Topic 24: Convexity and incentive design

Convex functions appear naturally in many incentive design problems.

24.1 Proper scoring rules

Scoring rules are used to elicit probabilistic beliefs from forecasters, and had their origins in the evaluation of weather forecasters.[1] The term comes from the use of "skill scores"[2] to evaluate weather forecasters. We shall describe the mathematical results here, but you should read the papers by Murphy and Winkler [11] and Savage [15] for a discussion of the practical aspects as well.

We start with a set S of states of the world, which for simplicity, we assume is finite. (Vectors in $\boldsymbol{R}^S$ have two interpretations: as random variables and as (signed) measures on S. For infinite state spaces, the space of random variables and the space of measures are not the same.) Let $\Delta = \{p \in \boldsymbol{R}^S_+ : p \cdot \mathbf{1} = 1\}$ denote the set of probability measures on S, viewed as vectors in $\boldsymbol{R}^S$. Thus the expected value of a random variable $x \in \boldsymbol{R}^S$ under the probability $p \in \Delta$ is simply $p \cdot x$.

An "expert" or "forecaster" is asked to state his or her subjective probability measure p and will be awarded a "score" $\xi_s(p)$ on the basis of the forecast p and what state s occurs. The expert is assumed to care only about the expectation of his score. If the forecaster's probability is p, and he reports q, his expected score is $p \cdot \xi(q) = \sum_s p_s \xi_s(q)$.

24.1.1 Definition *A **scoring rule** for Δ is a function $\xi \colon \Delta \to \boldsymbol{R}^S$. The value $\xi(p)$ is a random variable. The s^{th} component ξ_s is the **score** in state s. A scoring rule ξ is **proper** if for each $p \in \Delta$,*

$$p \cdot \xi(p) \geqslant p \cdot \xi(q) \quad \text{for all } q \in \Delta.$$

*The scoring rule ξ is **strictly proper** if for each $p \in \Delta$,*

$$p \cdot \xi(p) > p \cdot \xi(q) \quad \text{for all } q \neq p,\ q \in \Delta.$$

[1] The practice of expressing weather forecasts in terms of rough probabilities was initiated in Western Australia by W. E. Cooke in 1905 [3]. Interestingly, his idea was criticized by E. B. Garriot [4] of the U.S. because "the bewildering complication of uncertainties it involves would confuse even the patient interpolator" and "our public insist upon having our forecasts expressed concisely and in unequivocal terms."

[2] A skill score is intended to separate a forecaster's skill from the difficulty of the forecasting problem. For instance, Harris K. Telemacher, the "wacky weatherman" portrayed by Steve Martin in the film *L.A. Story* was able to accurately forecast the weather days in advance, not because he was a good forecaster, but because, as we all know, the weather here is always perfect. You can find various skill score measures described at the American Meteorological Society's Online Glossary.

In other words a proper scoring gives a risk neutral forecaster the incentive to truthfully report his subjective belief. A strictly proper scoring rules gives the forecaster a strict incentive to do so. (After all, a constant ξ is proper.)

Observe that if ξ is a (strictly) proper scoring rule, so is a positive affine transformation $\alpha\xi + \beta\mathbf{1}$, where $\alpha > 0$ and $\mathbf{1}$ is, as usual, the vector with all its components equal to 1.

A natural question is whether strictly proper scoring rules exist. It is generally asserted that this question was answered affirmatively in 1950 by Brier [2], but I find it hard to follow his argument. In 1951, Good [5] proved that

$$\xi_s(p) = \ln p_s$$

defined a strictly proper scoring rule. Finally in 1956, McCarthy [10] stated without proof the following result (more or less). A full proof, for a not necessarily finite state space, is provided by Hendrickson and Buehler [9].

24.1.2 Theorem *A function $\xi\colon \Delta \to \boldsymbol{R}^S$ is a strictly proper scoring rule if and only if there is a lower semicontinuous sublinear function $f\colon \boldsymbol{R}^S_+ \to \boldsymbol{R}$ that is strictly convex on Δ such that for every $p \in \Delta$,*

$$\xi(p) \in \partial f(p),$$

where as usual, $\partial f(p)$ denotes the subdifferential of f at p.

Proof: The proof of this result is surprisingly easy.

