TERNARY BELIEF ELICITATION[†]

PAUL J. HEALY[*] & GREG LEO[**]

ABSTRACT. There is an apparent trade-off in belief elicitation methodologies used in experimental economics. Of the two main mechanisms, binarized scoring rules have stronger incentives, but probability price lists require weaker preference assumptions. We resolve this trade-off by introducing a new methodology, the ternary price list. The ternary price list has incentives as strong as those of even the best binarized scoring rules but with the weaker incentive-compatibility requirements of price lists. Furthermore, in the limit, the uniform ternary price list's incentives are identical to those of the binarized quadratic scoring rule.

Keywords: Belief Elicitation; Experimental Methods

JEL Classification: C90, D83

Preliminary
January 4, 2025

[†] [*]Dept. of Economics, The Ohio State University; healy.52@osu.edu.
[**]Dept. of Economics, Loyola Marymount University; greg.leo@lmu.edu.

I. Introduction

There are two key mechanisms used to elicit probabilistic beliefs in experimental economics, binarized scoring rules and price lists.[1]

Binarized scoring rules map the elicited belief and realized/true value into an lottery which has a payoff-probability increasing in accuracy. For example, the binarized quadratic scoring rule [BQSR] incentivizes participants with a lottery where the payoff probability is inversely proportional to the squared prediction error.

In a probability price list, participants are given a set of N pairs of options and asked to choose one option in each pair. The pairs consist of an objective lottery, which pays a prize with an objective probability p, and an act that pays a prize if event E occurs. These pairs are often presented in a list sorted by probability in the objective lottery. To incentivize participants to report their true preferences in each row, they rewarded with their chosen option in a random row.

\$$x$ if E Occurs	\$$x$ with probability 0.1
\$$x$ if E Occurs	\$$x$ with probability 0.2
\$$x$ if E Occurs	\$$x$ with probability 0.3
\$$x$ if E Occurs	\$$x$ with probability 0.4
\$$x$ if E Occurs	\$$x$ with probability 0.5
\$$x$ if E Occurs	\$$x$ with probability 0.6
\$$x$ if E Occurs	\$$x$ with probability 0.7
\$$x$ if E Occurs	\$$x$ with probability 0.8
\$$x$ if E Occurs	\$$x$ with probability 0.9

TABLE I. A Probability Price List

Although binarized scoring rules are incentive compatible for a class of preferences that obey a form of reduction of compound lotteries known as **subjective-objective reduction** [S-O Reduction], price lists are incentive compatible under a weaker dominance assumption known as **statewise monotonicity** (Healy and Kagel, 2023).

This difference in the assumptions required for the two procedures is due to the different structure of the lotteries used for the incentives in each. This difference also makes it difficult to directly compare the strength of incentives. To account for this, we construct a binarized scoring rule that approximates the incentives of any probability price list. Using this, we demonstrate that probability price lists have incentives that

[1]See Schotter and Trevino (2014); Schlag et al. (2015); Charness et al. (2021) for surveys of belief elicitation in experimental economics.

in the limit[2] are exactly half as strong as those of the strongest binarized scoring rules (including the BQSR). Thus, in choosing between these procedures, there appears to be a trade-off between robustness and the strength of incentives.

We present a novel procedure that eliminates this trade-off, ternary price list. The ternary price list augments the probability price list by adding a third option: an act that pays if E *does not occur*.

$x if E Occurs	$x if E Does not Occur	$x with probability 0.5
$x if E Occurs	$x if E Does not Occur	$x with probability 0.6
$x if E Occurs	$x if E Does not Occur	$x with probability 0.7
$x if E Occurs	$x if E Does not Occur	$x with probability 0.8
$x if E Occurs	$x if E Does not Occur	$x with probability 0.9

TABLE II. A Ternary Price List

Like probability price lists, ternary price lists are incentive compatible under state-wise monotonicity and, in the limit, have incentives as strong as the best binarized scoring rules. In fact, in the limit, the incentives of a uniform[3] version of the ternary price list are identical to those of the BQSR. Thus, a ternary price list can be used to approximate the incentives of the BQSR with the more straightforward incentives of a price list.

