Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation

Robin Hanson *
Department of Economics
George Mason University†

January 2002

Abstract

In practice, scoring rules elicit good probability estimates from individuals, while betting markets elicit good consensus estimates from groups. Market scoring rules combine these features, eliciting estimates from individuals or groups, with groups costing no more than individuals. Regarding a bet on one event given another event, only logarithmic versions preserve the probability of the given event. Logarithmic versions also preserve the conditional probabilities of other events, and so preserve conditional independence relations. Given logarithmic rules that elicit relative probabilities of base event pairs, it costs no more to elicit estimates on all combinations of these base events.

Introduction

In theory, probability elicitation is hard. For expected utility maximizers, choices are jointly determined by utilities and probabilities, and so without additional constraints on utilities or probabilities subjective probabilities cannot be separated from event-dependent utilities (Kadane & Winkler, 1988). Some sophisticated approaches can in theory overcome this problem (Jaffray & Karni, 1999; Hanson, 2002b). Yet in practice, less sophisticated uses of scoring rules and related methods are widely and successfully used to elicit informative event probabilities from individuals in weather forecasting (Murphy & Winkler, 1984), economic forecasting (O'Carroll, 1977), risk analysis (DeWispelare, Herren, & Clemen, 1995), and the engineering of intelligent computer systems (Druzdzel & van der Gaag, 1995). Furthermore, simple conversational statements are often taken as reliable belief elicitations.

Given an ability to elicit probabilities, in theory it should be easy to induce individuals to aggregate their information into common estimates.[1] Given a finite state space, Bayesians with a common prior who repeatedly state their beliefs, and who have common knowledge about previous statements, must eventually achieve common knowledge of future statements (Geanakoplos & Polemarchakis, 1982), and thus common estimates (Aumann, 1976). Along the way, Bayesians

*For their comments, I thank Chris Hibbert and David Pennock.

†*rhanson@gmu.edu* http://hanson.gmu.edu 703-993-2326 FAX: 704-993-2323 MSN 1D3, Carow Hall, Fairfax VA 22030

[1] This can be true even if directly combining probability distributions is more difficult (Genest & Zidek, 1986).

cannot even predict the direction in which others will disagree with them (Hanson, 2002a). Even if they are tempted to strategically lie along the way, repeatedly paying each Bayesian by a scoring rule should also eventually produce a Nash equilibrium of this game (Kalai & Lehrer, 1993), and thus common knowledge of future statements. Similar results apply when common knowledge is replaced by common belief (Monderer & Samet, 1989), and for "Bayesian wannabes" (Hanson, 1997).

In practice, however, repeated exchanges of human opinion in conversation do not produce the degree of convergence that theory predicts for Bayesians (Cowen & Hanson, 2002). Even so, speculative markets, such as stock, commodities, and futures markets, seem to do a remarkable, if sometimes incomplete, job of aggregating available relevant information into market prices (Lo, 1997). Betting markets, in particular, create good probability estimates (Hausch, Lo, & Ziemba, 1994). And this seems to be true even though speculation in such markets seems to be irrational for most participants. For example, orange juice futures prices improve on government weather forecasts (Roll, 1984). Markets created within Hewlett-Packard Labs beat official corporate sales forecasts 6 out of 8 times, and tied the other time (Plott, 2000). The Iowa Electronic Markets, which predict U.S. presidential elections, were more accurate than major opinion polls in 451 out of 591 comparisons (Berg, Nelson, & Rietz, 2001). Even play money markets have done well at forecasting movie returns and scientific progress (Pennock, Giles, & Nielsen, 2001).

Since theory has not been a completely reliable guide here, we might look to the empirical successes of scoring rules and betting markets, and try to combine the best of these approaches while avoiding their failings. Simple scoring rules do not induce different individuals to form common estimates, while simple betting markets cannot create price estimates unless several people coordinate to bet on the same event, and it typically seems irrational to participate. Market scoring rules, in contrast, act like simple scoring rules when one person estimates an event once, yet can also act like a subsidized betting market with which many people can and rationally should repeatedly interact to produce a common estimate.

With a simple scoring rule, a person reports a probability for each event, and gets paid depending on that report and the actual event. Market scoring rules are scoring rules where anyone can change the current report, and be paid according to their new report, as long as he agrees to pay the last person reporting according to that person's report. This in effect lets anyone make any infinitesimal fair bet at the odds in the last report, and make any integral of such bets through changing betting odds. The cost to create a market scoring rule depends only on the informativeness of the last report given it, realtive to its starting position, and does not otherwise depend on how often or how many people use it. In contrast with a standard betting market, rational agents should expect a positive profit from participating, and do not need to find another person willing to make a matching bet.

