Eliciting Second-Order Beliefs

Subir Bose
University of Leicester
sb345@leicester.ac.uk

Arup Daripa
Birkbeck, University of London
a.daripa@bbk.ac.uk

September 27, 2019

Abstract

We study elicitation of subjective beliefs of an agent facing ambiguity (model uncertainty): the agent has a non-singleton set of (first-order) priors on an event and a second-order distribution on these priors. Such a two-stage decomposition of uncertainty and non-reduction of subjective compound lotteries resulting from non-neutrality to the second-order distribution plays an important role in resolving the Ellsberg Paradox. However, a key unanswered question is whether we can actually observe and separate the sets of first and second order subjective probabilities. We answer this question and show that it is indeed possible to pin the two sets down uniquely and therefore separate them meaningfully. We introduce prize variations and ensure that the tangent plane at any point on the surface of the certainty-equivalent function is reported truthfully. Any basis for the this plane consist of derivatives in two different directions, and these combine first and second order beliefs in different ways. We show that this is enough to ensure reporting both sets of beliefs truthfully is uniquely optimal. The technique requires knowledge of the changes in certainty equivalent of acts as a result of variations in prizes, which we also elicit.

JEL CLASSIFICATION: D81, D82

KEYWORDS: Ambiguity, second-order beliefs, elicitation of second-order beliefs (support and distribution), Klibanoff-Marinacci-Mukerji (2005) representation.
Second-order probabilities and attitude towards such probabilities play an important role in explaining choice that appear anomalous under expected utility. When the data generating process is known, an agent faces only aleatory uncertainty, captured by a (subjective) probability model. However, the data generating process itself might be uncertain, in which case the agent faces a second layer of uncertainty over possible probability models, giving rise to “model uncertainty” or “ambiguity”.\(^1\)

Model uncertainty may naturally give rise to the agent having subjective beliefs –referred to as second-order beliefs–over possible probability models. In some cases however – for example, when the agent’s preferences satisfy the Savage axioms – model uncertainty and beliefs over model uncertainty do not produce a situation qualitatively different from one where the agent is certain of the data generating process. In the latter case the agent essentially has a single first-order belief through reduction of subjective compound lotteries. However, for an important class of preferences, model uncertainty makes for a qualitatively different response, reflected in non-reduction of subjective compound lotteries. Introduced by Segal (1987,1990), such non-reduction is critical for explaining choice behavior categorised as the Ellsberg paradox (Ellsberg, 1961).

Consider an act \(a_E b\) which pays a monetary amount \(a\) if event \(E\) happens, and an amount \(b\) otherwise. Suppose the agent has a finite set of beliefs \(\{\pi_1, \ldots, \pi_m\}\) about \(E\), and suppose \(\mu_i\) is the second order belief on \(\pi_i\), \(i \in \{1, \ldots, m\}\). Here \(\pi_i \in [0,1]\), \(\mu_i \in (0,1]\) and \(\sum_i \mu_i = 1\). In the representation axiomatized by Klibanoff, Marinacci, and Mukerji (2005) (KMM), an agent evaluates this act according to

\[
\sum_{i=1}^{m} \phi(\pi_i u(a) + (1 - \pi_i)u(b)) \mu_i
\]

where \(\phi\) is an increasing function and \(u\) is the vN-M utility function. The function \(\phi\) captures attitude towards ambiguity: the agent’s preference exhibits ambiguity aversion, neutrality or fondness according as \(\phi\) is strictly concave, linear or strictly convex. Note that the first-order beliefs \(\pi_i\) are the belief-states\(^2\) with the second-order belief as the

---

\(^1\)Model uncertainty is a key aspect of empirical methods as well. As Marinacci (2015) notes, classical statistics posits a set of probability models over states to comprise an agent’s ex ante information (Wald (1950), Fisher (1957), Neyman (1957)). The one that best describes the variability in the states is then identified as the generating model.

\(^2\)In the literature, the belief-states are sometimes called second-order states, with the payoff-relevant
(unique) prior over these states.\textsuperscript{3}

A key question, then, is whether we can actually observe the two separate sets of subjective probabilities: the first-order beliefs and the second-order distribution on those. As we discuss below, there is no satisfactory answer to this question in the literature. In this paper, we assume the KMM representation and show that the beliefs $\pi_i$ and second-order probabilities $\mu_i$ can indeed be separated and uniquely pinned down. The fact that this separation is possible and each set can be uniquely identified is the principal theoretical contribution of the paper.

The problem of eliciting\textsuperscript{4} (first and second order) beliefs of an ambiguity-sensitive agent throws up challenges that are qualitatively different from those encountered when dealing with beliefs on observable events. Since the belief-states are fundamentally unobservable,\textsuperscript{5} standard elicitation techniques fail to work. Indeed, as discussed in the literature review in section 7, all available mechanisms in the literature depend on a reward based on an observable event and/or require the agent to be an SEU maximizer.

Therefore, elicitation of unobservable belief-states and second order probabilities on these states presents a significant challenge. In response, we introduce new techniques and conceptual innovations. Our mechanism introduces prize variations and ensures that for any belief report, the agent is given the incentive to ensure that the resultant change in certainty equivalent (CE) in any direction is reported truthfully. Essentially, then, the agent has the incentive to report truthfully the tangent plane at any point $(x_k, y_k)$ on the surface of the CE function. Now, the tangent plane at any point is a two-dimensional plane - so a basis for such a plane consists of two linearly independent vectors. The agent’s belief report must then satisfy the property that the directional derivatives corresponding to two linearly independent directions on the tangent plane coincide with the truth. In this case it would be possible for the report to ensure that directional derivatives coincide

\textsuperscript{3}Representations based on different axiomatisations by Nau (2006), Ergin and Gul (2009), Seo (2009) have a similar structure, with attitude towards dispersion of second-order probabilities capturing ambiguity attitude.

\textsuperscript{4}By “elicit” we mean “to make truthful reporting of beliefs the unique best response of the agent.”

\textsuperscript{5}Suppose an agent believes $E$ will happen with probability either 1/3 or 2/3 and that these two belief-states are equally likely. While a single observation determines whether $E$ occurs or not and hence resolves the aleatory uncertainty, it is not possible to ascertain whether $E$ happened with probability 1/3 or 2/3 if it does happen, implying that the model uncertainty cannot be resolved from such an observation.
with true values in all directions on the tangent plane. But the two directions combine \( \pi \) values and \( \mu \) values in different ways. Using this, we show that the required property can be satisfied if only if the report is truthful.

The literature on belief elicitation on observable states compares levels of utility from acts and lotteries, and such beliefs can be elicited without requiring knowledge of the utility function of the agent. Here, with unobservable states and ambiguity-sensitive agents, we need to consider how the CE changes as we vary the prizes. We elicit these changes in certainty equivalent as part of our mechanism.

However, the requirement to elicit both beliefs and the changes in CE complicates the elicitation mechanism as it needs to combine different elements for each task. We feel this makes the presentation of our key ideas for belief-elicitation somewhat difficult to follow. To resolve this problem, we structure our analysis as follows.

In section 3, we clarify our ideas for belief elicitation using examples. To focus clearly on belief elicitation, we assume that the function \( \phi \) is known to the mechanism designer - who can therefore calculate the changes in CE in response to prize variations. The ideas explained here comprise our principal contribution. Next, in sections 4 to 6, we remove the assumption that \( \phi \) is known and present our general mechanism - one that elicits both beliefs and CE changes.

### 2 Preliminaries

We are interested in eliciting the agent’s beliefs regarding an event \( E \).

Let a finite non-singleton set of probabilities represent the agent’s (first-order) beliefs about \( E \). The agent’s beliefs involve a second-order belief which is a probability distribution on the first-order beliefs. Any probability in the support of the second-order belief is called a belief-state.

Let the agent’s belief be denoted by \( B \equiv ((\pi_1, \mu_1), \ldots, (\pi_m, \mu_m)) \) where \((\pi_1, \ldots, \pi_m)\) are the belief-states, \( \pi_i \in [0, 1] \) and \( \pi_i \neq \pi_j \) for all \( i, j \in \{1, \ldots, m\} \). Further, \((\mu_1, \ldots, \mu_m)\) are the second-order beliefs, where \( \mu_i > 0 \) is the belief associated with the state \( \pi_i, i \in \{1, \ldots, m\} \) and \( \sum_i \mu_i = 1 \).

The task is to elicit \( B \), i.e. elicit both the belief-states and the second-order belief attached
to each state. We use direct revelation mechanisms for this task and “elicit X” means the agent’s unique best response is to truthfully report X when facing the mechanism.

As noted in the introduction, we consider eliciting beliefs of an agent with preference represented by the KMM model of smooth ambiguity.

The agent’s risk preference (as represented by the function $u(\cdot)$) plays no role in our analysis and our mechanisms for belief elicitation would work irrespective of the risk attitude of the agent. To simplify the analysis, we work with prizes in the utility-space rather than the money-space. In other words, expected utility from acts with money prizes are written simply as expected payoffs over corresponding vN-M utilities. So, for example, if the money prize is $\hat{x}$, we write the prize as $x = u(\hat{x})$. We assume the agent strictly prefers more money; in other words, if $\hat{x} > \hat{y}$, then $x > y$. Let $x \leq y$ denotes the (subjective) act that pays prize $x$ if event $E$ happens and prize $y$ otherwise. An agent with KMM preference evaluates this according to

$$
\phi \circ B = \sum_{i=1}^{m} \phi(\pi_i x + (1 - \pi_i) y) \mu_i
$$

We implicitly assume that the vN-M utility function $u$ is known. This makes for the simplest theoretical exposition of the idea of separating first and second order beliefs. However, for practical purposes, the literature offers several methods of approximating the function $u$.

In common with most of the literature, we assume that all departures from expected utility is due to ambiguity sensitivity. In particular, this means that for all types of ambiguity-sensitive preferences, the agent’s payoffs from objective compound lotteries is exactly the same as an expected utility maximizer’s with same risk preferences.

---

6See, for example, Wakker and Deneffe (1996). One method of approximating $u$ is as follows. Fix two outcomes $M > m$ and set $u(m) = 0, u(M) = 1$. Consider lotteries that award $M$ with probability $p$ and $m$ with $(1 - p)$. To elicit the CE of such lotteries across several values of $p$, set up the associated lotteries and ask the agent to report CE values. Pick one lottery randomly. Let $r$ denote the reported CE. The designer draws a value $q$ randomly (say, using the uniform distribution) from $[m, M]$ and awards the agent the lottery if $r > q$ and the prize $q$ otherwise. It is straightforward to see that the agent’s best response is to report CE truthfully. Once CE values are elicited, for each $p$ we have $u(\text{CE}(p)) = pu(M) + (1 - p)u(m) = p$. 