II. PRELIMINARIES

We make use of the following notation. Let $\overline{x}$ be a "high" prize and $\underline{x}$ be a "low" prize. For any $p \in [0,1]$

An **objective lottery** is a lottery that pays the high prize with probability p and the low prize otherwise.

$$L^p = (\overline{x}, p; \underline{x}, (1-p))$$

A **pure act** is a "bet" that pays the high prize if some event E occurs and and the low prize otherwise.

$$f^E = [\overline{x}, E; \underline{x}, E^c].$$

Let $\succsim$ be preferences over compound acts/lotteries and let $\mathscr{A}$ be the set of all pure acts. Throughout, we assume preferences meet the following assumptions:

Axiom 1 (Preference)**.** Preferences $\succsim$ over $\mathscr{A}$ are complete and transitive.

[2]As the number of rows grows to infinity.

[3]Where each row is chosen for payment with equal probability.

Axiom 2 (Monotonicity over Pure Lotteries)**.** $p \geq q$ if and only if $L^p \succsim L^q$.

Axiom 3 (Statewise Monotonicity over Pure Acts)**.** If $E \supseteq F$ then $f^E \succsim f^F$, and if $E \supset F$ then $f^E \succ f^F$.

Axiom 4 (Additive Beliefs)**.** $f^E \sim L^p$ if and only if $f^{E^c} \sim L^{1-p}$.

These axioms ensure that for any event E, a participant has a unique belief p defined by the indifference equation $f^E \sim L^p$ and that if a participant's belief about E is p then they belief about the complement of E is $1-p$. This follows by the straightforward inclusion of additive beliefs to *Healy and Leo (2024) Proposition 1*.

Binarized Scoring Rules

In a **binarized scoring rule**[4], a participant who states that their belief about the probability of an event E is q is compensated with a lottery that pays $\overline{x}$ with probability $s_1(q)$ if E occurs and pays $\overline{x}$ with probability $s_0(q)$ if E does not occur. Under S-O reduction, a participant is indifferent between this compound act and the objective lottery $L^{V(q|p)}$ where $V(q|p)$ is the reduced probability of $\overline{x}$:

$$V(q|p) = ps_1(q) + (1-p)s_0(q)$$

A scoring rule is **proper** if $p = q$ is a maximizer of the above expression. It is **strictly proper** if $p = q$ is the unique maximizer. Experimental economists often use the **binarized quadratic scoring rule** where $s_1(q) = 1-(1-q)^2$ and $s_0(q) = t(1-q) = 1-q^2$. However, there are other strictly incentive compatible choices for $s_1()$ and $s_0()$, for instance, the spherical rule: $s_1(q) = \frac{1-q}{(q^2+(1-q)^2)^{\frac{1}{2}}}$ and $s_0(q) = s_1(1-q)$.

Define $G(p) = ps_1(p) + (1-p)s_0(p)$, which represents the expected payment function under truth-telling.

Proposition 1. (*Gneiting and Raftery (2007) Theorem 1*) A scoring rule is [strictly] proper if and only if $G()$ is [strictly] convex.

Probability Price Lists

A **probability price list**[5] for event E consists of n pairs. Each pair has an objective lottery and a pure act (L^{l_i}, f^E) where probabilities $l_1, l_2, ..., l_n$ such that $l_i < l_{i+1}$. Participants are asked to choose one option from each pair. These pairs are arranged in rows

[4]See Healy and Leo (2024) for a history and more complete introduction to the theory and use of scoring rules in experimental economics.

[5]See Healy and Leo (2024) for a history and more complete introduction to the theory and use of price lists for eliciting beliefs in experimental economics.

that increase in l_i. To incentivize participants to reveal their favorite from each pair, a random row is selected. Row $i \in 1,...,n$ is chosen with probability π_i. The participant is paid according to their choice from that row.