There are a vast number of events about which one might want probability estimates. There are not only many simple events, but there are far more possible combinations of these simple events. How does the cost of a market scoring rule depend on the number of events it covers? Also, how modular can making changes be? That is, if different people specialize in estimating different events, how easily can each person make changes to estimates where they think they have expertise, while minimizing unintended changes to other estimates?

These considerations reflect favorably on logarithmic versions of market scoring rules. Once one has paid to create a logarithmic rule on some base set of event pairs, there is no additional cost to having that rule apply to all possible combinations of those base events. Furthermore, if someone makes a bet with a logarithmic rule on the conditional probability of one event given another, with a logarithmic rule the probability of the given event does not change, nor do the conditional probabilities of further events. Logarithmic rules thus only change the probabilities of events where the person betting took a risk. This supports the use of conditional independence relations, a popular mechanism humans use to manage the complexity of large probability spaces. With logarithmic rules, bets on some events do not change conditional independence relations between other events.

For example, given a logarithmic rule that induces an estimate of the chance of rain Monday, the chance of rain Tuesday, and the chance of rain Wednesday, it costs no more to also induce estimates such as whether it will rain Tuesday and Wednesday given that it rains Monday. And if someone bets that it will rain Tuesday given that it rains Monday, this bet will not change the estimated chance of rain Monday. If current estimates say that the chance of rain Wednesday is independent of Monday's chance, given Tuesday's chance, a bet on rain Tuesday given rain Monday also will not change this independence relation.

This paper will first review scoring rules, and then discuss market scoring rules, their costs, and their modularity.

Scoring Rules

Consider an expected-utility maximizing agent with subjective beliefs p_i regarding a complete set of I disjoint events i, where $\sum_i p_i = 1$. This agent might be paid a cash amount x according to a *proper scoring rule* $x_i = s_i(\vec{r})$ (Savage, 1971). Here x_i is the cash payment if i turns out to be the actual event, r_i is the probability the agent reports for the event i, and $\vec{r} = \{r_i\}_i$ is the full report. When $\vec{s} = \{s_i\}_i$ constitutes a proper[2] scoring rule, an agent who sets his reports r_i to maximize his expected monetary payoff will honestly report $r_i = p_i$. That is,

$$\vec{p} = \mathrm{argmax}_{\vec{r}} \sum_i p_i\, x_i = \sum_i p_i\, s_i(\vec{r}) \;\text{ given }\; \sum_i r_i = 1.$$

Expected monetary payoff will be maximized by risk-neutral agents with event-independent utility. Other agents can also be induced to maximize an expected monetary payoff.[3] Scoring rules can also give agents incentives to acquire information they would not otherwise possess (Clemen, 2002).

First and second order conditions for this maximization are that for all events j,

$$\lambda = \sum_i p_i \bigtriangledown_j s_i(\vec{p}) \tag{1}$$

$$0 \geq \sum_i p_i \bigtriangledown_j \bigtriangledown_j s_i(\vec{p}) \tag{2}$$

where λ is a Lagrange multiplier, and $\bigtriangledown_j$ is a partial derivative operator with respect to component j. Differentiating this first order condition with respect to p_k, and applying $\bigtriangledown_k \bigtriangledown_j = \bigtriangledown_j \bigtriangledown_k$, gives

[2] For a non-proper scoring rule, other reports also maximize an agent's expected payoff.

[3] Agents paid in tickets for a lottery with extreme consequences should maximize expected payoff if their utilities approach common lower and upper bounds, (Jaffray & Karni, 1999). Bayesians paid in lottery tickets after playing a certain insurance game should maximize their expected payoff relative to their objective beliefs, which are what they would believe had they updated a known common prior with their further personal information (Hanson, 2002b).

$\bigtriangledown_j s_i = \bigtriangledown_i s_j$, which implies $s_i(\vec{r}) = \bigtriangledown_i g(\vec{r})$ for some function $g(\vec{r})$ (Williamson, Crowell, & Trotter, 1972). The function $g(\vec{p})$ turns out to be the part of an agent's equilibrium expected payoff that depends on his report $\vec{r}$ (Savage, 1971).