4
3 Belief elicitation assuming $\phi$ known

As noted before, eliciting belief states and second-order beliefs requires knowledge of how CE changes in response to prize variations. This requires us to elicit CE changes along with beliefs. For ease of exposition, we first present our belief elicitation procedure assuming that the function $\phi$ is known (implying that CE changes are known). This allows us to present our belief-elicitation ideas clearly. Later, in sections 4 to 6, we remove this assumption and present the general mechanism that elicits both beliefs as well as the CE changes.

The agent has beliefs $B \equiv ((\pi_1, \mu_1), \ldots, (\pi_m, \mu_m))$ and reports beliefs $\hat{B} \equiv ((\hat{\pi}_1, \hat{\mu}_1), \ldots, (\hat{\pi}_n, \hat{\mu}_n))$. Note that the agent reports $n$ belief states, while the true number is $m$. Of course, reported number $n$ can be different from $m$, and belief states as well as second-order beliefs can be misreported even if $n = m$.

3.1 Eliciting derivatives - an example

As noted previously, a crucial part of our belief-elicitation mechanism is that agents have incentive to ensure that report-based calculations of changes to CE from prize variations coincide with the truth. To build intuition about how the belief-elicitation works, let us first explain this idea of “eliciting derivatives" using a simple example.

Consider a prize vector $(x, y)$ where $x > y$ and the act $x_Ey$. Consider the following prize variation: increase $x$ and reduce $y$ ($\frac{dy}{dx} < 0$). If the absolute value $|\frac{dy}{dx}|$ is high (low) enough, obviously the agent would be worse off (better off) from this variation. Let $D_y(0)$ be the value of $\frac{dy}{dx}$ such that the agent’s CE does not change from the variation.\footnote{Formally, this can be derived as follows. The certainty equivalent of the act $x_Ey$ is CE = $\phi^{-1} (\sum_{i=1}^{m} \phi(\pi_i x + (1 - \pi_i)y)\mu_i)$ and the change in CE from the variation being zero implies

$$
\frac{1}{\phi'(\text{CE})} \sum_{i=1}^{m} \phi'(\pi_i x + (1 - \pi_i)y)(\pi_i + (1 - \pi_i)\frac{dy}{dx})\mu_i = 0
$$

$D_y(0)$ is then the solution for $\frac{dy}{dx}$ from above.}
Let $\hat{D}_y(0)$ be the value calculated from the report $\hat{B}$. Consider the following scheme.

- Ask the agent to announce $B$. The agent announces $\hat{B}$. Calculate $\hat{D}_y(0)$.
- Pick a $d$ randomly (e.g. using a uniform distribution) from $(-\bar{d}, 0)$ where $\bar{d}$ is a large number. If $|d| \leq |\hat{D}_y(0)|$, offer the variation. Otherwise offer no variation (i.e. simply offer the original act).

Suppose that the agent’s reports are such that $|\hat{D}_y(0)| < |D_y(0)|$. If $|d| > |D_y(0)|$ no variation is offered under a truthful report $B$ or a non-truthful report $\hat{B}$. If $|d| \leq |\hat{D}_y(0)|$, the variation is offered under both reports. But if $|\hat{D}_y(0)| < d \leq |D_y(0)|$, no variation is received even though the variation would improve the agent’s payoff. Since the probability that $d$ is in the last region is strictly positive, a report of $\hat{B}$ is strictly dominated by a report of $B$. It can be easily checked that the same dominance holds also when $\hat{B}$ is such that $|\hat{D}_y(0)| > |D_y(0)|$. Thus the dominant strategy of the agent involves reporting $\hat{B}$ such that $\hat{D}_y(0) = D_y(0)$.

Thus by introducing report-dependent variations, we can create incentives for the agent to report beliefs in a way so that the calculated derivative value coincides with the true value. We exploit this idea below to set up a mechanism to elicit beliefs.

### 3.2 A mechanism with prize variations

We use the following two acts and prize variations on these acts. First, consider a prize vector $(x, y)$ where $x > y > 0$ and the act $x_E y$. Let $\text{Var}(x, y)$ denote a prize variation that raises $x$ and reduces $y$.

Next, consider a constant act $z_E z$, $z > 0$. Let $\text{Var}(z)$ denote the following prize variation: raise $z$ at the same rate as $x$ in the other variation, i.e. $dz = dx$.

It is worth noting at the outset that in all mechanisms specified, any randomization as part of the mechanism is done before (nature’s) randomization resolving the uncertainty regarding the realization of event $E$. 
The mechanism

The agent is asked to report $B$. The agent reports $\hat{B} \equiv \{ (\hat{\pi}_1, \hat{\mu}_1), \ldots, (\hat{\pi}_n, \hat{\mu}_n) \}$. Note that the cardinality of the set of reported belief-states is $n$.

After receiving the report the mechanism designer follows the procedure below.

1. Announce $n$ prize vectors $(x_k, y_k), k \in \{1, \ldots, n\}$ and randomly choose one. Let $(x, y)$ denote the chosen vector.

2. Choose a number $p$ randomly from a sub-interval of the unit interval and award the act $x E y$ with probability $p$ and the constant act $z E z$ with probability $(1 - p)$.

3. Using reports, calculate $\frac{dy}{dx}$ which makes the agent indifferent between receiving $\text{Var}(x, y)$ and receiving $\text{Var}(z)$. Let $\hat{D}_y(p)$ denote this value.

4. Next, choose a number $d$ randomly from $[-d, 0]$.
   - If (a) $x E y$ is picked in step 2 and (b) $|d| \leq |\hat{D}_y(p)|$, award the variation $\text{Var}(x, y)$ at $\frac{dy}{dx} = d$.
   - If (a) $z E z$ is picked in step 2 and (b) $|d| > |\hat{D}_y(p)|$, award the variation $\text{Var}(z)$.
   - In any other case, award no variation - i.e. simply award the original act picked in step 2.

We showed before in section 3.1 that the derivative calculated from the agent’s report must optimally coincide with the true value of the derivative. The same idea, applied to the mechanism above, shows that if the reported $\hat{B}$ is such that the value $\hat{D}_y(p)$ calculated from the report is not equal to the true value $D_y(p)$, the report is suboptimal.

Let us now calculate the (true) value $D_y(p)$ which makes the agent indifferent between receiving $\text{Var}(x, y)$ and receiving $\text{Var}(z)$. Note that in the mechanism, the objective uncertainty (about which act is chosen) is resolved before the subjective uncertainty about

---

8The precise sub-interval is specified in the formal presentation in section 4. The precise value is unimportant for understanding how the mechanism works.
the event $E$. Therefore, for any choice of $p$, the agent calculates CE as

$$p \phi^{-1} \left( \sum_{i=1}^{m} \phi(\pi_i x + (1 - \pi_i) y) \mu_i \right) + (1 - p) z$$

In the rest of the analysis, we calculate changes to CE from prize variations. For economy of algebra, we define

$$a_i \equiv \frac{\phi'(\pi_i x + (1 - \pi_i) y)}{\phi'(\text{CE}|B)}$$

When calculated from reported beliefs, we denote the term by $\hat{a}_i$.

Now, CE change from variation $\text{Var}(x, y)$ is

$$p \sum_{i=1}^{m} a_i \left( \pi_i + (1 - \pi_i) \frac{dy}{dx} \right) \mu_i dx$$

and that from $\text{Var}(z)$ is $(1 - p) dz$. Using $dz = dx$, it follows that $D_y(p)$ is given implicitly by

$$\sum_{i=1}^{m} a_i \left( \pi_i + (1 - \pi_i) D_y(p) \right) \mu_i = (1 - p) / p. \quad (3.1)$$

Since, the value of $D_y(p)$ calculated from reports must be the same, we have

$$\sum_{i=1}^{n} \hat{a}_i \left( \hat{\pi}_i + (1 - \hat{\pi}_i) D_y(p) \right) \hat{\mu}_i = (1 - p) / p. \quad (3.2)$$

Since the right hand sides are equal, we can equate the left hand sides and rearrange terms to obtain the following.

$$\sum_{i=1}^{n} \hat{a}_i \hat{\pi}_i \hat{\mu}_i - \sum_{i=1}^{m} a_i \pi_i \mu_i = D_y(p) \left( \sum_{i=1}^{n} \hat{a}_i (1 - \hat{\pi}_i) \hat{\mu}_i - \sum_{i=1}^{m} a_i (1 - \pi_i) \mu_i \right)$$

Note that the only term that involves $p$ is $D_y(p)$. Since $D_y(p)$ changes with $p$, the above equality can only hold if the coefficient of $D_y(p)$ is zero, which also implies that the left hand side is zero. Simplifying from these two equations, we get

$$\sum_{i=1}^{n} \hat{a}_i \hat{\pi}_i \hat{\mu}_i = \sum_{i=1}^{m} a_i \pi_i \mu_i$$

$$\sum_{i=1}^{n} \hat{a}_i \hat{\mu}_i = \sum_{i=1}^{m} a_i \mu_i \quad (3.3)$$

Note that the first is the condition that the directional derivative in the direction $(1,0)$ coincide with the truth, and the second requires the directional derivative in the direction $(1,1)$ coincide with the truth.$^9$

---

$^9$CE change from variation $\text{Var}(x, y)$ is $p \sum_{i=1}^{m} a_i \left( \pi_i + (1 - \pi_i) \frac{dy}{dx} \right) \mu_i dx$. In the direction $(1,0)$, only $x$ changes, so that $dy/dx = 0$. In the direction $(1,1)$, $dy/dx = 1$. Using these, one can get the values of the derivatives of CE. The term $p$ appears on both sides and cancels out.
Recall that the mechanism announces \( n \) prize vectors \( x_k, y_k, k \in \{1, \ldots, n\} \), and randomly chooses one: \( (x_k, y_k) \). The above equations must be satisfied for any choice of \( (x_k, y_k) \), so we get two sets of conditions, each set comprising \( n \) conditions.

To accommodate multiple prize vectors, let us change the notation slightly. Let

\[
\hat{a}_{ki} \equiv \frac{\phi'(\pi_i x_k + (1 - \pi_i) y_k)}{\phi'(CE_k|\hat{B})}
\]

Similarly define \( \hat{a}_{ki} \) using reported values \( \hat{\pi}_i \) and \( CE|\hat{B} \). Since \( \phi \) is known, the designer can calculate values \( \hat{a}_{ki} \). Let

\[
\hat{A} = \begin{pmatrix}
\hat{a}_{11} & \ldots & \hat{a}_{12} \\
\vdots & & \vdots \\
\hat{a}_{n1} & \ldots & \hat{a}_{nn}
\end{pmatrix}
\]

The designer must choose the \( n \) prize vectors \( x_k, y_k, k \in \{1, \ldots, n\} \) so that the above matrix has full rank. As we will show later (2 in section 4), this is indeed in general possible so long as \( \phi \) has derivatives of orders high enough.