($\Longrightarrow$) Assume ξ is a strictly proper scoring rule. Define $f(0) = 0$ and for nonzero $p \in \boldsymbol{R}^S+$, define

$$f(p) = p \cdot \xi(p/\mathbf{1} \cdot p).$$

Then f is homogeneous of degree one by construction. From the definition of a strictly proper scoring rule, for distinct $\bar{q}$ and $\bar{p}$ in Δ and $\lambda > 0$, letting $q = \lambda\bar{q}$, we have

$$f(q) = f(\lambda\bar{q}) = \lambda\bar{q} \cdot \xi(\bar{q}) > \lambda\bar{q} \cdot \xi(\bar{p}) = q \cdot \xi(\bar{p}), \tag{1}$$

where the strict inequality is true because ξ is a strictly proper scoring rule. Note that when $\bar{p} = \bar{q}$ we have equality in (1). This tells us two things. First,

$$f(q) = \max_{\bar{p}\in\Delta} \ell_{\bar{p}}(q), \qquad \text{where } \ell_{\bar{p}}\colon q \mapsto \xi(\bar{p}) \cdot q \text{ is linear in } q.$$

Thus f is convex and lower semicontinuous on $\boldsymbol{R}^S_+$, since it is the pointwise maximum of linear functions.

The second thing we learn from (1) is that for $q = \lambda\bar{q}$,

$$\begin{aligned} f(q) > \lambda\bar{q} \cdot \xi(\bar{p}) &= \underbrace{f(\bar{p}) - \bar{p} \cdot \xi(\bar{p})}_{=0} + \lambda\bar{q} \cdot \xi(\bar{p}) \\ &= f(\bar{p}) + \xi(\bar{p}) \cdot (q - \bar{p}), \end{aligned}$$

for all distinct $\bar{q}, \bar{p}$ in Δ, which is just the subgradient inequality for $\xi(\bar{p})$ at $f(\bar{p})$. (This inequality holds as an equality for $q = 0$, since $f(\bar{p}) = \xi(\bar{p}) \cdot \bar{p}$.) That is,

$$\xi(\bar{p}) \in \partial f(\bar{p}) \quad \text{for all } \bar{p} \in \Delta.$$

Moreover, since the subgradient inequality is strict, Proposition 14.1.3 shows that f is strictly convex on Δ.

($\Longleftarrow$) Assume now that $f\colon \boldsymbol{R}^S_+ \to \boldsymbol{R}$ is a lower semicontinuous sublinear function that is strictly convex on Δ, and for each $p \in \Delta$,

$$\xi(p) \in \partial f(p).$$

By Euler's Theorem for subgradients 14.1.12, f satisfies $f(\lambda p) = \lambda p \cdot \xi(p)$ for all $p \in \Delta$. In particular,

$$f(p) = p \cdot \xi(p) \quad \text{for all } p \in \Delta.$$

We can now run the previous argument in reverse.

Let $p, q \in \Delta$, with $q \neq p$. Since f is strictly convex, the subgradient inequality holds strictly (again by Proposition 14.1.3), that is,

$$f(p) > f(q) + \xi(q) \cdot (p - q).$$

Replacing $f(p)$ by $p \cdot \xi(p)$ and $f(q)$ by $\xi(q) \cdot q$, this reduces to

$$p \cdot \xi(p) > p \cdot \xi(q).$$

In other words, ξ is a proper scoring rule. ∎

24.1.3 Remark When I first sat down to prove this result, I didn't see the point of defining f on all of $\boldsymbol{R}^S_+$, and requiring it to be homogeneous. In fact, the first part of the proof shows that if ξ is a strictly proper scoring rule, then the $f(p) = \xi(p) \cdot p$ is strictly convex on Δ, and that $\xi(p) \in \partial f(p)$. It is the second part of the theorem where we need the fact that f is defined on $\boldsymbol{R}^S_+$ and not just Δ. Here is the key point: Regarded as a function with domain Δ, $f|_\Delta$ may have more subgradient vectors than f as a function on $\boldsymbol{R}^S_+$. This is because the subgradient inequality $f(q) > f(p) + \xi(p)(q - p)$ need only hold for q in the domain of f. With a smaller domain, more vectors ξ can be subgradients, and these extra vectors may not be part of a strictly proper scoring rule. The homogeneity forces f to satisfy $f(p) = \xi(p) \cdot p$ for $p \in \Delta$, which is crucial to showing that the subgradient inequality reduces to the definition of strict properness of ξ.