Examples of proper scoring rules include

$$\begin{aligned}
\text{Quadratic} \quad s_i = &\quad a_i + br_i - b\sum_j r_j^2/2, \\
\text{Spherical} \quad s_i = &\quad a_i + b\,r_i/(\sum_j r_j^2)^{1/2}, \\
\text{Logarithmic} \quad s_i = &\quad a_i + b\log(r_i), \\
\text{Power Law} \quad s_i = &\quad a_i + b\alpha\int_0^{r_i} \rho_i^{\alpha-2} d\rho_i - b\sum_j r_j^\alpha
\end{aligned}$$

Power law rules are proper scoring rules for $\alpha \geq 1$ (Selten, 1998), and both the quadratic (Brier, 1950) and logarithmic rules (Good, 1952) are special cases of this, for α of 2 and 1 respectively.

The quadratic rule satisfies a number of desirable properties (Selten, 1998). The logarithmic rule also satisfies desirable properties (von Holstein, 1970). For example, it is the only rule satisfying the constraint $\bigtriangledown_j s_i = 1_{ij}\nu_i$ for some ν_i (where $1_{ij} = 1$ when $i = j$ and 0 otherwise), which implies $s_i(\vec{r}) = s_i(r_i)$, i.e., that an agent's payoff depends only on the probability he assigned to the actual event (Savage, 1971). This rule is thus the only one that can simultaneously reward agents and evaluate them via standard likelihood methods (Winkler, 1969).

It can be convenient to extend scoring rules $s_i(\vec{r})$, defined so far for normalized $\vec{r}$ satisfying $\sum_i r_i = 1$, to cases where $\sum_i r_i \neq 1$. This can be done by substituting $r_i/\sum_j r_j$ for p_i. For example, the extended logarithmic rule is $s_i = a_i + b_i \log(r_i/\sum_j r_j)$. An equivalent approach is to require, for all positive α, that $\vec{s}(\alpha\vec{r}) = \vec{s}(\vec{r})$. This implies $0 \leq \bigtriangledown_i \bigtriangledown_i g$ and $0 = \sum_i p_i \bigtriangledown_i \bigtriangledown_j g$, and that $\lambda = 0$ in equation 1. For example, for the extended logarithmic rule, $g(\vec{r}) = b\sum_i r_i \log(r_i/\sum_j r_j)$.

Market Scoring Rules

In many existing markets, all trades are made with one or a few central actors called market makers. These actors always have public offers to buy or to sell, and update these prices in response to trades. Human market makers have been found to do as well as the standard double auction market form at aggregating information (Krahnen & Weber, 1999). For the price of a modest subsidy, automated market makers can also play this role in supporting trade. For example, automated market makers have been used successfully by the Hollywood Stock Exchange (www.hsx.com) to promote speculation on the prospects of thousands of movies and movie stars.

It has long been known that in one dimension, allowing an agent to interact with a proper scoring rules is equivalent to allowing him to choose a quantity from a continuous offer demand schedule (Savage, 1971). This paper shows how this equivalence also holds in higher dimensions. A simple scoring rule variation, the market scoring rule, acts like a continuous automatic market maker with which an arbitrary number of agents can have an arbitrary number of interactions, at no additional cost over that of the one last interaction.

For any scoring rule $s_i(\vec{r})$, an agent should voluntarily agree to accept a payment of the form

$$x_i = \Delta s_i(\vec{r}, \vec{\rho}) = s_i(\vec{r}) - s_i(\vec{\rho})$$

for any value of $\vec{\rho}$. After all, an agent can ensure himself no effect ($\vec{x} = 0$) by setting $\vec{r} = \vec{\rho}$, and expects a positive (and maximal) profit if he sets $\vec{r} = \vec{p} \neq \vec{\rho}$ (all normalized). Thus if $\vec{\rho}$ is repeatedly set to the last report made, an arbitrary number of agents can be allowed an arbitrary number of interactions. That is, let agents make their reports $\vec{r}_t$ one at a time to a *market scoring rule*, where each report is paid $x_{it} = \Delta s_i(\vec{r}_t, \vec{r}_{t-1})$, with $\vec{r}_0$ an initial reference report. The total cost to pay for T reports,

$$x_i = \sum_{t=1}^{T} x_{it} = \sum_{t=1}^{T} (s_i(\vec{r}_t) - s(\vec{r}_{t-1})) = s_i(\vec{r}_T) - s(\vec{r}_0),$$

depends only on the initial and final reports, and is thus the same as the cost for one final report with the same final values r_i.