Under the additional preference assumption of statewise monotonicity (see Healy and Leo, 2024) a price list identifies a range $[l_i, l_{i+1}]$ that must contain a participant's true belief. This range is identified by their "switching point". Suppose a participant chooses the objective lottery L^{l_i} over the act f^E but f^E over the objective lottery $L^{l_{i+1}}$ then $L^{l_{i+1}} \succsim f^E \succsim L^{l_i}$ (with at least one of these strict) and so their belief p where $F^E \sim L^p$ must be between l_i and l_{i+1}.

By far the most common type of price list in experimental economics is a **uniform price list**. A uniform price list with n rows uses rows chosen with the same chance $\pi_i = \frac{1}{n}$ and probabilities that are equally spaced such that $l_i = \frac{i}{n+1}$ as in Table I.

Ternary Price Lists

A **ternary price list** extends a probability price list by adding a third option to each "row" that pays $\overline{x}$ if E *does not* occur. That is, each row contains the triplet is (L^{l_i}, f^E, f^{E^c}). Furthermore, in a ternary price list, the row probabilities are all above 0.5: $l_i \geq 0.5$.

Because a ternary list is still a procedure where each participant chooses one option from a set of menus, it is incentive compatible under the weaker statewise monotonicity, like the probability price list. However, compared to probability price lists, ternary lists require the assumption of additive beliefs (Axiom 4). In fact, it is precisely the assumption of additivity that makes the ternary list more efficient relative to a traditional probability price list.

When beliefs are additive, you can measure the probability of E by measuring the probability of E^c. For example, if you know that the probability of E^c is 0.75 then the probability of E is 0.25. In fact, the probability price list in Table I and the ternary price list in Table II categorize the beliefs of the participants in the same partition, but the ternary price list uses almost half as many rows.

III. Comparing Incentives

Approximating Incentives of Lists with Scoring Rules

Our goal is to compare the incentives of price lists, ternary price lists, and scoring rules. There are two main barriers to this comparison. First, the lotteries used for incentives in price/ternary price lists and scoring rules have a different structure and different incentive compatibility requirements. Second, scoring rules elicit beliefs precisely while price/ternary price lists elicit beliefs coarsely, categorizing a participant's belief into an interval. In this section, we introduce "reduced infinite probability price lists" and "reduced infinite ternary price lists" that approximate the incentives of probability price lists and ternary lists using binarized scoring rules. This allows for a direct comparison.

Consider the following modification to the probability price list. Instead of choosing a "paying" row, reduce the lottery for the participants and pay an objective lottery that provides the high prize $\overline{x}$ with the *weighted average of the probability of the high prize across the rows* and conditional on the state. This is a lottery that has the same form as the lottery used in a binarized scoring rule.

A participant who reduces all lotteries that mix subjective and objective risk to the simple lottery that pays the high prize with their subjective expected probability of receiving that prize across the states is indifferent between being paid the lottery from the reduced probability price list and the compound lottery that results from the analogous traditional probability price list. In this sense, this modification maintains the overall incentives of the price list, but with the reduced structure of a binarized scoring rule.

Our next hurdle is to make price lists and scoring rules elicit the same information. Theoretically, a scoring rule elicits a participant's precise belief p, while a price list elicits an interval $[p_i, p_{i+1}]$ that must contain a participant's belief. The endpoints of this interval are two consecutive row probabilities of the probability price list. As the number of rows increases, the gap between each pair of probabilities shrinks. Thus, the interval identified by the switching point that must contain the participant's true belief becomes more precise. In the limit, the probability is precisely identified.

These two modifications lead to our notion of an infinite reduced probability price list. Formally, the mechanism is an infinite set of pairs of lotteries L^l for $l \in [0,1]$ and a weighting distribution F. Have the participant submit their belief q. For each l, determine which of $f^E \succsim L^l$ a participant with belief q would prefer. Pay the participant with an objective lottery that provides the high prize $\overline{x}$ with the weighted (by some distribution F) average of the probability of the high prize across all l and conditional on the state.