24.1.4 Remark By the way, defining $g(p, q) = \xi(p) \cdot q$, we see that f is the optimal value function for the parametrized optimization problem of maximizing $g(p, q)$ with respect to q. If f is differentiable at p, the Envelope Theorem tells us that $f'(p) = \xi(p)$. This theorem extends that to the case where f is not necessarily differentiable.

24.1.5 Remark We have seen in Theorem 21.4.1 that every lower semicontinuous positively homogeneous function is the support function of a closed convex set. In this case, f is clearly the support function of $\overline{\mathrm{co}}\{\xi(p) : p \in \Delta\}$.

24.1.6 Example I asserted earlier that Good [5] proved that

$$\xi_s(p) = \ln p_s$$

defined a strictly proper scoring rule. So according to the proof of Theorem 24.1.2, it should be the subgradient of the differentiable function

$$f(p) = p \cdot \xi(p/\mathbf{1} \cdot p) = \sum_{s=1}^{S} p_s \ln\left(\frac{p_s}{\mathbf{1} \cdot p}\right) = \sum_{s=1}^{S} p_s \ln p_s - \sum_{s=1}^{S} p_s \ln(\mathbf{1} \cdot p).$$

Tedious differentiation[3] reveals that

$$\frac{\partial f(p)}{\partial p_s} = \ln p_s - \ln(\mathbf{1} \cdot p).$$

So when $p \in \Delta$, we have $\mathbf{1} \cdot p = 1$, and $\ln 1 = 0$, so

$$\nabla f(p) = \xi(p),$$

as promised. □

24.1.7 Exercise Verify that the **quadratic scoring rule** given by $\xi_s(p) = p_s - \frac{1}{2}p_s^2$ is indeed a strictly proper scoring rule. What is the associated sublinear function? □

24.1.8 Exercise A scoring rule ξ is called **symmetric** or **local** if the there is a function h such that for each state $s \in S$, $\xi_s(p) = h(p_s)$. If $|S| \geqslant 3$, show that the logarithmic scoring rule is the unique (up to positive affine transformation) local scoring rule. □

24.2 Private information and trade

The next example of the role of convexity is based on Myerson and Satterthwaite [13].

Consider the problem of trading a discrete object. The **seller** initially owns the object and values it at s. The **buyer** has value b. These values are not knowable

[3] Write $t(p) = \mathbf{1} \cdot p$, so $\partial t(p)/\partial p_k = 1$, $\partial \ln t(p)/\partial p_k = 1/t(p)$, for each k. Then writing $f(p) = \sum_i p_i \ln p_i - \sum_i p_i \ln t(p)$, we have

$$\frac{\partial f(p)}{\partial p_k} = (1 + \ln p_k) - \sum_i \bigl[p_i\bigl(1/t(p)\bigr) + \delta_{ik} \ln t(p)\bigr] = \ln p_k - \ln t(p).$$

to the other agent. If the buyer's value exceeds the seller's value, $b > s$, then they can trade at an intermediate price p and both will be better off. A mechanism for trading must determine whether to trade and at what price. Moreover it can rely only on information provided (perhaps indirectly) by the buyer and seller.

The current approach to modeling mechanisms like these was proposed by John Harsanyi [6, 7, 8]. His approach is called the **Bayesian game** model. In this model the buyer and seller act as if nature moves first and assigns the buyer and seller their values by drawing them from known probability distributions. A **strategy** is then a function from each trader's information (his value) to an action, depending on the rules of exchange. (Think of nature as dealing them a hand from a deck of cards. Traders should be prepared to act on whatever hand they possess.) A **Bayes–Nash equilibrium** is simply a Nash equilibrium of a Bayesian game.

We shall refer to actions in this game as **bids**. But a strategy is a **bidding function**. The bidding functions will be denoted $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}$, where $\boldsymbol{\beta}(b)$ is the bid made by a buyer with value b and $\boldsymbol{\sigma}(s)$ is the bid made by a seller with value s.

For concreteness we shall assume that the buyer believes that nature chooses s from a cumulative distribution function G, and the seller believes b is drawn from distribution F. That is, the buyer believes that $\operatorname{Prob}\{s \leqslant t\} = G(t)$ and the seller believes that $\operatorname{Prob}\{b \leqslant t\} = F(t)$. To simplify things, let's suppose that the distributions F and G have common support $[0, 1]$, and are continuously differentiable. This ensures that they have densities, so calculations are simpler. Let's also assume that the buyer and seller are risk neutral, and that prices and values are commensurable, and that b and s are stochastically independent.