Not only can the total movement from $\vec{r}_0$ and $\vec{r}_T$ be split at no additional cost into smaller movements from $\vec{r}_{t-1}$ to $\vec{r}_t$, each smaller movement can be thought of an integral of infinitesimal movements $d\vec{r}$ along a line of changing reports $\vec{r}(t)$ as t varies continuously. If an agent changes his report at a rate of $q_i = dr_i/dt$, the rate of change in his asset amounts are $y_i = dx_i/dt = \sum_j q_j \bigtriangledown_j s_i$. When this agent has beliefs $\vec{p}$, the rate of change in his expected payoff is thus

$$\frac{d}{dt} \sum_i p_i x_i = \sum_i p_i y_i = \sum_i p_i \sum_j q_j \bigtriangledown_j s_i(\vec{r}) = \sum_j q_j (\sum_i p_i \bigtriangledown_j s_i(\vec{r})).$$

When $\vec{r} = \vec{p}$, this last term in parentheses is zero according to the first order condition (equation 1). Thus the first order condition is really a local "fair bet" condition, saying that the assets exchanged as one changes one's report are locally a fair (i.e., zero expected value) bet at the current "market" prices $\vec{r}$. An agent pays assets of the form "Pays \$1 if event i holds" in exchange for other assets of the same form.

We can thus think of a market scoring rule as a continuous inventory-based automated market maker. Such a market maker has a zero bid-ask spread, at least for infinitesimal trades, has an internal state described entirely by its inventory of assets $\vec{x}$, and offers an instantaneous price of $\vec{p} = \vec{m}(\vec{x})$, where $\sum_i m_i(\vec{x}) = 1$. That is, such a market maker will accept any "fair bet" infinitesimal trade $d\vec{x} = \vec{y}\, dt$, such that

$$\sum_i y_i \, m_i(\vec{x}) = 0, \tag{3}$$

and accept any finite trade that is an integral of such infinitesimal trades.

The main task of any market maker is to extract the information implicit in the trades others make with it, in order to infer new rational prices (O'Hara, 1997). In response to an infinitesimal trade, the "inference rule" of a market scoring rule is $\bigtriangledown_i m_j$, which determines fair price changes via $dp_i/dt = q_i = \sum_j y_j \bigtriangledown_j m_i$. This inference rule should satisfy $\bigtriangledown_i m_i < 0$ to compensate for an expected adverse selection in trades. That is, people buying suggests that the market maker's price is probably too low, and people selling suggests the price is too high.

Market scoring rules produce consensus estimates in the same way that betting markets produce consensus estimates. While each person is always free to change the current estimate, doing so requires taking on more risk, and eventually everyone reaches a limit where they do not want to make further changes, at least not until they receive further information. At this point the market can be said to be in equilibrium.

We can extend $m(\vec{x})$ to all possible $\vec{x}$ by requiring that $\vec{m}(\vec{x}+\alpha\vec{1}) = \vec{m}(\vec{x})$ for all α, where $\vec{1} = \{1\}_i$. This says that changing the cash reserves of a market maker does not change its prices. This and $\sum_i m_i = 1$ imply

$$\sum_i \nabla_i m_j = 0 \tag{4}$$

$$\sum_i \nabla_j m_i = 0. \tag{5}$$

The market maker $\vec{m}$ that is equivalent to a particular proper scoring rule then satisfies $\vec{p} = \vec{m}(-\vec{s}(\vec{p}))$ when $\sum_i p_i = 1$. The negative sign is because an agent's gains are the market maker's losses. For example, a logarithmic scoring rule $s_i = a_i + b\log(p_i)$ corresponds to an exponential market maker

$$m_i(\vec{x}) = \frac{\exp((-a_i - x_i)/b)}{\sum_j \exp((-a_j - x_j)/b)}, \tag{6}$$

which is characterized by the differential equation

$$\nabla_i m_j = -m_i(1_{ij} - m_j)/b. \tag{7}$$

Note that an equivalence between these market makers and scoring rules ensures that such a market maker cannot be exploited for arbitrarily large profits; traders can as a group only take what they could get by making some report $\vec{r}$ to the equivalent proper scoring rule.

Costs of Market Scoring Rules

The net assets given to all the agents who interact with a market scoring rule is $x_i = s_i(\vec{r}_T) - s(\vec{r}_0)$, where $\vec{r}_T$ is the last report made, and $\vec{r}_0$ is an initial reference report. A principal funding the scoring rule minimizes her expected payments by setting the initial report to her initial beliefs $\vec{\pi}$, as in $\vec{r}_0 = \vec{\pi}$.

