2 states of the world and 2 signals), and provide a complete characterization of comparable pairs.

Definition 1. Given two 2×2 information structures α and β, the function $\phi_{\alpha,\beta} : [0,1] \longrightarrow \mathbb{R}$ is defined as follows:

$$\phi_{\alpha,\beta}(\cdot) := I(\beta, \cdot) - I(\alpha, \cdot).$$

Theorem 2. *For 2×2 information structures α and β, α investment-dominates β independently of the prior if and only if*

$$\phi'_{\alpha,\beta}\left(0^+\right) \le 0, \text{ and}$$
$$\phi'_{\alpha,\beta}\left(1^-\right) \ge 0.$$

Furthermore, α and β are comparable using this partial order if and only if

$$\phi'_{\alpha,\beta}\left(0^+\right)\phi'_{\alpha,\beta}\left(1^-\right) \le 0.$$

Proof. By definition, α investment-dominates β independently of the prior if and only if $\phi_{\alpha,\beta}$ is non-positive on the interval $[0,1]$. My proof is a generalization of the investigation of the function ϕ_{α_1,α_2} in the proof of Theorem 1. The main step is to show that $[0,1]$ can be divided into two intervals, $[0,t]$ and $[t,1]$ (with t possibly equal to 0 or 1), such that in one of these intervals $\phi_{\alpha,\beta}$ is convex and in the other it is concave. In other words, $\phi_{\alpha,\beta}$ is either convex, concave, or "S-shaped": convex on one side of t and concave on the other side. From this analysis and the fact that $\phi_{\alpha,\beta}(0) = \phi_{\alpha,\beta}(1) = 0$, one readily concludes the first part of the theorem. The second part follows from the first, since $\phi_{\alpha,\beta} = -\phi_{\beta,\alpha}$.

I now turn to the proof that for every pair of 2×2 structures α and β, there exist $t \in [0,1]$ such that $\phi_{\alpha,\beta}$ is convex in one of the intervals $[0,t]$ and $[t,1]$ and concave in the other.

Denote

$$\alpha = \begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}, \ \beta = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix},$$

for some x, y, a, $b \in [0,1]$. The function $\phi_{\alpha,\beta}$ is continuously twice differentiable on $[0,1]$. Assume first that neither of the information structures is degenerate, that is, neither equals $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ or $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. Direct computation shows that

$$\phi''_{\alpha,\beta}(p) = \frac{(x-y)^2 p_\beta(1)p_\beta(2) - (a-b)^2 p_\alpha(1)p_\alpha(2)}{p_\alpha(1)p_\alpha(2)p_\beta(1)p_\beta(2)}. \tag{2.0.2}$$

The denominator is positive on $(0,1)$, as a product of four positive factors. The numerator is an affine function in p: it is the difference between two quadratic polynomials that have the same quadratic term.[11] This implies that the entire expression could change signs at most once. This concludes the proof in the non-degenerate case.

[11]The numerator is equal to $(x-y)^2(b+(a-b)p)(1-b+(b-a)p) - (a-b)^2(y+(x-y)p)(1-y+(y-x)p)$.

If both α and β are degenerate, then $\phi_{\alpha,\beta} \equiv 0$; if only α is degenerate, then

$$\phi''_{\alpha,\beta}(p) = \frac{-(a-b)^2}{p_\beta(1)p_\beta(2)};$$

and if only β is degenerate, then

$$\phi''_{\alpha,\beta}(p) = \frac{(x-y)^2}{p_\alpha(1)p_\alpha(2)}.$$

In all three of the degenerate cases $\phi''_{\alpha,\beta}$ has the same sign throughout the interval $(0,1)$. $\square$

Remark. Using this theorem and the Monte Carlo method, I conclude that two 2×2 information structures drawn uniformly at random are comparable with probability approximately 0.94, compared with a 2/3 probability that they are comparable using Blackwell's criterion.

The following theorem provides a sufficient condition that is much simpler to verify and illustrate than the condition of Theorem 2.