Let \( A \) be the \( n \times m \) matrix with typical element \( a_{ki}, k \in \{1, \ldots, n\} \) and \( i \in \{1, \ldots, m\} \). Next, let \( P \) be an \( m \times m \) diagonal matrix of belief-states:

\[
P \equiv \begin{pmatrix}
\pi_1 & \ldots & 0 \\
\vdots & & \vdots \\
0 & \ldots & \pi_m
\end{pmatrix}
\]

Let \( \hat{P} \) be the similar \( n \times n \) diagonal matrix of reported belief-states. Next, let \( M \) denote the \( m \times 1 \) vector of second-order beliefs

\[
M' = (\mu_1, \ldots, \mu_m)
\]

and similarly define \( \hat{M} \) as the \( n \times 1 \) vector of reported second-order beliefs.

The conditions given by (3.3) can then be written as

\[
\hat{A} \hat{P} \hat{M} = A P M.
\]

\[
\hat{A} \hat{M} = A M.
\]

Solving for \( \hat{M} \) from each equation and equating, \( \hat{A}^{-1} AM = \hat{P}^{-1} \hat{A}^{-1} A PM \) which implies that

\[
\hat{A}^{-1} A = \hat{P}^{-1} \hat{A}^{-1} A P.
\]
We now show that it is not possible to satisfy equation (3.4) unless \( \hat{M} = M \) and \( \hat{P} = P \), implying that truthtelling is uniquely optimal.

To see this, we work through an example. General proofs are supplied in later sections.

An example

Suppose the agent reports two belief-states \( \hat{\pi}_1 \) and \( \hat{\pi}_2 \). We now consider three cases: the true set of belief states consists, first, a unique belief; second, two belief-states; and third, three belief states. (The argument in the first case generalises to general \( n > m \), the second case \( n = m \) and the third, \( n < m \).)

Case 1: \( n = 2, m = 1 \) Suppose the agent has a unique prior \( \pi_1 \), but reports 2 priors \( \hat{\pi}_1 \) with weight \( \hat{\mu}_1 \) and \( \hat{\pi}_2 \) with weight \( \hat{\mu}_2 \), where \( \hat{\mu}_i > 0 \) and \( \sum_i \hat{\mu}_i = 1 \).

Let \( z_{ij} \) denote a typical element of the matrix \( \hat{A}^{-1}A \). Since the matrix is \( n \times m \), we have \( i \in \{1, \ldots, n\} \) and \( j \in \{1, \ldots, m\} \).

Applied to this case, equation (3.4) implies

\[
\begin{pmatrix}
  z_{11} \\
  z_{21}
\end{pmatrix}
= \begin{pmatrix}
  \frac{\pi_1}{\hat{\pi}_1}z_{11} \\
  \frac{\pi_1}{\hat{\pi}_2}z_{21}
\end{pmatrix}
\] (3.5)

Since \( \pi_1 \) cannot be equal to both \( \hat{\pi}_1 \) and \( \hat{\pi}_2 \), the only way this can hold is if, first, one of the reported belief-states is equal to \( \pi_1 \). Suppose \( \hat{\pi}_1 = \pi_1 \). Then we need \( z_{21} = 0 \). But recall that \( \hat{M} = \hat{A}^{-1}A \). Here \( M \) is simply 1, so \( \begin{pmatrix}
  \hat{\mu}_1 \\
  \hat{\mu}_2
\end{pmatrix} = \begin{pmatrix}
  z_{11} \\
  z_{21}
\end{pmatrix} \). If \( z_{21} = 0, \hat{\mu}_2 = 0 \), but this cannot be the case. Therefore equation (3.5) cannot hold.\(^{10}\)

This rules out reporting more belief states than the true number.

Case 2: \( n = 2, m = 2 \)

Next, suppose the agent has two belief states \( \pi_1 \) and \( \pi_2 \) and reports two belief states \( \hat{\pi}_1 \) and \( \hat{\pi}_2 \) (with positive weights \( \hat{\mu}_1 \) and \( \hat{\mu}_2 \), respectively, which add up to 1).

\(^{10}\)In the general case of \( m \) values and \( n > m \) reports, exactly the same problem emerges. All \( n \) rows of \( \hat{A}^{-1}A \) must contain positive elements. This then implies that for equation (3.4) to hold, some \( \pi \) values would need to be equal to more than one \( \hat{\pi} \) value, which is impossible.
In this case, equation (3.4) implies

\[
\begin{pmatrix}
z_{11} & z_{12} \\
z_{21} & z_{22}
\end{pmatrix} =
\begin{pmatrix}
\frac{\pi_1}{\hat{\pi}_1}z_{11} & \frac{\pi_2}{\hat{\pi}_1}z_{12} \\
\frac{\pi_1}{\hat{\pi}_2}z_{21} & \frac{\pi_2}{\hat{\pi}_2}z_{22}
\end{pmatrix}
\]

First, note that \(\hat{\pi}_i\) must be equal to true \(\pi_j\). Otherwise, it is impossible to have \(z_{ij} = \frac{\pi_j}{\hat{\pi}_i}\) for any choice of \(i, j\) where \(i \in \{1, 2\}\) and \(j \in \{1, 2\}\). Therefore we have \(\hat{\pi}_1 = \pi_1\) and \(\hat{\pi}_2 = \pi_2\).

Second, we further require that \(z_{12} = z_{21} = 0\). If this is not true, the proof is already done, since in that case equation (3.4) can only hold for truthful reporting. Suppose then that \(z_{12} = z_{21} = 0\). This implies the following.

**Lemma 1.** Suppose \(z_{12} = z_{21} = 0\). This implies that \(z_{11} = z_{22} = 1\), so that \(\hat{A}^{-1}A\) is the identity matrix.

The formal proof is in the appendix, but the intuition is simple. \(z_{12} = z_{21} = 0\) is only possible if \(\hat{A}\) and \(A\) are the same matrices, so that the off-diagonal elements of \(\hat{A}^{-1}A\) are 0. But then, of course, \(\hat{A}^{-1}A\) is the identity matrix.

Now, recall that \(\hat{M} = \hat{A}^{-1}AM\). Since \(\hat{A}^{-1}A\) is the identity matrix, it follows that \(\hat{M} = M\), so that second order beliefs are also truthfully reported.

**Case 3: \(n = 2, m = 3\)**

Finally, suppose two belief states \(\hat{\pi}_1\) and \(\hat{\pi}_2\) (with positive weights \(\hat{\mu}_1\) and \(\hat{\mu}_2\), respectively, which add up to 1), while the agent has 3 belief states \(\pi_i, i \in \{1, \ldots, 3\}\) with second-order beliefs \(\mu_i > 0, \sum_i \mu_i = 1\). In this case, equation (3.4) implies

\[
\begin{pmatrix}
z_{11} & z_{12} & z_{13} \\
z_{21} & z_{22} & z_{23}
\end{pmatrix} =
\begin{pmatrix}
\frac{\pi_1}{\hat{\pi}_1}z_{11} & \frac{\pi_2}{\hat{\pi}_1}z_{12} & \frac{\pi_3}{\hat{\pi}_1}z_{13} \\
\frac{\pi_1}{\hat{\pi}_2}z_{21} & \frac{\pi_2}{\hat{\pi}_2}z_{22} & \frac{\pi_3}{\hat{\pi}_2}z_{23}
\end{pmatrix}
\]

First, note that, as above, \(\hat{\pi}_i\) must be equal to some \(\pi_j\). Suppose, in this case,

\[\hat{\pi}_1 = \pi_2\quad\text{and}\quad\hat{\pi}_2 = \pi_3.\]
We further need all terms $z_{ij}$ other than $z_{12}$ and $z_{23}$ to be zero. If any of these terms are not zero, the proof is already done, since the only way to satisfy equation (3.4) is by reporting truthfully. So suppose that all these terms are indeed zero. As in the step above, this implies $\phi'(CE_i|\hat{B}) = \phi'(CE_i|B)$ for $i \in \{1, 2\}$. Therefore the matrix $\hat{A}$ can be rewritten as

$$
\hat{A} = \begin{pmatrix}
a_{12} \\
\phi'(CE_1|B) a_{13} \\
\phi'(CE_2|B) \\
\phi'(CE_2|B) a_{23}
\end{pmatrix}
$$

In other words, $\hat{A}$ is formed by the second and third columns of $A$. It follows directly that

$$
\hat{A}^{-1} A = \hat{P}^{-1} \hat{A}^{-1} AP = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
$$

But $\hat{M} = \hat{A}^{-1} AF$. Therefore,

$$
\begin{pmatrix}
\hat{\mu}_1 \\
\hat{\mu}_2
\end{pmatrix} = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
\mu_1 \\
\mu_2 \\
\mu_3
\end{pmatrix}
$$

But that implies $\hat{\mu}_1 = \mu_2$ and $\hat{\mu}_2 = \mu_3$. In other words, if the agent only reports $\pi_2$ and $\pi_3$, they must also only report second order $\mu_2$ and $\mu_3$ in order to match derivative values.

Of course, this reporting does not work, since report requires $\mu$ values adding to 1, whereas here $\sum \hat{\mu}_i = \mu_2 + \mu_3 < 1$.

From cases 1 and 3, we know that it is not possible to have $n \neq m$. Thus the only possibility is case 2, in which $n = m$, and as we showed, in that case beliefs (belief states and second-order beliefs) must be truthfully reported. This shows that under our mechanism with prize variations, truth-telling is uniquely optimal.
4 Belief Elicitation: The General Case

In the previous section we showed how beliefs can be elicited by assuming that the function $\phi$ is known and then using an example (which assumed $n$, the number of reported belief states is 2). Here we remove the assumption that $\phi$ is known, remove any restriction on $n$, and present a mechanism to elicit both beliefs and CE changes.

Further, in the previous section we also used a result that the matrix $\hat{A}$ of CE changes can be constructed to be of full rank without proof. Here we prove such a result formally (Theorem 2).

For ease of exposition, we break up the argument as follows. In this section, we introduce some preliminary concepts and a mechanism with prize variations. We then assume that CE changes are known and show that a matrix of reported CE changes (the matrix $\hat{A}$) can be constructed to be of full rank. Using this, and the mechanism with prize variations, we show that beliefs can be elicited. The arguments here are a more formal and general version of those in the last section. Next, in section 5, we construct a mechanism to elicit the CE changes. Finally, in section 6, we use the mechanisms presented in this section and section 5 to construct a grand mechanism that elicits beliefs and CE changes simultaneously.