In the extreme case where agents eventually become sure of the actual state i, the principal's expected payment is $\sum_i \pi_i \Delta s_i(\vec{1}_i, \vec{\pi})$, where $\vec{1}_i = \{1_{ij}\}_j$. For the logarithmic scoring rule, this maximum expected payment is the entropy, $-b\sum_i \pi_i \log(\pi_i)$, of the initial distribution $\vec{\pi}$. In less than extreme cases, if the principal accepts the probabilities estimates of the final report, her expected payment is proportional to the difference between the entropies of the initial and final distributions. (For the quadratic scoring rule, the maximum expected payment is $b - b\sum_i \pi_i^2$.) Of course these cost calculations neglect costs to communicate current prices, to compute price and asset changes, and to implement transactions.

A space of possible events i might be constructed as the *combinatorial* product space of N base variables n, each of which has V_n possible values v. In this case there would be $I = \prod_n V_n$ possible events i, and events could be written as $i = \{v_n\}_n$, where v_n is a particular value of base variable n. If agents reported only on the probability of base variable values, they would make reports r_{nv} such that $\sum_v r_{nv} = 1$. These base-only reports would have a maximum expected cost of $\sum_{nv} p_{nv} \tilde{s}_v(\vec{1}_v, \vec{p}_n)$, where p_{nv} is a sum of p_i over events i which satisfy $v_n = v$. For the logarithmic scoring rule, the maximum expected cost for the full combinatorial report $\vec{r} = \{r_i\}_i$, which reports on the probability of all base variable value combinations, is *no more* than the cost for the base-only reports, at least when the parameter b is held constant. After all, the entropy of the full distribution is no more than the sum of the entropies of the marginal distributions.

While there is no direct additional financial cost to funding combinatorial reports, it remains difficult to bound the computational cost of updating prices and assets. The computational complexity of such updates can be large, being worse than polynomial in the worst case (i.e., is NP-complete) (Cooper, 1990). It remains an open question how to devise market scoring rules that minimize such computational costs.

Modularity of Market Scoring Rules

There are a vast number of events, combinatorial and otherwise, for which one might want probability estimates. An important practical consideration for market scoring rules is thus how well they help people manage large event spaces, and help people change the estimates where they think they have relative expertise, while minimizing unintended and uninformed changes to other estimates.

These considerations can be formalized in terms of how well trade regarding some events preserves conditional independence relations regarding other events. Conditional independence is a central tool for managing complex probability distributions. For variables $\mathcal{A}, \mathcal{B}, \mathcal{C}$, we say that in distribution P, variable $\mathcal{A}$ is independent of $\mathcal{C}$ given $\mathcal{B}$, and write $I(\mathcal{A}, \mathcal{B}, \mathcal{C})$, when

$$P(A_i|B_jC_k) = P(A_i|B_j)$$

for all values A_i of $\mathcal{A}$, B_j of $\mathcal{B}$, and C_k of $\mathcal{C}$. ($I(\mathcal{A}, \mathcal{B}, \mathcal{C})$ implies $I(\mathcal{C}, \mathcal{B}, \mathcal{A})$). Humans often find it difficult to state probability estimates, yet can typically express conditional independence relations quickly and confidently. Humans also change probability estimates more often than they change estimated independence relations. Such relations typically determine a sparse graph of related events, producing a vast reduction in the dimensionality of the resulting probability space. Humans can also frequently locate their areas of relative expertise within such graphs (Pearl, 1988; Pennock & Wellman, 2000).

Consider the case of an agent who bets with a market scoring rule, and bets only on the event A given the event B. That is, this agent gains assets of the form "Pays \$1 if A and B hold", in trade for assets of the form "Pays \$1 if B holds and A does not. Recall that $y_i = dx_i/dt$. If we describe these events A and B as sets of finer events i, this implies $y_i = y_j$ for all $i, j \in A \cap B$, $y_i = y_j$ for all $i, j \in \bar{A} \cap B$ (where $\bar{A}$ denotes the complement of A), and $y_i = 0$ for all $i \in \bar{B}$. In general, depending on the particular market scoring rule, such a bet might change any probability estimate p_i, and thus change any event probability $p(C) = \sum_{i \in C} p_i$. It seems preferable, however, for this bet to change as little as possible besides $p(A|B)$ (and of course $p(\bar{A}|B) = 1 - p(A|B)$).

That is, the market maker's inference rule should assume that a new bet on A given B in general gives it new information only about how the event A might depend on the event B, but no information on the probability of B, or on events unrelated to how A might depend on B. For such unrelated events, previous independence relations should be preserved.