An infinite reduced probability price list with weighting distribution $F()$ is a scoring rule with $s_1(q) = \int_0^q f(x)dx + \int_q^1 xf(x)dx$ and $s_0(q) = \int_q^1 xf(x)dx$. A participant whose preferences obey S-O reduction treats this as a lottery that pays $\overline{x}$ with the following probability:

$$V(q|p) = p\left(\int_0^q f(x)dx + \int_q^1 xf(x)dx\right) + (1-p)\left(\int_q^1 xf(x)dx\right) \tag{1}$$

Proposition 2. A reduced infinite probability price list with weighting distribution $F()$ is a proper scoring rule, strictly so if $F()$ is continuous.

We can repeat a similar process with a ternary price list to convert it into a **infinite reduced ternary price list** which is a binarized scoring rule that approximates the incentives of a ternary price list. Let $F()$ be the weighting distribution, a CDF with support $[0.5, 1]$, the result is a scoring rule with s_1 and s_0 as follows:

$$s_1(q) = \begin{cases} \int_{0.5}^{q} f(x)dx + \int_{q}^{1} x f(x)dx, & \text{if } q \geq 0.5 \\ \int_{0.5}^{1-q} f(x)dx + \int_{1-q}^{1} x f(x)dx, & \text{if } q < 0.5 \end{cases} \tag{2}$$

$$s_0(q) = \begin{cases} \int_{q}^{1} x f(x)dx, & \text{if } q \geq 0.5 \\ \int_{1-q}^{1} x f(x)dx, & \text{if } q < 0.5 \end{cases} \tag{3}$$

A participant whose preferences obey S-O reduction treats this as a lottery that pays $\overline{x}$ with the following probability:

$$V(q|p) = \begin{cases} p\left(\int_{0.5}^{q} f(x)dx + \int_{q}^{1} x f(x)\right) + (1-p)\left(\int_{q}^{1} x f(x)dx\right), & \text{if } q \geq 0.5 \\ p\left(\int_{0.5}^{1-q} f(x)dx + \int_{1-q}^{1} x f(x)dx\right) + (1-p)\left(\int_{1-q}^{1} x f(x)dx\right), & \text{if } q < 0.5 \end{cases} \tag{4}$$

Proposition 3. A reduced infinite ternary price list with weighting distribution $F()$ is a proper scoring rule, strictly so if $F()$ is continuous.

Strength of Incentives

One way to measure the incentives for truth-telling in a scoring rule is to measure how quickly $V(q|p)$ decreases as q changes from p. That is, by the second derivative of V with respect to q evaluated at $q = p$. In other words, the **strength of incentives for truth-telling** at p for a proper scoring rule is given by: $\left.\frac{\partial^2 V(q|p)}{(\partial q)^2}\right|_{q=p}$.

Proposition 4. (*Healy and Leo (2024) Lemma 1*) The strength of incentives for a proper scoring rule is given by the second derivative of G. That is, $\left.\frac{\partial^2 V(q|p)}{(\partial q)^2}\right|_{q=p} = G''(p)$.

Healy and Leo (2024) demonstrate that, for proper binarized scoring rules, the average incentives must be weakly less than 2. That is $\int_0^1 G''(p)dp \leq 2$. Calculating the average for the infinite reduced ternary price lists demonstrates that they attain the bound, while probability price lists offer average incentives only half-as-strong. We demonstrate this fact using an alternative approach in Appendix Section .

Procedure	**Strength of Incentives**
Any Binarized Scoring Rules	$\int_0^1 G''(p) \leq 2$ (Average)
Any Ternary Price Lists	$\int_0^1 G''(p) = 2$ (Average)
Any Probability Price Lists	$\int_0^1 G''(p) = 1$ (Average)
BQSR	$G''(p) = 2$ Everywhere
Uniform Ternary price list	$G''(p) = 2$ Everywhere
Uniform Probability price list	$G''(p) = 1$ Everywhere

TABLE III. Average Strength of Incentives for Binarized Scoring Rules and Infinite Reduced Price/ternary price lists.