24.2.1 The revelation principle

A **mechanism** for trading sets up a Bayesian game by specifying an action set for each trader and an outcome function. Let B be the buyer's set of actions. (In the split-the-difference rule, B is just the set of possible bids.) Let S denote the seller's action set. The mechanism specifies two functions:

$$t(\beta, \sigma) = \begin{cases} 1 \text{ if buyer gets object} \\ 0 \text{ if seller keeps object} \end{cases}$$
$$p(\beta, \sigma) = \text{payment from buyer to seller},$$

where $\beta \in B$ is the buyer's chosen action and $\sigma \in S$ is the seller's. Now let $\boldsymbol{\beta}^* \colon b \mapsto \boldsymbol{\beta}^*(b)$ denote the buyer's Bayes–Nash equilibrium strategy. That is, a buyer with value b chooses action $\boldsymbol{\beta}^*(b)$ in equilibrium. Similarly, $\boldsymbol{\sigma}^* \colon s \mapsto \boldsymbol{\sigma}^*(s)$ is the seller's equilibrium strategy. When the values are b and s, the buyer acquires the object if and only if $t\Big(\boldsymbol{\beta}^*(b), \boldsymbol{\sigma}^*(s)\Big) = 1$ and pays $p\Big(\boldsymbol{\beta}^*(b), \boldsymbol{\sigma}^*(s)\Big)$.

Consider another mechanism where $\hat{B}$ is the set of buyer's values and $\hat{S}$ is the set of seller's values. Define the outcome functions by

$$\hat{t}(b,s) = g\Big(\boldsymbol{\beta}^*(b), \boldsymbol{\sigma}^*(s)\Big)$$
$$\hat{p}(b,s) = p\Big(\boldsymbol{\beta}^*(b), \boldsymbol{\sigma}^*(s)\Big).$$

It follows from the definition of equilibrium that an equilibrium strategy for the buyer is to choose action b when his value is b, and for the seller to choose s when his value is s. In other words, the first mechanism is equivalent to a mechanism in which actions correspond to values and the equilibrium action choice is to choose the true value. This observation is known as the **revelation principle**.

24.2.2 Revelation mechanisms

Because of the revelation principle the only mechanisms that need to be considered are the **incentive compatible direct revelation mechanisms**. That is, mechanisms where strategies are values, and the equilibrium bidding functions are truthful, that is, $\boldsymbol{\beta}(b) = b$ and $\boldsymbol{\sigma}(s) = s$. Note that these bidding functions are strictly increasing and continuously differentiable.

For the sake of concreteness again take $B = S = [0,1]$, and assume

$$F(0) = G(0) = 0 \qquad \text{and} \qquad F(1) = G(1) = 1.$$

Let

$$\pi^B(b,\beta) = \int_0^1 \Big[b \cdot t(\beta,s) - p(\beta,s)\Big] G'(s)\,ds$$

be the buyer's expected payoff when his value is b and he bids β, and the seller bids his true value. Similarly let

$$\pi^S(s,\sigma) = \int_0^1 \Big[\Big(1 - t(b,\sigma)\Big)s + p(b,\sigma)\Big] F'(b)\,db$$

be the seller's expected payoff when his value is s and he bids σ and the buyer bids his true value.

By incentive compatibility, bidding you value is optimal, so

$$\pi^B(b,b) = \max_\beta \pi^B(b,\beta),$$

and denote this value by $V_B(b)$. Likewise

$$\pi^S(s,s) = \max_\sigma \pi^S(s,\sigma) = V_S(s).$$

Define

$$\bar{t}(b) = \int_0^1 t(b,s)G'(s)\,ds, \qquad \text{and} \qquad \bar{p}(b) = \int_0^1 p(b,s)G'(s)\,ds.$$