The logarithmic market scoring rule is *local* in this strong sense. Consider binary variables $\mathcal{A}, \mathcal{B}, \mathcal{C}$, with values $A, \bar{A}$ for $\mathcal{A}$, values $B, \bar{B}$ for $\mathcal{B}$, and values $C, \bar{C}$ for $\mathcal{C}$. We can prove the following (proofs in the Appendix).

Theorem 1 *Logarithmic rule bets on A given B preserve $p(B)$, and for any event C preserve $p(C|AB)$, $p(C|\bar{A}B)$, and $p(C|\bar{B})$, and thus preserve $I(\mathcal{A}, \mathcal{B}, \mathcal{C})$, and $I(\mathcal{B}, \mathcal{A}, \mathcal{C})$.*

When there are at least three events i, the inverse also holds; the logarithmic market scoring rule is the only local rule, even in the weak sense that a bet on A given B preserves $p(B)$. Any rule that satisfies this constraint must do so in the simplest case of only two events. In this simple case, $A = \{j\}$, $B = \{j, k\}$, only dx_j and dx_k are non-zero, and only dp_j and dp_k should be non-zero. We can prove that only the logarithmic rule satisfies this constraint.

Theorem 2 *For $I \geq 3$, if $y_i = 0$ for $i \notin \{j, k\}$ implies $q_i = 0$ for $i \notin \{j, k\}$, the rule is logarithmic.*

Thus the logarithmic market scoring rule is unique in having a local inference rule. When someone makes a bet on A conditional on B, and hence takes no risk regarding whether B is true, all other market scoring rules will sometimes change their estimate of the probability of B. The logarithmic rule, in contrast, not only preserves $p(B)$, but also preserves $p(C|AB)$, $p(C|\bar{A}B)$, and $p(C|\bar{B})$ for all events C.

Conclusion

Simple scoring rules are regularly used to elicit probability estimates from individuals, and more sophisticated versions are available if needed to overcome risk-aversion and state-dependent utility. In theory, repeatedly eliciting and announcing individual estimates should be sufficient to induce common estimates, but in practice estimate differences remain. In practice, however, standard betting markets do elicit common estimates that seem to aggregate individual information well, even though participation seems irrational and requires a coordination of trading activity.

With a simple scoring rule, a person reports a probability for each event, and gets paid depending on that report and the actual event. Market scoring rules are scoring rules where anyone can change the official report, and be paid according to that new report, as long as they are willing to pay the last person reporting according to their report. This in effect lets anyone make any infinitesimal fair bet at the odds in the last report, with no need to find another person willing to make a matching bet, as in ordinary betting markets. Market scoring rules thus combine the advantages of both simple scoring rules and betting markets, inducing each individual to make an estimate and inducing common estimates.

Market scoring rules cost no more to implement than simple scoring rules. The cost does depend on the number of base events for which probability estimates are invited, but for logarithmic rules there is no additional cost to elicit estimates on all combinations of these base events. The logarithmic rule is also unique in making very local inferences from the trades it sees; regarding a bet on one event given another event, only a logarithmic rule preserves the probability of the given event. Logarithmic versions also preserve the conditional probabilities of further events, and so preserve conditional independence relations.

The computational costs of updating market scoring rules in combinatorial event spaces can be large, however, and it remains an open question how to best minimize and allocate such costs.

Appendix

Proof of Theorem 1

Logarithmic market scoring rules are exponential market makers which satisfy equation 7. For a bet on A given B, $y_i = dx_i/dt = 0$ for $i \in \bar{B}$. Putting these together, for $i \in \bar{B}$, we have

$$dp_i/dt = \sum_j y_j \bigtriangledown_j m_i = -m_i \sum_j y_j 1_{ij}/b + m_i \sum_j y_j m_j/b.$$

The first sum is zero because $y_i = 0$, and the second sum is zero by the fair bet equation 3. Thus the $p(\bar{B}) = \sum_{i \in \bar{B}} p_i$ does not change, nor does $p(B) = 1 - p(\bar{B})$.

According to equation 6, a transition from old prices p_i to new prices p'_i due to asset changes Δx_i satisfies $p'_i = p_i K \exp(-\Delta x_i/b)$, for some $K(\Delta\vec{x})$. Thus

$$p'(C|AB) = \frac{\sum_{i \in A\cap B\cap C} p'_i}{\sum_{i \in A\cap B} p'_i} = \frac{\sum_{i \in A\cap B\cap C} p_i K \exp(-\Delta x_i/b)}{\sum_{i \in A\cap B} p_i K \exp(-\Delta x_i/b)} = \frac{\sum_{i \in A\cap B\cap C} p_i}{\sum_{i \in A\cap B} p_i} = p(C|AB),$$

since Δx_i is the same everywhere within $A \cap B$. Since Δx_i is also the same everywhere within $\bar{A} \cap B$ and within $\bar{B}$, $p'(C|\bar{A}B) = p(C|\bar{A}B)$ and $p'(C|\bar{B}) = p(C|\bar{B})$ as well. Thus the logarithmic rule preserves $P(C|AB)$, $P(C|\bar{A}B)$, and $P(C|\bar{B})$ for any C.