Theorem 3. *For a non degenerate information structure* $\alpha = \begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$ *and an information structure* $\beta = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix}$, *$\alpha$ investment-dominates β independently of the prior if*

$$(x-y)^2 a(1-a) - (a-b)^2 x(1-x) \geq 0,$$

and

$$(x-y)^2 b(1-b) - (a-b)^2 y(1-y) \geq 0.$$

Note that Theorem 1 follows from Theorem 3. The condition in Theorem 3 specifies an intersection of two ellipses, which is a strictly convex set; therefore any non extreme point on the relative boundary of $\mathrm{Dom}(\alpha)$ is an internal point of the set of information structures investment-dominated by α independently of the prior.

Proof. It is sufficient to show that $\phi_{\alpha,\beta}$ is convex in $[0,1]$, or, equivalently, $\phi''_{\alpha,\beta}$ is non negative in $(0,1)$. As seen in (2.0.2), since the denominator is always positive, it is sufficient to show that the numerator is non negative. Namely,

$$L(p) = (x-y)^2 p_\beta(1)p_\beta(2) - (a-b)^2 p_\alpha(1)p_\alpha(2) \geq 0 \quad \forall p \in (0,1).$$

Since $L(p)$ is a linear function of p, one need only verify that the two end points, at $p = 0, 1$, are non negative, which is exactly the condition of the theorem. $\square$

Remark. Using this sufficient condition, one can show that two 2×2 information structures drawn uniformly at random are comparable with probability approximately 0.85.

3 Discussion

Several previous papers extend Blackwell's partial order by restricting attention to a certain class of decision-making problems (Lehmann, 1988; Persico, 2000; Athey and Levin, 2001; Jewitt, 2007). Cabrales, Gossner and Serrano (2013) consider a different class of decision problems and, for a fixed prior, derive a complete order that is represented by the expected decrease in entropy from the prior to the posteriors. This paper shows that while the prior-free version of their order is incomplete, it still suggests an interesting new approach for extending Blackwell's order.

References

Arrow, Kenneth J. 1965. *Aspects of the Theory of Risk-Bearing.* Helsinki: Yrjö Jahnssonin Säätiö.

Arrow, Kenneth J. 1971. *Essays in the Theory of Risk Bearing.* Chicago: Markham.

Arrow, Kenneth J. 1972. "The Value of and Demand for Information." *Decision and Organisation. Londres: North-Holland.*

Athey, Susan, and Jonathan Levin. 2001. "The Value of Information in Monotone Decision Problems." *Stanford University.*

Aumann, Robert J., and Roberto Serrano. 2008. "An Economic Index of Riskiness." *Journal of Political Economy*, 116(5): 810–836.

Binswanger, Hans P. 1981. "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India." *The Economic Journal*, 91(364): 867–890.

Blackwell, David. 1953. "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics*, 24(2): 265–272.

Cabrales, Antonio, Oliver Gossner, and Roberto Serrano. 2014. "The Inverse Demand for Information and the Appeal of Information Transactions." *Working paper.*

Cabrales, Antonio, Olivier Gossner, and Roberto Serrano. 2013. "Entropy and the Value of Information for Investors." *The American Economic Review*, 103(1): 360–377.

Foster, Dean P., and Sergiu Hart. 2009. "An Operational Measure of Riskiness." *Journal of Political Economy*, 117(5): 785–814.

Hart, Sergiu. 2011. "Comparing Risks by Acceptance and Rejection." *Journal of Political Economy*, 119(4): 617–638.

Jewitt, Ian. 2007. "Information Order in Decision and Agency Problems." *Nuffield College.*

LaVallе, Irving H. 1968. "On Cash Equivalents and Information Evaluation in Decisions under Uncertainty: Part I: Basic Theory." *Journal of the American Statistical Association*, 63(321): 252–276.

Lehmann, Erich L. 1988. "Comparing Location Experiments." *The Annals of Statistics*, 521–533.

Leshno, Moshe, and Yishay Spector. 1992. "An Elementary Proof of Blackwell's Theorem." *Mathematical Social Sciences*, 25(1): 95–98.

Persico, Nicola. 2000. "Information Acquisition in Auctions." *Econometrica*, 68(1): 135–148.

Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show." *American Economic Review*, 98(1): 38–71.

Pratt, John W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica*, 32(1/2): 122–136.

Shorrer, Ran I. 2015. "Consistent Indices." *Mimeo.*

Willinger, Marc. 1989. "Risk Aversion and the Value of Information." *Journal of Risk and Insurance*, 56(1): 104–112.