That is, $\bar{t}(b)$ is the probability that the buyer receives the object when his type is b, and $\bar{p}(b)$ is his expected payment (given that the seller is bidding his value). Then

$$\pi^B(b, \beta) = b\bar{t}(\beta) - \bar{p}(\beta).$$

So incentive compatibility for the buyer can be written as

$$V_B(b) = \pi^B(b, b) = b\bar{t}(b) - \bar{p}(b) \geqslant b\bar{t}(b') - \bar{p}(b') = b\bar{t}(b') - \Big[b'\bar{t}(b') - V_B(b')\Big],$$

or

$$V_B(b) \geqslant V_B(b') + \bar{t}(b')(b' - b). \tag{2}$$

This implies in fact

$$V_B(b) = \sup_{b'} V_B(b') + b'\bar{t}(b') - \bar{t}(b')b.$$

Now the function $b \mapsto V_B(b') + b'\bar{t}(b') - \bar{t}(b')b$ is an affine function of b, so V_B is convex, being the pointwise supremum of affine functions. Moreover, by (2), we see that for every point b', $\bar{t}(b')$ is a supergradient of V_B at b'. Thus for all but countably many values of b, the function V_B is differentiable and

$$V_B'(b) = \bar{t}(b).$$

In particular, since convex functions are integrals of their supergradients, we also know that

$$\begin{aligned} V_B(b) &= V_B(0) + \int_0^b \bar{t}(x)\,dx. \\ &= V_B(0) + \int_0^b \int_0^1 t(x, s)G'(s)\,ds\,dx \end{aligned} \tag{3}$$

A similar argument shows that

$$V_S(s) = V_S(1) - \int_s^1 \int_0^1 t(b, x)F'(b)\,db\,dx. \tag{4}$$

24.2.3 Inefficiency is inevitable when trade is voluntary

We now show that in general, voluntary participation and efficiency are incompatible. To do this, we need to make sure that the informational problem is nontrivial, in that we are not ex ante sure whether trade should occur. A sufficient condition is the following nondegeneracy condition:

$$\int_0^1 G(t)\Big(1 - F(t)\Big)\,dt > 0. \tag{5}$$

What this condition does is guarantee that there is a set of positive measure of values of t satisfying Prob $\{s \leqslant t\} > 0$ and Prob $\{b > t\} > 0$. This implies that there is positive probability that trade is optimal.

Traders will voluntarily participate only if their expected payoff is at least as great as not participating. That is,

$$V_B(b) \geqslant 0 \quad \text{and} \quad V_S(s) \geqslant s \quad \text{for all} b, s.$$

Now let us see what efficiency demands. An efficient outcome demands that

$$t(b,s) = 1 \iff b > s, \tag{6}$$

or

$$t(b,s) = \mathbf{1}_{b>s}.$$

Now by (3), (4), and (6) we have

$$\begin{aligned} V_B(b) &= V_B(0) + \int_0^b \int_0^1 t(x,s)G'(s)\,ds\,dx \\ &= V_B(0) + \int_0^b \int_0^1 \mathbf{1}_{x>s}G'(s)\,ds\,dx \\ &= V_B(0) + \int_0^b G(x)\,dx. \end{aligned}$$

Similarly

$$V_S(s) = V_S(1) - \int_s^1 F(b)\,db$$

Taking expectations,

$$EV_B = V_B(0) + \int_0^1 \int_0^b G(s)\,ds\,dF(b) \quad \text{and} \quad EV_S = V_S(1) - \int_0^1 \int_s^1 F(b)\,db\,dG(s).$$

Integrating by parts,

$$\int_0^1 \left\{ \int_0^b G(s)\,ds \right\} dF(b) = \int_0^1 G(s)\,ds - \int_0^1 G(b)F(b)\,db = \int_0^1 G(x)\big(1 - F(x)\big)\,dx$$

$$\int_0^1 \left\{ \int_s^1 F(b)\,db \right\} dG(s) = 0 + \int_0^1 G(x)F(x)\,dx.$$

Therefore

$$EV_B + EV_S = \left\{ V_B(0) + \int_0^1 G(x)\big(1 - F(x)\big)\,dx \right\} + \left\{ V_S(1) - \int_0^1 G(x)F(x)\,dx \right\}. \tag{7}$$

Now consider what happens when the buyer's value is b and the seller's value is s. The sum of the ex post payoffs is

$$b\mathbf{1}_{b>s} - p(b,s) + s(1 - \mathbf{1}_{b>s}) + p(b,s) = \max\{b, s\}.$$