4.1 Preliminaries

We first introduce a few elements that are used in the construction of the mechanism. The acts and prize variations are as in section 3.2.

**Ambiguous Act:** For prizes $x > y$, consider a lottery that pays $x$ with probability $p$ (and hence $y$ with probability $1 - p$) and it is announced that $p$ is chosen from a set $\mathcal{P} \equiv \{p_1, \ldots, p_n\}$. No further information as to how $p$ is chosen from this set is provided. The actual, physical construction of the ambiguous act can be done following techniques well-known in the literature and used in laboratory experiments. For example, a two coloured urn is constructed, consisting of red and white balls of unknown proportion and the act pays $x$ if the colour of the ball drawn from the urn is red and pays $y$ if the colour is white. It is announced that the proportion of red and white balls is

\[\text{The actual, physical construction of the ambiguous act can be done following techniques well-known in the literature and used in laboratory experiments. For example, a two coloured urn is constructed, consisting of red and white balls of unknown proportion and the act pays } x \text{ if the colour of the ball drawn from the urn is red and pays } y \text{ if the colour is white. It is announced that the proportion of red and white balls is} \]

13
state $s_i$ associated with probability $p_i$ of receiving payoff $x$. Let $\{h_1, \ldots, h_n\}$ denote the agent’s second-order belief over $\mathcal{P}$, with $h_i$ being the probability weight attached to ambiguity state $s_i$.

**Second-Order Act (SOA):** For some prizes $w_1, \ldots, w_n$, a second-order act pays $w_i$ if $p$ is chosen to be $p_i$.

**Change in certainty equivalent from SOA prize variation** Next, let us introduce the following change in certainty equivalent (CE) of an SOA as specified above. The CE is given by $CE = \phi^{-1} (\sum_{i=1}^n \phi(w_i) h_i)$. Therefore the change in CE from a change in $w_i$ is given by

$$t_i = \frac{\phi'(w_i)}{\phi'(CE)} h_i$$

(4.1)

In what follows, all variables with a “hat” are either values reported directly or values derived based on reports.

Recall from section 3 that the derivative $\hat{D}_y(p)$ is calculated from reports and this plays an important part in the mechanism. A fully formal derivation is as follows.

**Construction of derivative $\hat{D}_y(p)$ from reports** $\hat{D}_y(p)$ is given by the solution to

$$p \sum_{i=1}^n \frac{\hat{t}_i}{\hat{h}_i} \left( \hat{\pi}_i + (1 - \hat{\pi}_i) \frac{dy}{dx} \right) \hat{\mu}_i = (1 - p)$$

(4.2)

Let $\underline{p}$ be the value of $p$ from above when $dy/dx = 0$. In other words,

$$\underline{p} \equiv \frac{1}{\hat{T}_1 + 1}$$

(4.3)

where $\hat{T}_1 \equiv \sum_{i=1}^n \frac{\hat{t}_i}{\hat{h}_i} \hat{\pi}_i \hat{\mu}_i$. Note that $\underline{p} \in (0, 1)$, so that the interval $(\underline{p}, 1]$ is non-empty. It is straightforward to verify that so long as $\underline{p} < p \leq 1$, there is a unique solution $\hat{D}_y(p)$.

We are now ready to present the mechanism. It is worth noting at the outset that in all mechanisms specified, any randomization is done before the event $E$ is realized.
4.2 A mechanism with prize variations

I. Reporting stages

The agent is asked to report $B$. The agent reports $\hat{B} \equiv \{ (\hat{\pi}_1, \hat{\mu}_1), \ldots, (\hat{\pi}_n, \hat{\mu}_n) \}$. Note that the cardinality of the set of reported belief-states is $n$. Let $s_1, s_2, \ldots, s_n$ denote the announced belief-states.

After receiving the report the mechanism designer follows the procedure below.

1. Announce $n$ prize vectors $x_i, y_i, i \in \{1, \ldots, n\}$.

2. Construct $n$ second order acts $k = 1, \ldots, n$ as follows. Act $k$ has the prize profile $\omega_i^k, \ldots, \omega_n^k$. For $k = 1, \ldots, n$, these are constructed using the reports about belief-states:

\[
\omega_i^k \equiv \hat{\pi}_i x_k + (1 - \hat{\pi}_i) y_k
\]

for $i \in \{1, \ldots, n\}$. We defined the CE change value $t_i$ above (identity (4.1)). Let $t_i^k$ denote the CE change for the $k$-th prize profile:

\[
t_i^k \equiv \frac{\phi'(w^k_i)}{\phi'(CE_k)} h_i
\]

(4.4)

Ask agents to report all values $t_i^k$ and $h_i$ for $i, k \in \{1, \ldots, n\}$. The agent reports $\hat{t}_i^k$ and $\hat{h}_i$.

II. Prize award and variations

3. Choose a number $p$ randomly from $[p, 1]$ (where $p$ is given by equation (4.3)) and award the act $x_E y$ with probability $p$ and the constant act $z_E z$ with probability $(1 - p)$.

4. Same as steps 3 and 4 in the mechanism specified in section 3.

The first result is that the agent does not want $\hat{D}_y(p)$ calculated from his report to be different from the true value $D_y(p)$. 15
Theorem 1. Under the mechanism with prize variations stated above, if $\hat{D}_y(p)$ (derived using the reported $\hat{B}$ and reported values $\hat{t}_i, \hat{h}_i$ for $i \in \{1, \ldots, n\}$) differs from the “true” value $D_y(p)$ for any $p \in [p, 1]$, the report is suboptimal.

While the formal proof is in the appendix, the discussion in section 3.1 explains the idea of the proof, which relies on a Vickrey-type dominance argument. Further, note that the act received by the agent does not depend on the agent’s report. The report only influences the calculation of the derivative $\hat{D}_y(p)$ and therefore only influences the variation received by the agent. This, combined with theorem 1 above, shows that any report that ensures $\hat{D}_y(p) = D_y(p)$ is optimal. It follows that truth-telling is an optimal strategy. In other words, the agent cannot do better than telling the truth by lying, and is at best indifferent between reporting truthfully or lying. We note this in the following corollary, which is useful for later arguments.

Corollary 1. Under the mechanism with prize variations, truth-telling is a best response. In other words, consider a report $(\hat{B})$ and $(\hat{t}^k_i, \hat{h}_i)_{i,k \in \{1, \ldots, n\}}$ where either $\hat{B} \neq B$, or $\hat{t}^k_i \neq t^k_i$, or $\hat{h}_i \neq h_i$ for any $i, k \in \{1, \ldots, n\}$ (with at least one of the reported elements untruthful). The agent must be indifferent between submitting any such untruthful report and reporting truthfully.

4.3 Eliciting Beliefs

Corollary 1 above says that truth-telling is an optimal strategy.

To focus on belief elicitation first, we now simplify the problem by assuming that the CE change values $t^k_i$ as well as $h_i$ values are known. This allows us to split the exposition. In this section we simply show how beliefs are elicited assuming CE changes are known. We show that under the mechanism with prize variations presented in the previous section augmented by an additional property, reporting $B$ truthfully is the uniquely optimal strategy. In section 5, we then show how $t^k_i$ and $h_i$ values are elicited. Since this latter part does not refer to beliefs, there is no contamination. To establish that the split is purely for expositional ease, we then combine the two mechanisms (belief elicitation and CE-change elicitation) and present a “grand mechanism” which elicits all beliefs and CE changes at the same time.

As noted above, the following assumption holds throughout this section.

Assumption 1. The numbers $t^k_i$ given by (4.4) and the beliefs $h_i$ are known.
Before we state the main result of this section on eliciting $B$, we need to introduce a matrix $\hat{A}$ of CE variations which plays a crucial role in belief elicitation.

Let $\hat{\pi} \equiv (\hat{\pi}_1, \ldots, \hat{\pi}_n)$ and $h \equiv (h_i, \ldots, h_n)$. Let $\text{CE}_k|\hat{\pi}, h$ denote the certainty equivalent of a second-order act with $\omega_i$ constructed from $\hat{\pi}$ and given beliefs $h$ over the ambiguity states.

$$\hat{A} \equiv \begin{pmatrix}
\frac{\phi'(\hat{\pi}_1 x_1 + (1 - \hat{\pi}_1) y_1)}{\phi'(\text{CE}_1|\hat{\pi}, h)} & \cdots & \frac{\phi'(\hat{\pi}_n x_1 + (1 - \hat{\pi}_n) y_1)}{\phi'(\text{CE}_1|\hat{\pi}, h)} \\
\vdots & \ddots & \vdots \\
\frac{\phi'(\hat{\pi}_1 x_n + (1 - \hat{\pi}_1) y_n)}{\phi'(\text{CE}_n|\hat{\pi}, h)} & \cdots & \frac{\phi'(\hat{\pi}_n x_n + (1 - \hat{\pi}_n) y_n)}{\phi'(\text{CE}_n|\hat{\pi}, h)}
\end{pmatrix} \quad (4.5)$$

As will become clear when we prove the main result on belief elicitation (theorem 3 below), it is crucial for the proof that the $n \times n$ matrix $\hat{A}$ above can be constructed to be of full rank by choosing the $n$ prize vectors appropriately. We show next that a sufficient condition for $\hat{A}$ to have full rank is that the function $\phi$ has non-zero derivatives (almost everywhere) of order at least $n$. Since the agent can report any number of belief-states, a sufficient condition is that $\phi$ has non-zero derivatives of all orders almost everywhere. We assume this:

**Assumption 2.** The function $\phi(\cdot)$ has non-zero derivatives of all orders almost everywhere on its domain (set of all possible expected utilities from the set of lottery prize vectors).

Note that this class includes exponential and log functions, power functions such as $x^t$ where $t < 0$ or $t$ is a positive fraction, functions such as $x^x$, trigonometric functions. Essentially the only class excluded is polynomials of positive integer degree. For this latter class, our proof works only so long as the degree exceeds the cardinality of the announced set of belief-states.

**Theorem 2.** Under assumption 2, it is possible to construct the matrix $\hat{A}$ to have full rank for any finite $n$.

The proof shows that the question can be translated into one about the rank of a Vandermonde matrix and then proven using its properties. We relegate this to the appendix.
Of course, to be able to choose prize vectors so that $\hat{A}$ has full rank, the mechanism designer must be able to observe the elements of $\hat{A}$. However, this is straightforward, since this is simply the matrix of the elements $t_{i}^{k}/h_{i}$ for $i, k \in \{1, \ldots, n\}$. Thus the mechanism designer can simply construct the matrix of reported values $\hat{t}_{i}^{k}/\hat{h}_{i}$ and if these values are reported truthfully (which is assumed in this section, and their truthful elicitation will be shown later), the matrix coincides with $\hat{A}$.