Note as well that the particular instances of each class of mechanism in the table above: the BQSR, Uniform Infinite Reduced ternary price list, and Uniform infinite reduced probability price list all attain their class's average bound across all beliefs. That is they each maximizes the minimum strength of incentives within their relative class of mechanism. They each have **uniformly strongest incentives** in their class.

From a strength perspective, we also see that the BQSR and the uniform infinite reduced ternary price list have the same strength of incentives everywhere. This is no coincidence; they are, in fact, the same thing.

Proposition 5. The incentives of the uniform infinite ternary price list are identical to those of the BQSR.

IV. TWO-PART LISTS

The ternary list improves the incentives of a price list by cutting the number of rows in half (or in the case of an infinite list cutting the range in half) to double the probability/likelihood each row is randomly chosen for payment. Eliminating these rows means that the ternary list can only probabilities in the range 0.5 to 1. However, a participant has to believe that one of E and E^c is at least 50% likely. Thus, by including both options, the ternary list can both determine which of E and E^c the participant thinks is at least 50% likely and elicit the probability of that in one list.

Having three options in each row would not be necessary if we knew which of E and E^c the participant thought was the most likely. We could then use a standard probability price list (with two options in each row) starting at 0.5 to elicit the probability of whichever of E and E^c they believe is most likely.

Interestingly, we can elicit this information without providing any additional incentives or watering down the resulting probability elicitation. First, ask the participant which they prefer f^E or f^{E^c} (or which of E of E^c they think is more likely). Then elicit the probability of the relevant event with a probability price list using row probabilities each at least 0.5. We call this procedure the **two-part price list**. An example of a two-part price list is shown in Table IV.

The elicitation itself is incentive compatible under Axioms 1-4 and statewise monotonicity as before. The preliminary question is incentive compatible under the same conditions. Note that the best a participant can do if they choose their less preferred option is to then choose the objective lottery in each row of the subsequent price list. However, by telling the truth both in the preliminary question and in the subsequent price list, the outcome in each state of the resulting compound lottery is at least as good since the objective lottery remains available. If the act is chosen in any of the rows, then the outcome in that state is strictly better. Thus, truth-telling in both the preliminary question and subsequent probability elicitation statewise dominates any other set of choices.

x if E	x if E^c

Which do you prefer?

Choice	\$$x$ with probability 0.5
Choice	\$$x$ with probability 0.6
Choice	\$$x$ with probability 0.7
Choice	\$$x$ with probability 0.8
Choice	\$$x$ with probability 0.9

TABLE IV. A Two-Part Price List

V. CONCLUSION

The ternary price list and two-part price list provides an efficient tool for belief elicitation and extends the application of price list methods by matching the incentive strength of scoring rules without sacrificing robustness. These procedures can be used to implement (or approximate) the incentives of the BQSR in a simpler way, without any complex instructions or calculations and while improving the robustness of the incentives.

Ternary lists can provide a substantial amount of information about beliefs in a few simple choices. At the more extreme end of this, the following two-row ternary price list elicits beliefs into five intervals: $\{[0,0.20],[0.2,0.4],[0.4,0.6],[0.6,0.8],[0.8,1]\}$. In this sense, a two-row ternary list can act as a sort of objective Likert-scale for beliefs.

\$x if E	\$x if E'	\$x with probability 0.8
\$x if E	\$x if E'	\$x with probability 0.6

TABLE V. A Two-Question Ternary Price List that Partitions Beliefs into Five Intervals $\{[0,0.20],[0.2,0.4],[0.4,0.6],[0.6,0.8],[0.8,1]\}$.

In fact, if the first row is eliminated, the categorization is $\{[0,0.4],[0.4,0.6],[0.6,0.8]\}$ which may be a sufficient categorization of beliefs in some contexts and is elicited with just one question, and essentially universally incentive compatible, though by placing this elicitation in the context of a larger experiment, statewise monotonicity is likely still required. See Azrieli et al. (2018) for a more detailed discussion.

\$x if E	\$x if E'	\$x with probability 0.6

TABLE VI. A Two Question Ternary Price List that Partitions Beliefs into Three Intervals $\{[0,0.4],[0.4,0.6],[0.6,1]\}$.