Taking expectations with respect to both b and s,

$$\begin{aligned}
E(V_B + V_S) &= E \max\{b, s\} \\
&= \int_0^1 \int_0^1 \max\{b, s\} F'(b) G'(s)\, db\, ds \\
&= \int_0^1 \int_0^b b F'(b) G'(s)\, ds\, db + \int_0^1 \int_0^s s F'(b) G'(s)\, db\, ds \\
&= \int_0^1 b F'(b) G(b)\, db + \int_0^1 s F(s) G'(s)\, ds \\
&= \int_0^1 x \Big[F'(x) G(x) + G'(x) F(x) \Big]\, dx \\
&= x F(x) G(x) \Big|_0^1 - \int_0^1 F(x) G(x)\, dx \\
&= 1 - \int_0^1 F(x) G(x)\, dx \qquad (8)
\end{aligned}$$

Equating the two expressions (7) and (8) for $EV_B + EV_S$ gives

$$V_B(0) + \int_0^1 G(x)\Big(1 - F(x)\Big)\, dx + V_S(1) - \int_0^1 G(x) F(x)\, dx = 1 - \int_0^1 F(x) G(x)\, dx$$

so

$$V_B(0) + V_S(1) = 1 - \int_0^1 G(x)\Big(1 - F(x)\Big)\, dx$$

but by the voluntary participation constraints, $V_S(1) \geqslant 1$ and $V_B(0) \geqslant 0$, so

$$1 - \int_0^1 G(x)\Big(1 - F(x)\Big)\, dx \geqslant 1,$$

which contradicts nondegeneracy (5). ∎

References

[1] P. Bardsley. 2011. Duality in contracting. Working paper, Department of Economics, University of Melbourne.
http://www.hss.caltech.edu/~kcb/Ec181/pdf/Bardsley2011.pdf

[2] G. W. Brier. 1950. Verification of forecasts of expressed in terms of probability. *Monthly Weather Review* 78(1):1–3.
http://docs.lib.noaa.gov/rescue/mwr/078/mwr-078-01-0001.pdf

[3] W. E. Cooke. 1906. Forecasts and verifications in Western Australia. *Monthly Weather Review* 34(1):23–24.
http://docs.lib.noaa.gov/rescue/mwr/034/mwr-034-01-0023.pdf

[4] E. B. Garriott. 1906. Note by Prof. E. B. Garriott. *Monthly Weather Review* 34(1):24.
http://docs.lib.noaa.gov/rescue/mwr/034/mwr-034-01-0023.pdf

[5] I. J. Good. 1952. Rational decisions. *Journal of the Royal Statistical Society. Series B (Methodological)* 14(1):107–114.
http://www.jstor.org/stable/2984087

[6] J. C. Harsanyi. 1967. Games with incomplete information played by "Bayesian" players, I–III. Part I. The basic model. *Management Science* 14:159–182. http://www.jstor.org/stable/2628393

[7] ——— . 1968. Games with incomplete information played by "Bayesian" players, I–III. Part II. Bayesian equilibrium points. *Management Science* 14:320–334. http://www.jstor.org/stable/2628673

[8] ——— . 1968. Games with incomplete information played by "Bayesian" players, I–III. Part III. The basic probability distribution of the game. *Management Science* 14:486–502. http://www.jstor.org/stable/2628894

[9] A. D. Hendrickson and R. J. Buehler. 1971. Proper scores for probability forecasters. *Annals of Mathematical Statistics* 42:1916–1921.
http://www.jstor.org/stable/2240117

[10] J. McCarthy. 1956. Measures of the value of information. *Proceedings of the National Academy of Sciences, U.S.A.* 42(9):654–655.
http://www.pnas.org/content/42/9/654.full.pdf

[11] A. H. Murphy and R. L. Winkler. 1970. Scoring rules in probability assessment and evaluation. *Acta Psychologica* 34:273–286.
DOI: 10.1016/0001-6918(70)90023-5

[12] M. Mussa and S. Rosen. 1978. Monopoly and product quality. *Journal of Economic Theory* 18(2):301–317. DOI: 10.1016/0022-0531(78)90085-6

[13] R. B. Myerson and M. A. Satterthwaite. 1983. Efficient mechanisms for bilateral trading. *Journal of Economic Theory* 29(2):265–281.
DOI: 10.1016/0022-0531(83)90048-0

[14] J.-C. Rochet. 1987. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics* 16(2):191–200. DOI: 10.1016/0304-4068(87)90007-3

[15] L. J. Savage. 1971. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66(336):783–801.
http://www.jstor.org/stable/2284229

[16] R. V. Vohra. 2011. *Mechanism design: A linear programming approach*. Econometric Society Monographs. Cambridge and New York: Cambridge University Press.