Consider the following mechanism.

$\Gamma_{B}$: A mechanism to elicit $B$: set of belief states and second-order beliefs

This is the same as the mechanism with prize variations outlined in the previous section with the following augmentation: in step 2, construct the matrix $\hat{A}$ (as noted above, this is possible). Check if this has full rank. If not, change some prize vectors and repeat the steps until step 2 until $\hat{A}$ has full rank (we know this is possible from Theorem 2 above).

We now state the main belief-elicitation result. This is a general version of the same result shown using examples in section 3.

**Theorem 3.** Suppose assumption 2 holds, and $\hat{A}$ is known. Then under the mechanism $\Gamma_{B}$, the agent's unique best response is to report $B$ truthfully, i.e. to report truthfully the belief states and the second-order belief attached to each state.

5 Eliciting CE changes

So far we have described the belief elicitation mechanism assuming that CE change values $t_{i}^{k}$ (given by equation (4.4)) and beliefs $h_{i}$ are known. We now discuss the mechanism that elicits these values. The mechanism below does not use any element of reports on beliefs $B$, and therefore there is no contamination of incentives for eliciting $B$. 

18
**An outline of the idea**  Before we describe the mechanism, let us describe the idea behind it. Suppose the agent reports 2 belief states, \( \hat{\pi}_1 \) and \( \hat{\pi}_2 \). The mechanism designer then chooses 2 prize vectors, and constructs 2 second-order acts using these: act \( k \) has the prize profile \( \omega^k_1, \omega^k_2 \) for \( k = 1, 2 \), where

\[
\omega^k_i \equiv \hat{\pi}_i x_k + (1 - \hat{\pi}_i) y_k
\]

The designer then also constructs a third second-order act, not using reports, but picking randomly one of the two prize vectors (say \( (x_1, y_1) \) gets picked for concreteness) and two numbers \( r_1, r_2 \) from \((0, 1)\). The act (act 3) then has prize profile \( \omega^3_1, \omega^3_2 \) where \( \omega^3_i = r_i x_1 + (1 - r_i) y_1 \) for \( i = 1, 2 \). Then the 3 acts are announced (in any order, so the agent does not know which act has prize profile constructed from \( r \) values) and for each act \( k = 1, 2, 3 \), the agent is asked to report \( t^k_i \). There are 2 ambiguity states associated with the prizes \( \omega^k_1 \) and \( \omega^k_2 \). The agent forms a belief \( h_1, h_2 \) on these (where \( h_1 + h_2 = 1 \)). The agent is also asked to report the beliefs \( h_i \).

Once \( t^k_i \) values and \( h_i \) are reported, the agent is presented with two schemes, A and B, and one of these is chosen randomly. Scheme A is essentially the mechanism in Karni (2009) and, as we clarify, a standard Vickrey-type dominance argument shows that it can elicit beliefs \( h_i \). For scheme B, the designer identifies the second-order act constructed using \( r \) values and, using reported \( t^3_i \) values, calculates \( D_y(0) \) - the prize variation that leaves the agent indifferent and then gives them the variation using the scheme as described in section 3.1. As we know, the agent has incentive to ensure that the value of \( D_y(0) \) calculated from reports coincides with the true value. To this end, the agent would want to truthfully report \( t^3_i \) values. But, as we show, the agent cannot be certain which prize vector is constructed from reports and which from chosen \( r \) values. Any belief the agent forms puts strictly positive probability that any of the acts presented to the agent is constructed using \( r \) values. Further, no advantage can be gained by misreporting \( t_i \) values when the prize vector is constructed from reported beliefs rather than \( r \) values. It is therefore uniquely optimal for the agent to report \( t^k_i \) values truthfully for all \( k \) under scheme B. Since reports are made before scheme A or B is chosen, reporting \( t^k_i \) and \( h_i \) values truthfully is uniquely optimal under the mechanism.

We now proceed to state the mechanism and derive results formally.
The mechanism makes use of an objective lottery: for $q \in [0, 1]$, let $\ell(q; w, z)$ denote the objective lottery that pays $w$ with probability $q$ and $z$ with probability $1 - q$.

$\Gamma_{\Delta}$: A mechanism for eliciting CE changes and beliefs on second-order acts

I. Reporting stages

1. As in the mechanism for belief elicitation, introduce $n$ prize vectors.

2. Construct $n + 1$ second order acts $k = 1, \ldots, n + 1$ as follows. Act $k$ has the prize profile $\omega^k_1, \ldots, \omega^k_n$ where $\omega^k_i$ is the wealth level associated with ambiguity state $s_i$. For $k = 1, \ldots, n$, wealth levels are constructed using the reports about belief-states:

$$\omega^k_i \equiv \hat{\pi}_i x_k + (1 - \hat{\pi}_i) y_k$$

for $i \in \{1, \ldots, n\}$.

For the remaining second order act, randomly pick a point $(r_1, \ldots, r_n)$ from the $n - 1$ dimensional simplex. Also pick one of the $n$ prize vectors randomly. Let $(x_s, y_s)$ be the prize vector picked. Then construct $\omega^{n+1}_i \equiv r_i x_s + (1 - r_i) y_s$ for $i \in \{1, \ldots, n\}$.

Recall that $h_i$ denotes the second-order belief that ambiguity state $s_i$ would be chosen and $t_i^k$ (given by equation (4.4)) denotes the change in CE$_k$ from change in $\omega^k_i$.

3. Choose a random re-ordering of the numbers $1, \ldots, n + 1$. Suppose $j$ is in position 1, $j'$ in position 2 etc.

Using the chosen re-ordering, announce each SOA sequentially to the agent (i.e. announce $k = j$ first, followed by $k = j'$ etc). For each SOA announced, ask the agent to report $t_i^k$ for all $i \in \{1, \ldots, n\}$.\footnote{Note that the agent is told only the prize profile of an act, but not told whether any given profile is constructed from reported belief-states or from selected $r_i$ values (this is the purpose of the re-ordering). As we show later, this plays a role in generating incentives.} The agent reports $n \times (n + 1)$ values $\hat{t}_i^k$.

Also ask the agent to report $h_i$ for all $i = 1, \ldots, n$, where $h_i$ is the second-order belief that ambiguity state $s_i$ would be chosen. The agent reports $\hat{h}_i$.\footnote{Note that the agent is told only the prize profile of an act, but not told whether any given profile is constructed from reported belief-states or from selected $r_i$ values (this is the purpose of the re-ordering). As we show later, this plays a role in generating incentives.}
II. Elicit $t_i^k$ and beliefs $h_i$

Announce the SOA where the prize vector is $(x_s, y_s)$ and ambiguity state $s_i$ involves picking $x_s$ with probability $r_i$, $i \in \{1, \ldots, n\}$.

Choose scheme A or B with probability 1/2 each.

1. **scheme A:** Select a state $s_i$ randomly and select $q \in [0, 1]$. Then if $q \geq h_i$ award objective lottery $\ell(q, x_s, y_s)$ and otherwise award the SOA.

2. **scheme B:** Using the reported $t_{i}^{n+1}$ values, calculate $D^*_y(0)$ (the prize variation that leaves the agent indifferent) such that

$$\sum_i \hat{\pi}_{i}^{n+1} \left( r_i + (1 - r_i)D^*_y(0) \right) = 0$$

Then pick $d$ randomly from $[-\overline{d}, 0]$ and award the variation if and only if $|d| \leq |D^*_y(0)|$. Otherwise award the original prize vector $(x_s, y_s)$.

§

We now state the elicitation result.

**Theorem 4.** Under the mechanism $\Gamma_\Phi$, the agent’s unique best response is to report $t_i^k$ and $h_i$ truthfully for $i \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, n + 1\}$.

The proof in the appendix follows the idea described at the start of this section.
6 A grand mechanism to elicit $B$ and CE changes concurrently

We have so far constructed the mechanism $\Gamma_B$ in section 4.3 to elicit $B$ assuming $t^k_i$ and $h$ are known and $\Gamma_\phi$ in section 5 to elicit $t^k_i$ and $h$ (the latter does not require any knowledge of $B$). In this section we use these to construct a grand mechanism to elicit $B, t^k_i$ and $h$ together.

The grand mechanism is a straightforward combination of the two mechanisms discussed so far:

$\Gamma$: A grand mechanism to elicit $B, t^k_i$ and $h$ concurrently

I. Reporting stages

These are as in the mechanism $\Gamma_\Delta$: the mechanism for eliciting CE changes and beliefs on second-order acts in section 5. Also, as in mechanism $\Gamma_B$, assume $\hat{A}$ has been constructed to have full rank.

II. Design of incentives

Run

- part II of the mechanism $\Gamma_B$ with probability 1/2, and
- part II of the mechanism $\Gamma_\Delta$ with probability 1/2.

Corollary 1 shows that under the mechanism with prize variations introduced in section 4.2 – and therefore also under mechanism $\Gamma_B$ – reporting beliefs $B$, CE change values $t^k_i$ and belief on SOA $h$ truthfully is an optimal strategy. It follows that at worst the agent is indifferent between lying and telling the truth.

Next, under mechanism $\Gamma_\Delta$, reporting $t^k_i$ and $h$ truthfully is uniquely optimal (Theorem 4). Also note that this mechanism does not use any information on $B$. Since mechanism $\Gamma_\Delta$ is offered with positive probability, truthful report of $t^k_i$ and $h$ is uniquely optimal.
It follows that we need to only consider reports of the type \((t^k_i, h, \hat{B})\), and then rule out \(\hat{B} \neq B\). This is accomplished by Theorem 3, which shows that if \(t^k_i\) and \(h\) are known and assumption 2 holds, reporting \(B\) truthfully is uniquely optimal under mechanism \(\Gamma_B\). Since \(\Gamma_B\) is offered with positive probability, reporting \(B\) truthfully is uniquely optimal.

It follows that under the grand mechanism \(\Gamma\), it is uniquely optimal to report \(t^k_i, h\) and \(B\) truthfully. This proves the result below.

**Theorem 5.** Suppose assumption 2 holds. Under the grand mechanism \(\Gamma\) above, the uniquely optimal response of the agent is to report \(B\) and \(t^k_i, h\) truthfully.

## 7 Literature review

Let us now describe the relevant literature briefly. A large literature on scoring rules addresses the problem of eliciting beliefs on observable events. See Gneiting and Raftery (2007) for a discussion of this literature. Recent surveys by Schotter and Trevino (2014) and Schlag, Tremewan, and van der Weele (2015) provide superb accounts of theoretical and practical issues with procedures used in experiments to elicit subjective beliefs. The procedures in this literature are however not related to our belief-elicitation procedure since our task is to elicit unobservable belief-states and beliefs on such states.