For a binary variable $\mathcal{C}$, we can write $P(A|BC) = P(A|B)$ as $P(A|BC) = P(A|B\bar{C})$. For binary variables, $I(\mathcal{A}, \mathcal{B}, \mathcal{C})$ (equivalent to $I(\mathcal{C}, \mathcal{B}, \mathcal{A})$) requires $p(C|AB) = p(C|\bar{A}B)$ and $p(C|A\bar{B}) = p(C|\bar{A}\bar{B})$, and similarly for $\bar{C}$ instead of C. A logarithmic rule preserves $p(C|AB) = p(C|\bar{A}B)$ by preserving each side of this equation. A logarithmic rule also preserves both sides of $p(C|A\bar{B}) = p(C|\bar{A}\bar{B})$ because, for example, $p(C|A\bar{B}) = p(CA|\bar{B})/P(A|\bar{B})$, and both numerator and denominator here are of the form $p(D|\bar{B})$ that is preserved. $I(\mathcal{B}, \mathcal{A}, \mathcal{C})$ (equivalent to $I(\mathcal{C}, \mathcal{A}, \mathcal{B})$) requires $p(C|AB) = p(C|A\bar{B})$ and $p(C|\bar{A}B) = p(C|\bar{A}\bar{B})$, and similarly for $\bar{C}$ instead of C. But all of these forms are similarly preserved by a logarithmic rule. QED.

Proof of Theorem 2

Let the events j, k be $1, 2$. Since we've assumed $y_i = 0$ for $i \notin \{1, 2\}$, the fair bet equation (3) requires $\sum_i y_i m_i = y_1 m_1 + y_2 m_2 = 0$. By assumption, when this is true we must have, for $i \notin \{1, 2\}$, $0 = q_i = y_1 \bigtriangledown_1 m_i + y_2 \bigtriangledown_2 m_i$. These together imply

$$\frac{\bigtriangledown_1 m_i}{m_1} = \frac{\bigtriangledown_2 m_i}{m_2}.$$

Since our choice of $1, 2$ as the non-zero variables here was arbitrary, we can call the value of this equation $h_i(\vec{x})$. Thus for all $i \neq j$ we have $\bigtriangledown_j m_i = m_j h_i$. For $i = j$ we can define a residual μ_i from this expression, as in $\bigtriangledown_i m_i = h_i m_i + m_i \mu_i(\vec{x})$. Substituting these last two equations into equation 4 gives $h_j = -\mu_i m_j$, which implies that $\mu_i = \mu$, independent of i. This fact allows us to write, for all i, j,

$$\bigtriangledown_i m_j = \mu\, m_i (1_{ij} - m_j). \quad (8)$$

If we knew μ to be a constant function, this would be the same as the differential equation (7) that characterizes the logarithmic rule, and so we would be done. We can show that μ is constant by substituting equation 8 into this equation

$$\bigtriangledown_k \bigtriangledown_i m_j = \bigtriangledown_i \bigtriangledown_k m_j, \tag{9}$$

which must hold for all functions m. Considering the case $i \neq j \neq k$ gives $(\bigtriangledown_k \mu)/m_k = (\bigtriangledown_i \mu)/m_i$, which implies $\bigtriangledown_i \mu = m_i \eta$ for some $\eta(\vec{x})$. Substituting equation 8 into equation 9 and summing over j, k gives $\sum_k \bigtriangledown_k \mu = 0$. These together imply $\nu = 0$, which makes μ constant. QED.

References

Aumann, R. (1976). Agreeing to Disagree. *The Annals of Statistics*, *4*(6), 1236–1239.

Berg, J., Nelson, F., & Rietz, T. (2001). Accuracy and Forecast Standard Error of Prediction Markets. Tech. rep., University of Iowa, College of Business Administration.

Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, *78*, 1–3.

Clemen, R. T. (2002). Incentive Contracts and Strictly Proper Scoring Rules. *Test*, *11*, 195–217.