By choosing the row probabilities carefully in these small ternary lists, it would be possible to elicit sufficient information for many research scenarios where absolute precision is not required. Smaller lists like this are simple and, since they use fewer rows, provide stronger incentives in each of the rows that remains.

Future research is needed to empirically test the performance of these procedures in various experimental settings.

ACKNOWLEDGMENTS

REFERENCES

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.

Charness, G., Gneezy, U., and Rasocha, V. (2021). Experimental methods: Eliciting beliefs. *Journal of Economic Behavior & Organization*, 189:234–256.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378.

Healy, P. J. and Kagel, J. (2023). Testing Elicitation Mechanisms Via Team Chat.

Healy, P. J. and Leo, G. (2024). Belief elicitation: A user's guide. working paper.

Schlag, K. H., Tremewan, J., and Van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3):457–490.

Schotter, A. and Trevino, I. (2014). Belief elicitation in the laboratory. *Annu. Rev. Econ.*, 6(1):103–128.

VI. APPENDIX

Proof of Proposition 2

Proof. Let $H()$ be the anti-derivative of F. Setting $q = p$ in (1) and simplifying the result:

$$G(p) = 1 - H(1) + H(p) \tag{5}$$

Since $G''(p) = f(p) \geq 0$, $G(p)$ is convex [strictly-so if $F()$ is continuous] and so the price list is a [strictly] proper scoring rule by Proposition 1. □

Proof of Proposition 3

Proof. Let $H()$ be the anti-derivative of F. Setting $q = p$ in (4) and simplifying the result:

$$G(p) = \begin{cases} 1 - H(1) + H(p) & \text{if } p \geq 0.5 \\ 1 - H(1) + H(1-p) & \text{if } p < 0.5 \end{cases} \tag{6}$$

Since $H(p)$ is the anti-derivative of an increasing function, $G(p)$ is increasing and convex for $p \in [0.5, 1]$. Since $G(p)$ is symmetric around 0.5 it is decreasing convex for $p \in [0, 0.5]$.

□

Proof of Proposition 5

Proof. Let $V_s(q|p)$ and $V_t(q|p)$ be the probability of receiving $\overline{x}$ given belief p and claimed belief q in the BQSR and reduced infinite ternary price list respectively.

$$V_s(q|p) = 2\left(pq - \frac{q^2}{2} + \frac{1}{2}\right) - p \tag{7}$$

We can construct V_t using (4) with $f(p) = 2$ and $F(p) = 2p - 1$. For $q \geq 0.5$:

$$\begin{aligned} V_t(q|p) &= \int_q^1 2\pi d\pi + p \int_{0.5}^q 2d\pi \\ &= 2\left(pq - \frac{q^2}{2} + \frac{1}{2}\right) - p = V_s(q|p) \end{aligned} \tag{8}$$

For $q \leq 0.5$:

(9)
$$V_t(q|p) = \int_{1-q}^{1} 2\pi d\pi + (1-p)\int_{0.5}^{1-q} 2d\pi$$
$$= 2\left(pq - \frac{q^2}{2} + \frac{1}{2}\right) - p = V_s(q|p)$$

□

Incentives of Uniform Price List and BQSR

Below, let $V_s(q|p)$ and $V_l(q|p)$ be the probability of receiving $\overline{x}$ given belief p and elicited belief q in the BQSR and reduced infinite uniform price list respectively.

(10)
$$V_s(q|p) = 2\left(pq - \frac{q^2}{2} + \frac{1}{2}\right) - p$$

Simplifying (1) with $F(p) = p$:

(11)
$$V_l(p,q) = pq - \frac{q^2}{2} + \frac{1}{2}$$

We introduce the *trivial scoring rule* that ignores the report and always pays the price when E such that $V_t(q|p) = p$. From (10) and (11), $V_l = \frac{1}{2}V_s + \frac{1}{2}V_t$. That is, the induced incentives of the uniform price list can be thought of as an equal mixture between the BQSR and the "Trivial" scoring rule that ignores the report completely.