The literature on eliciting beliefs of agents who have non-standard preferences is relatively small. Chambers (2008) studies belief-elicitation of agents with maxmin expected utility preference using proper scoring rules. The main result is that the agent announces a single probability belonging to her set of priors. Bose and Daripa (2016) address the problem of belief-elicitation under the more general \(\alpha\)-maxmin preferences, and, in contrast to the above, elicit the entire set of beliefs as well as the relevant preference parameter for \(\alpha\)-maxmin (elicitation of the entire set of beliefs under maxmin is a special case).

The papers mentioned above do not face the question of eliciting beliefs over unobservable states. This question does arise in Karni (2016), who considers elicitation of beliefs of decision makers who have non-standard preferences but different from the one we focus on. In Karni’s work, the agent faces a multi-stage environment. The agent faces
Knightian uncertainty in period 0, but is an SEU maximizer in period 1 when this uncertainty is resolved. The mechanism elicits the beliefs in period 0 by allowing the choice between acts and lotteries to be delayed to period 1, allowing SEU payoff comparisons across choices. The fact that agents are SEU when making a decision is crucial for elicitation of beliefs in this work. In contrast, we have the more standard problem where the agent is sensitive to ambiguity, and must make choices while facing ambiguity. For this problem, the procedure in Karni (2016) cannot be applied, and our methods are quite different.

To summarise, all existing belief-elicitation procedures depend on observability of the belief-relevant event and/or agents being SEU maximizers for making choices. The problem we study is a departure from this, and we present novel belief-elicitation methods to address it.

8 Conclusion

We study elicitation of subjective beliefs of an agent facing model uncertainty or ambiguity. The agent has non-singleton (first-order) priors on an event. Each prior (belief-state) refers to a different underlying data generating process and the corresponding probability model. The agent then has a second-order prior on these first-order belief-states. The agent’s preference exhibits non-neutrality to the second-order distribution. As discussed at the outset, such a two-stage decomposition of uncertainty and non-reduction of compound lotteries resulting from non-neutrality to the second-order distribution plays an important role in the literature in resolving the Ellsberg Paradox.

The large belief-elicitation literature focuses on eliciting beliefs on events that are observable. The entire scoring rule literature fits into this category. Here, on the other hand, we elicit (second-order) beliefs on unobservable belief-states. A few recent papers do also address problems pertaining to belief-elicitation on unobservable states, but these require the agent to be SEU maximizer at some decision making stage. In this paper, in contrast, we consider the more standard problem of eliciting first and second-order be-

14Either the agent has incomplete preferences (à la Bewley (2002)) in period 0 and beliefs are determined in period 1, or the agent is a Bayesian decision maker who in period 0 entertains the possibility of a range of possible posteriors for period 1 with a prior in period 0 over (information signals corresponding to) the set of the posteriors. In either case, the agent is an SEU maximizer (with a unique prior) in period 1.
liefs of agents who are ambiguity-sensitive and do in fact face ambiguity when making a decision. Here we adopt the smooth ambiguity representation of KMM for specificity. However, as discussed, several other axiomatisations result in a similar representation structure.

The literature on belief elicitation on observable states compares levels of utility from acts and lotteries. Here, with unobservable states and ambiguity-sensitive agents, we show that we need to consider how utility changes under report-dependent prize variations. Using such variations, we construct a novel direct revelation mechanism that induces truthful reporting of the first-order belief states as well as the second-order distribution on the belief-states as the unique best response. The precise attitude towards ambiguity is immaterial: we elicit the beliefs for any increasing function $\phi$ that satisfies a certain “smoothness” property. The mechanism requires knowledge of the responsiveness of certainty equivalents of acts to variations in prizes, and we also elicit these CE changes.

While we consider a smooth ambiguity representation to elicit second-order beliefs, our idea of using report-dependant prize-variations is quite general, and should be applicable for eliciting beliefs on unobservable states more broadly.
A Appendix

A.1 Proof of Lemma 1

To see this, let us calculate $z_{12}$. Let $\phi'_k \equiv \phi'(\pi_ix_k + (1 - \pi_i)y_k)$. The $k_i$-th element of $A$ is $a_{ki} = \phi'_k / \phi'(CE_i|B)$. Given $\hat{\pi}_1 = \pi_1$ and $\hat{\pi}_2 = \pi_2$, $\hat{a}_{ki} = \phi'_k / \phi'(CE_i|\hat{B})$. Note that $a_{ki}$ and $\hat{a}_{ki}$ have the same numerator, but potentially different denominators. We have

$$A = \begin{pmatrix} \phi'_{11} / \phi'(CE_1|B) & \phi'_{12} / \phi'(CE_1|B) \\ \phi'_{21} / \phi'(CE_2|B) & \phi'_{22} / \phi'(CE_2|B) \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} \phi'_{11} / \phi'(CE_1|\hat{B}) & \phi'_{12} / \phi'(CE_1|\hat{B}) \\ \phi'_{21} / \phi'(CE_2|\hat{B}) & \phi'_{22} / \phi'(CE_2|\hat{B}) \end{pmatrix}$$

It follows that

$$z_{12} = T \left( \frac{\phi'(CE_2|\hat{B})}{\phi'(CE_2|B)} - \frac{\phi'(CE_1|\hat{B})}{\phi'(CE_1|B)} \right)$$

where $T = \frac{\phi'_{11} \phi'_{22} - \phi'_{21} \phi'_{12}}{\phi'_{11} \phi'_{22} - \phi'_{12} \phi'_{21}}$. To have $z_{12} = 0$ for all choices of $(x_1,y_1)$ and $(x_2,y_2)$, the ratios $\frac{\phi'(CE_i|\hat{B})}{\phi'(CE_i|B)}$, $i \in \{1,2\}$, must be constant across choices of prize vector $x,y$. But as $y \to x$, in the limit the CE is $x$ in both numerator and denominator, so that the ratio is equal to 1. Since the ratio is continuous in $x$ and $y$, and goes to the limit in a continuous manner, the ratio must be equal to 1 for all values of $x,y$.\footnote{Otherwise there will be a region of prize values for which the ratio changes with $x,y$, which would imply that for some choices of prize vectors, the ratio is different from 1 for $i = 1$ while equal to 1 for $i = 2$ implying $z_{12} \neq 0$.}

Since $\phi'(CE_i|\hat{B}) = \phi'(CE_i|B)$ for $i \in \{1,2\}$, it follows that $A = \hat{A}$, implying that $\hat{A}^{-1}A$ is the identity matrix.\|

A.2 Proof of Theorem 1

Recall that $D_y(p)$ is given by equation (3.1). The expression on the left hand side is the change in CE from variation $\text{Var}(x,y)$ and that on the right hand side is the change in CE from variation $\text{Var}(z)$. The derivative $D_y(p)$ is such that the two are equal. Note that the
change from \( \text{Var}(x, y) \) is preferred for \(|d| < |D_y(p)|\) and that from \( \text{Var}(x, y) \) is preferred for \(|d| > |D_y(p)|\).

Suppose \(|\hat{D}_y(p)| > |D_y(p)|\).

If either \(|d| > |\hat{D}_y(p)|\) or \(|d| \leq |D_y(p)|\) reporting truthfully or misreporting either beliefs or CE changes (or both) does not make any difference. In the first case, the agent receives (if \(z \in z \) is picked, which happens with probability \((1 - p)\)) the change from \( \text{Var}(z) \) whether the report is truthful or not. In the second case, the agent receives (with probability \(p\)) the change from \( \text{Var}(x, y) \) whether the report is truthful or not. However, if \(|\hat{D}_y(p)| \leq |d| > |D_y(p)|\), the agent receives \( \text{Var}(x, y) \) (with probability \(p\)) but this in fact reduces the agent’s payoff. If the agent reported truthfully, the agent would receive (with probability \((1 - p)\)) \( \text{Var}(z) \) which would improve payoff. Since this last case happens with strictly positive probability, any report resulting in \(|\hat{D}_y(p)| > |D_y(p)|\) is strictly dominated by a truthful report.

A similar argument shows that any report resulting in \(|\hat{D}_y(p)| < |D_y(p)|\) is strictly dominated by a truthful report. This completes the proof.

A.3 Proof of Theorem 2

We show here that under assumption 2, it is possible to find prize vectors \((x_1, y_1), \ldots, (x_n, y_n)\) such that \(\hat{A}\) (given by equation (4.5)) has full rank.

In what follows, all \(\pi\)-values are reports (i.e. they all have a “hat” on top). Since there is no possibility of confusion, for economy of notation, we remove all “hat” symbols in the proof that follows. Further, since we need to repeatedly use expressions of the form \(\phi'(\hat{\pi}_i x_k + (1 - \hat{\pi}_i)y_k)\), we shorten it to \(\phi'(\pi_i, x_k, y_k)\) (again, note that we are also removing the hat on \(\pi_i\)).

Since multiplying any row by a scalar does not change rank, instead of carrying the expressions in the denominator of each term, we simply show that the following matrix \(A\) can be constructed to have full rank. It then follows directly that the matrix \(\hat{A}\) in equation (4.5) can be constructed to have full rank.
The rest of the proof now shows that we can construct $A$ to have full rank $n$.

Let $M(k)$ be the following $k \times k$ minor of $A$, where $2 \leq k \leq n$.

$$M(k) = \begin{pmatrix}
\phi'(\pi_1, x_1, y_1) & \phi'(\pi_2, x_1, y_1) & \cdots & \phi'(\pi_n, x_1, y_1) \\
\vdots & \vdots & \ddots & \vdots \\
\phi'(\pi_1, x_n, y_n) & \phi'(\pi_2, x_n, y_n) & \cdots & \phi'(\pi_n, x_n, y_n)
\end{pmatrix}$$

Note that a sufficient condition for the rank of $A$ to be at least $k$ is that $M(k)$ is full rank.

To show that it is possible to construct $A$ so that it has full rank $n$, we use a proof by induction. Let $R(X)$ denote the rank of matrix $X$. We show that for $k = 2$, the minor $M(2)$ has full rank, establishing that $R(A) \geq 2$. We then show that if $M(k-1)$ has full rank, then we can construct $M(k)$ so that the latter has full rank $k$, for any $k \in \{3, \ldots, n\}$. In other words, $R(A) \geq k - 1$ implies $R(A) \geq k$. Since $M(2)$ has rank 2, this shows that $M(3) \ldots M(n)$ can be constructed to have full rank. But $M(n)$ is simply the matrix $A$, which proves that $R(A) = n$.