Cooper, G. F. (1990). The computational complexity of probabilistic inference using Bayes belief networks. *Artificial Intelligence*, *42*, 393–405.

Cowen, T., & Hanson, R. (2002). Are Disagreements Honest?. Tech. rep., George Mason University Economics.

DeWispelare, A. R., Herren, L. T., & Clemen, R. T. (1995). The use of probability elicitation in the high-level nuclear waste regulation program. *International Journal of Forecasting*, *11*(1), 5–24.

Druzdzel, M. J., & van der Gaag, L. C. (1995). Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information. In *Uncertainty in Artificial Intelligence*, pp. 141–148.

Geanakoplos, J., & Polemarchakis, H. (1982). We Can't Disagree Forever. *Journal of Economic Theory*, *28*, 192–200.

Genest, C., & Zidek, J. V. (1986). Combining Probability Distributions: A Critique and an Annotated Bibliography. *Statistical Science*, *1*(1), 114–135.

Good, I. J. (1952). Rational Decisions. *Journal of the Royal Statistical Society. Series B (Methodological)*, *14*(1), 107–114.

Hanson, R. (2002a). Disagreement is Unpredictable. *Economics Letters*, *77*, 365–369.

Hanson, R. (2002b). Eliciting Objective Probabilities via Lottery Insurance Games. Tech. rep., George Mason University Economics. http://hanson.gmu.edu/elicit.pdf.

Hanson, R. D. (1997). *Four Puzzles in Information and Politics: Product Bans, Informed Voters, Social Insurance, and Persistent Disagreement*. Ph.D. thesis, California Institute of Technology.

Hausch, D. B., Lo, V. S., & Ziemba, W. T. (1994). *Efficiencey of Racetrack Betting Markets*. Academic Press, San Diego.

Jaffray, J.-Y., & Karni, E. (1999). Elicitation of Subjective Probabilities When the Initial Endowment is Unobservable. *Journal of Risk and Uncertainty*, *18*(1), 5–20.

Kadane, J. B., & Winkler, R. L. (1988). Separating Probability Elicitation From Utilities. *Journal of the American Statistical Association*, *83*(402), 357–363.

Kalai, E., & Lehrer, E. (1993). Rational learning leads to Nash equilibrium. *Econometrica*, *61*(5), 1019–1045.

Krahnen, J. P., & Weber, M. (1999). Does information aggregation depend on market structure? Market makers vs. double auction. *Zeitschrift fur Wirtschaftsund Sozialwissenschaften*, *119*, 1–22.

Lo, A. W. (1997). *Market efficiency: Stock market behaviour in theory and practice*. Elgar, Lyme.

Monderer, D., & Samet, D. (1989). Approximating Common Knowledge with Common Beliefs. *Games and Economic Behavior*, *1*, 170–190.

Murphy, A. H., & Winkler, R. L. (1984). Probability Forecasting in Meterology. *Journal of the American Statistical Association*, *79*(387), 489–500.

O'Carroll, F. M. (1977). Subjective Probabilities and Short-Term Economic Forecasts: An Empirical Investigation. *Applied Statistics*, *26*(3), 269–278.

O'Hara, M. (1997). *Market Microstructure Theory*. Blackwell.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kauffmann, San Mateo, California.

Pennock, D. M., Giles, C. L., & Nielsen, F. A. (2001). The Real Power of Artificial Markets. *Science*, *291*, 987–988.

Pennock, D. M., & Wellman, M. P. (2000). Compact Secutiries Markets for Pareto Optimal Reallocation of Risk. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pp. 481–488 San Francisco. Morgan Kaufmann. http://rome.exp.sis.pitt.edu/UAI/Abstract.asp?articleID=56.

Plott, C. R. (2000). Markets as Information Gathering Tools. *Southern Economic Journal*, *67*(1), 2–15.

Roll, R. (1984). Orange Juice and Weather. *American Economic Review*, *74*(5), 861–880.

Savage, L. J. (1971). Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association*, *66*(336), 783–801.

Selten, R. (1998). Axiomatic Characterization of the Quadratic Scoring Rule. *Experimental Economics*, *1*(1), 43–62.

von Holstein, C.-A. S. S. (1970). Measurement of Subjective Probability. *Acta Psychologica*, *34*, 146–159.

Williamson, R., Crowell, R., & Trotter, H. (1972). *Calculus of Vector Functions* (Third edition). Prentice-Hall, Englewood Cliffs, New Jersey.

Winkler, R. L. (1969). Scoring Rules and the Evaluation of Probability Assessors. *Journal of the American Statistical Association*, *64*(327), 1073–1078.