### A.3.1 Step 1: $M(2)$ has full rank

We first show that $M(2)$ can be constructed to have rank 2.

Suppose not. Suppose for all possible choices of $(x_2, y_2)$, $M(2)$ has rank 1.

It follows that there are numbers $\beta_1$ and $\beta_2$ not both zero such that

$$\begin{pmatrix}
\phi'(\pi_1, x_1, y_1) & \phi'(\pi_2, x_1, y_1) \\
\phi'(\pi_1, x_2, y_2) & \phi'(\pi_2, x_2, y_2)
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\
\beta_2
\end{pmatrix}
= \begin{pmatrix}
0 \\
0
\end{pmatrix}$$

Since

$$\beta_1 \phi'(\pi_1, x_2, y_2) + \beta_2 \phi'(\pi_2, x_2, y_2) = 0 \quad (A.1)$$

28
for all possible choices of \( x_2, y_2 \), it follows that a necessary condition for equation (A.1) to hold is that the derivative of the expression on the left hand side with respect to \( x_2 \) must also be zero:

\[
\beta_1 \pi_1 \phi''(\pi_1, x_2, y_2) + \beta_2 \pi_2 \phi''(\pi_2, x_2, y_2) = 0. \tag{A.2}
\]

Further, suppose \( \beta_1 + \beta_2 \neq 0 \). We can write equation (A.1) as

\[
\beta_1 \left( \phi'(\pi_1, x_2, y_2) - \phi'(\pi_2, x_2, y_2) \right) = - (\beta_1 + \beta_2) \phi'(\pi_2, x_2, y_2).
\]

But then by choosing \( y_2 \) arbitrarily close to \( x_2 \), we can make the left hand side arbitrarily close to zero, while the right hand side is bounded away from zero, thus violating equation (A.1), which is a contradiction.

It follows that for equation (A.1) to hold for all possible choices of \( x_2, y_2 \), a necessary condition is

\[
\beta_1 + \beta_2 = 0 \tag{A.3}
\]

Starting from equation (A.2) and using exactly the same argument, it follows that a further necessary condition for equation (A.1) to hold for all possible choices of \( x_2, y_2 \) is given by

\[
\pi_1 \beta_1 + \pi_2 \beta_2 = 0. \tag{A.4}
\]

However, it is easy to see that the necessary conditions (A.3) and (A.4) can only be satisfied if both \( \beta_1 \) and \( \beta_2 \) are 0. To see this, write the necessary conditions (A.3) and (A.4) as

\[
\begin{pmatrix}
1 & 1 \\
\pi_1 & \pi_2
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\
\beta_2
\end{pmatrix}
=
\begin{pmatrix}
0 \\
0
\end{pmatrix}
\]

Since the first matrix on the left hand side has rank 2 (the determinant is \( \pi_2 - \pi_1 > 0 \)), the only solution is \( \beta_1 = \beta_2 = 0 \).

It follows that equation (A.1) cannot be satisfied for all possible choices of \( x_2, y_2 \) unless all \( \beta \)-values are 0, implying that it must be possible to have some value \( x_2, y_2 \) for which \( M(2) \) has rank 2.

**A.3.2 Step 2: \( M(k) \) has full rank if \( M(k - 1) \) has full rank**

Suppose \( M(k - 1) \) is full rank but \( M(k) \) has rank \( k - 1 \) for all possible values of the prize vector \((x_k, y_k)\). Then there exists \((\beta_1, \ldots, \beta_k)\) not all zero such that for every value
of \((x_k, y_k)\),

\[
\sum_{i=1}^{k} \beta_i \tilde{c}_i = 0, \quad (A.5)
\]

where \(\tilde{c}_i\) is the \(t\)-th column vector

\[
\left( \begin{array}{c} 
\phi' (\pi_t, x_1, y_1) \\
\vdots \\
\phi' (\pi_t, x_k, y_k) 
\end{array} \right)
\]

Since \(\sum_{i=1}^{k} \beta_i \phi' (\pi_t, x_k, y_k) = 0\) for all possible values of \((x_k, y_k)\) where not all \(\beta_t\) values are 0, we get the following result.

**Lemma 2.** Consider any \(k \in \{1, \ldots, n\}\). The assumption that equation (A.5) holds for all values of the prize vector \((x_k, y_k)\) implies that for any \(\ell \in \{1, 2, \ldots, k\}\),

\[
\sum_{i=1}^{k} \beta_i \pi_i^\ell-1 \phi^{(\ell)} (\pi_t, x_k, y_k) = 0, \quad (A.6)
\]

where \(\phi^{(\ell)} (\cdot)\) denotes the \(\ell\)-th partial derivative of \(\phi (\cdot)\) with respect to \(x_k\), and where not all values of \(\beta_t\) are zero.

**Proof:** Note that we have assumed that all derivatives of \(\phi (\cdot)\) of orders up to \(k\) exist and are non-zero almost everywhere.

Also note that since we assume equation (A.5) holds, the equation is true for \(\ell = 1\) by assumption.

Fix any \(y_k\). Let

\[
g(x, \ell) \equiv \sum_{i=1}^{k} \beta_i \pi_i^\ell-1 \phi^{(\ell)} (\pi_t, x, y_k).
\]

and suppose that \(g(x, 1) = 0\) for all values of \(x\).

Consider the case of \(\ell = 2\). Suppose \(g(x, 2) \neq 0\) for some value of \(x\). Specifically, suppose \(g(x, 2) > 0\) for some \(x\). By continuity, there is some non-empty interval \([a, b]\) of values of \(x\) such that \(g(x, 2) > 0\) on \((a, b)\). In this case if we choose \(x \in (a, b)\), then by raising \(x\) slightly \(g(x, 1)\) can be raised. Therefore it cannot be true that \(g(x, 1) = 0\) for all values of \(x\), which is a contradiction.

Now, for any \(\ell \in 2, \ldots, k\), suppose \(g(x, \ell - 1) = 0\) for all values of \(x\). This implies that \(g(x, \ell) = 0\) for all values of \(x\). To see this, suppose \(g(x, \ell) > 0\) for some value of \(x\). As
above, this implies that \( g(x, \ell) > 0 \) for values of \( x \) in some interval \((a, b)\). Choosing \( x \) in the interval and then raising \( x \) would then raise \( g(\cdot, \ell - 1) \) above 0, which is a contradiction.

This proves that if \( g(x, \ell - 1) = 0 \) for all values of \( x \), then \( g(x, \ell) = 0 \) for all values of \( x \).

Since we have shown that \( g(x, 1) = g(x, 2) = 0 \) for all values of \( x \), it follows that \( g(x, \ell) = 0 \) for all values of \( \ell = 1, \ldots, k \). This completes the proof. \( \| \)

The next result derives some necessary conditions for equation (A.5) to hold for all prize vectors.

**Lemma 3.** Consider any \( k \in \{2, \ldots, n\} \). The following conditions are necessary for equation (A.5) to hold for all prize vectors \((x_k, y_k)\). For any \( \ell \in \{1, \ldots, k\} \),

\[
\sum_{t=1}^{k} \beta_t \pi_t^{\ell-1} = 0,
\]

(A.7)

where not all values of \( \beta_t \) are 0.

**Proof:** Fix any \( \ell \in \{1, \ldots, k\} \) and suppose

\[
\sum_{t=1}^{k} \beta_t \pi_t^{\ell-1} \neq 0.
\]

(A.8)

From Lemma 2, we know that a necessary condition for equation (A.5) to hold for all prize vectors \((x_k, y_k)\) is equation (A.6):

\[
\sum_{t=1}^{k} \beta_t \pi_t^{\ell-1} \phi^{(\ell)}(\pi_t, x_k, y_k) = 0.
\]

The equation above can be written as

\[
\sum_{t=1}^{k-1} \beta_t \pi_t^{\ell-1} \left( \phi^{(\ell)}(\pi_t, x_k, y_k) - \phi^{(\ell)}(\pi_k, x_k, y_k) \right) = - \sum_{t=1}^{k} \beta_t \pi_t^{\ell-1} \phi^{(\ell)}(\pi_k, x_k, y_k)
\]

It follows from equation (A.8) that the right hand side is not 0.

By choosing \( y_k \) very close to \( x_k \) we can make the left hand side as small as we like, while the right hand side gets close to

\[
- \sum_{t=1}^{k} \beta_t \pi_t^{\ell-1} \phi^{(\ell)}(\pi_k, x_k, x_k)
\]
which is bounded away from zero. It follows that equation (A.6) is violated. This is a contradiction. This proves the result.

Writing the necessary conditions given by lemma 3 in matrix form, we get

\[
\begin{pmatrix}
1 & 1 & \ldots & 1 \\
\pi_1 & \pi_2 & \ldots & \pi_k \\
\pi_1^2 & \pi_2^2 & \ldots & \pi_k^2 \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1^{k-1} & \pi_2^{k-1} & \ldots & \pi_k^{k-1}
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_k
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

Now, let \( V(k) \) denote the first matrix on the left hand side. The transpose of \( V(k) \) is a \( k \times k \) Vandermonde matrix. It follows that \( \det(V(k)^T) = \prod_{1 \leq i < j \leq n} (\pi_j - \pi_i) \).

Since \( \pi_j \neq \pi_i \) for any \( j \neq i \), the determinant is non-zero. Therefore the rank of \( V(k)^T \) is \( k \). Since taking a transpose does not change rank (or, indeed, the determinant), \( V(k) \) has full rank as well.

Since \( V(k) \) has full rank, there is a unique solution for the \( \beta \)-values. Since \( \beta_t = 0 \) for all \( t = 1, \ldots, k \), is a solution, this must be the only solution.

This implies that the null-space of the minor \( M(k) \) contains only the zero-vector, implying that \( M(k) \) has \( k \) linearly independent columns, indicating that \( R(M(k)) = k \).

Thus, starting from the assumption that \( M(k - 1) \) has full rank, we have shown that we can find a prize vector \( x_k, y_k \) such that \( M(k) \) has full rank. We have also shown that \( M(2) \) has full rank. It follows by induction that \( M(k) \) can have full rank for all values of \( k \in \{2, \ldots, n\} \). Since \( M(n) \) is simply the matrix \( A \), this proves that we can find \( n \) prize vectors such that \( A \) has full rank.

### A.4 Proof of Theorem 3

The proof delineates a general version of the same arguments as in section 3, which established the same result using an example.

\[16\] This is a standard result in matrix algebra. See, for example, Horn and Johnson (2013), chapter 0.9.11.
Using exactly the same arguments as in section 3, we obtain that to maintain \(D_y(p)\) at its true value, the agent’s reports must satisfy the following two conditions:

\[
\hat{A}\hat{P}\hat{M} = APM.
\]

\[
\hat{A}\hat{M} = AM.
\]

These imply that we must have

\[
\hat{A}^{-1}A = \hat{P}^{-1}\hat{A}^{-1}AP. \tag{A.9}
\]

Let us now show that it is not possible to satisfy equation (A.9) unless \(\hat{M} = M\) and \(\hat{P} = P\), implying that truth-telling is uniquely optimal.

Note that \(\hat{A}^{-1}A\) is \(n \times m\) with typical element \(z_{ij}\), where \(i \in \{1, \ldots, n\}\) and \(j \in \{1, \ldots, m\}\).

The matrix \(\hat{P}^{-1}\hat{A}^{-1}AP\) is also \(n \times m\) with typical element \(\frac{\pi_j}{\hat{\pi}_i}z_{ij}\). Thus condition (A.9) requires

\[
z_{ij} = \frac{\pi_j}{\hat{\pi}_i}z_{ij} \tag{A.10}
\]

for all \(i \in \{1, \ldots, n\}\) and \(j \in \{1, \ldots, m\}\). Note that this either requires \(z_{ij} = 0\), or if \(z_{ij} \neq 0\), then \(\pi_j = \hat{\pi}_i\).

Since \(\hat{M} = \hat{A}^{-1}AM\), and \(\hat{M}\) is \(n \times 1\) with each element non-zero, \(\hat{A}^{-1}A\) must have at least one non-zero element in each row. It follows that for equation (A.10) to hold, the following must hold:

each \(\hat{\pi}_i\) must be equal to some \(\pi_j\). \hspace{1cm} (\ast)

**Case A:** \(n > m\): If \(n > m\), it is not possible to satisfy condition (\ast) for all \(\hat{\pi}_i\) since there are more values \(\hat{\pi}_i\) than values \(\pi_j\).

**Case B:** \(n < m\): Let \(Z\) be a \(n \times (m - n)\) matrix with typical element denoted by \(z_{ij}\). Let \(I_n\) denote the \(n \times n\) identity matrix. Let \(k_i, i \in \{1, \ldots, n\}\) be \(n\) numbers such that \(1 \leq k_1 < k_2 \ldots < k_n \leq m\). Let \(Z \oplus I_n\) denote an augmented matrix that inserts columns from \(I_n\) in the \(Z\) matrix as follows: the first column of \(I_n\) is inserted after the \((k_1 - 1)\)-th column of \(Z\), second column of \(I_n\) inserted after the \((k_2 - 2)\)-th column of \(Z\) and so on until the last column of \(I_n\) inserted after the \((k_n - n)\)-th column of \(Z\).

Let \(Z_1\) denote the matrix formed by the first \((k_1 - 1) \geq 0\) columns of \(Z\), \(Z_2\) denote the matrix formed by the next \((k_2 - k_1 - 1) \geq 0\) columns of \(Z\) and so on until \(Z_{n+1}\) formed by
the last \((k_m - n - k_{m-n-1} - 1) \geq 0\) columns of \(Z\). The augmented matrix has the following form:

\[
Z \oplus I_n \equiv \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{pmatrix}
\]

The proof proceeds through the following Lemma. Let \(Z_0\) denote an \(n \times (m - n)\) matrix where each element in the matrix is 0.

**Lemma 4.** To satisfy equation (A.9), a necessary condition is that the matrix \(\tilde{A}^{-1} A\) is of the form \(Z_0 \oplus I_n\).

**Proof:** To satisfy condition \((\ast)\), let \(k_i, i \in \{1, \ldots, n\}\) be such that \(\tilde{\pi}_i = \pi_{k_i}\), where \(1 \leq k_1 < k_2 \ldots < k_n \leq m\).

Recall that \(z_{ij}\) denotes the \(ij\)-th element of \(\tilde{A}^{-1} A\). In this case, we must have \(z_{ij} = 0\) for columns \(j \neq k_i\) for any \(i \in \{1, \ldots, n\}\). But any element \(z_{ij}\) is the product of a row vector (of \(\tilde{A}^{-1}\)) and a column vector (of \(A\)) and thus a sum of \(n\) terms, with term \(\ell\) weighted by \(\phi(CE_{i\ell}^\prime|\tilde{\pi}, h)|\phi(CE_{i\ell}^\prime|B)|\), \(\ell \in \{1, \ldots, n\}\). As shown in the proof of lemma 1, each of these terms must be equal to 1 to ensure \(z_{ij} = 0\).

Given this, it follows that the matrix \(\tilde{A}\) is simply the matrix formed by columns \(k_1, k_2, \ldots, k_n\) of matrix \(A\). It follows that \(\tilde{A}^{-1} A\) is of the form \(Z \oplus I_n\).

Therefore, each row has one element 1. Consider row \(i\). The element in the \(k_i\)-th column is 1. Suppose some \(z_{ij} \neq 0\) where \(j \neq k_i\). Then equation (A.10) would require \(\tilde{\pi}_i = \pi_j\). This is impossible since we already have \(\tilde{\pi}_i = \pi_{k_i}\). It follows that row \(i\) must have \(z_{ij} = 0\) for all \(j \neq k_i\). This implies that all elements of the matrix \(Z\) must be 0. It follows that \(\tilde{A}^{-1} A\) is of the form \(Z_0 \oplus I_n\). This completes the proof of Lemma 4.

Since \(\tilde{M} = \tilde{A}^{-1} A F\), and \(\tilde{A}^{-1} A\) is of the form \(Z_0 \oplus I_n\) (from Lemma 4 above), it follows that

\[
\tilde{\mu}_i = \mu_{k_i}.
\]

But since \(n < m\), \(\{k_1, \ldots, k_n\}\) is a strict subset of \(\{1, \ldots, m\}\), it follows that \(\sum_{i=1}^{n} \tilde{\mu}_i = \sum_{i=1}^{n} \tilde{\mu}_{k_i} < 1\), which is impossible. Therefore we can rule out the case \(n < m\).
Above we ruled out $n > m$ and $n < m$. This proves that we cannot have $n \neq m$. Then $n = m$. In this case, $z_{ij} = \frac{\pi_j}{\pi_i} z_{ij}$ requires each $\pi_i$ to be equal to a distinct $\pi_j$. Since $n = m$, the reported vector must coincide with the true vector. In other words, we have $\hat{P} = P$. But then $\hat{A} = A$. Therefore, $\hat{A}^{-1} A$ is the $m \times m$ identity matrix, and $\hat{M} = \hat{A}^{-1} A M = M$. This completes the proof.||

### A.5 Proof of Theorem 4

**Step 1** First, let us show that under scheme A, reporting $h_i$ truthfully for all $i$ is uniquely optimal. Suppose for some $i \in \{1, \ldots, n\}$, $\hat{h}_i \neq h_i$. The state associated with $h_i$ is chosen with strictly positive probability. Suppose this is chosen, and suppose $h_i > \hat{h}_i$. Then if $q \geq h_i$ or $q < \hat{h}_i$, the misreport does not affect payoff. However, if $\hat{h}_i \leq q < h_i$, the agent receives the objective lottery while the SOA would be preferred. Similarly, $h_i < \hat{h}_i$ is suboptimal. It follows that scheme A elicits $h_i$ for $i \in \{1, \ldots, n\}$. Scheme A is essentially the mechanism in Karni (2009) and its chief advantage is that it elicits beliefs $h_i$ on an SOA while allowing for the mechanism to remain agnostic about CE changes.

**Step 2** Next, let us show that for any SOA $k$, scheme B elicits $t^k_i$ for all $i$.

**Step 2.1** First, consider the case in which $\omega_i = r_i x_s + (1 - r_i) y_s$ for some chosen profile $r_1, \ldots, r_n$ from the $n - 1$ dimensional simplex and some prize vector $x_s, y_s$. Let us show that in this case it is optimal to report $t_i$ values truthfully. To see this, note that it is optimal for the agent to report $\hat{t}_i$ to ensure that $D^*_y(0)$ is the same as under truthtelling. But since $\hat{t}_i$ are announced before the agent learns the chosen $r_i$ values, if $\hat{t}_i \neq t_i$ for some $i$, it is not possible to maintain $D^*_y(0)$ at its true value for all possible choices of the vector $(r_1, \ldots, r_n)$ from the $n - 1$ dimensional simplex.

**Step 2.2** Next, note that at the reporting stages, the agent is simply presented the prize profile $\omega^k_1, \ldots, \omega^k_n$ for the $k$-th SOA, $k \in \{1, \ldots, n + 1\}$. Now, misreporting $t_i$ values has no payoff impact if $\omega_i \neq r_i x_s + (1 - r_i) y_s$ for some chosen profile $r_1, \ldots, r_n$. However, as we show below, for any given prize profile, it is not possible for the agent to be certain that $\omega_i \neq r_i x_s + (1 - r_i) y_s$. In other words, whatever belief the agent forms over how $\omega_i$ is constructed, the agent must place strictly positive probability on the case analyzed in

---

17 This is true for the same Vickrey-type dominance reason explained in section 3.1 and formalized in Theorem 1. We omit the details.
step 2.1. This then implies that reporting $t_i$ values truthfully is uniquely optimal.

To see that the agent must place strictly positive probability on the case analyzed in step 2.1, note that for any announced prize profile $\omega_1, \ldots, \omega_n$, the agent can try to solve $n$ equations of the form $\omega_i = m_i x + (1 - m_i) y$. Further, $\sum_i m_i = 1$. There are $n$ values of $m_i$ plus $x$ and $y$ - a total of $n + 2$ unknowns, but only $n + 1$ equations. The 1 degree of freedom suggests it is impossible that the set of solutions only contain $m_i = \hat{\pi}_i$ (this is true even if $\omega_i$ is indeed formed as $\hat{\pi}_i x + (1 - \hat{\pi}_i) y$ for some prize vector $x, y$). In other words, whatever beliefs the agent forms about the possible values of $m_i$, any such belief must attach positive probability to $m_i \neq \hat{\pi}_i$, i.e. positive weight to the possibility that $\omega_i$ arise from some vector $(r_1, \ldots, r_n)$ chosen from the $n - 1$ dimensional simplex and some prize vector $x, y$.

**Step 3** Step 1 shows that under scheme A, reporting $h_i$ values truthfully is uniquely optimal, and step 2 shows that reporting $t_i$ values truthfully is uniquely optimal. Note that under scheme B, there is no advantage gained by misreporting $h_i$, and under scheme A, no advantage is gained by misreporting $t_i$ values. Since either scheme is chosen with strictly positive probability, it follows that under the mechanism, it is uniquely optimal for the agent to report $h_i$ and $t_i$ values truthfully. ||
References


