Precision May Harm: The Comparative Statics of Imprecise Judgement

Sean Horan
University of Montreal and CIREQ

Paola Manzini
University of Sussex and IZA

Marco Mariotti
Queen Mary University of London

This version: 25 May 2018

Abstract

We consider an agent whose information about the objects of choice is imperfect in two respects: first, the values of the objects are perceived with error; and, second, the realised values cannot be discriminated with perfect precision. Reasons for imprecise discrimination include coarseness in sensory perception, memory function, or the technology that experts use to communicate with decision-makers. We study the effect on the quality of decision making of increasing precision. When values are perceived without error, more precision is unambiguously beneficial. We show when this ceases to be true if perception errors can arise. Our results have practical implications. They define, for example, conditions where requiring experts to use a finer classification scheme has the opposite of the intended effect.

JEL codes: D01.

Keywords: stochastic choice, imprecise perception.

1 Introduction

Precision captures the extent to which a decision-maker is able to discriminate between the perceived values of the alternatives available to him. A more precise agent can discriminate between values that a less precise agent lumps together. We are interested in the impact of precision on the quality of decision-making. Specifically: does increased precision enhance the probability of choosing the best alternative?

*We are grateful to Helios Herrera, Ian Jewitt, Mathieu Marcoux and Chris Tyson for useful discussions and advice about this project. Horan’s participation was financially supported by FRQSC.
The idea that individual decision-makers find it difficult to distinguish among values is well-known in psychology, dating back to the work of Fechner [11] and Thurstone [25, 26]. It also has a long tradition in economics dating back to Georgescu-Roegen [12], Luce [16, 17] and Quandt [23]. More recently, experiments by Butler and Loomes [5, 6] document the presence of “imprecision intervals” in lottery evaluations: an agent may be able to state confidently that a lottery is worth somewhere between $3 and $5, but may balk at making a more precise evaluation. Moreover, as Bayrak and Hey [4] argue, the discrepancy between willingness to pay and willingness to accept may also result from imprecise valuation intervals—with the agent offering the low end of the interval when in the role of a buyer and accepting the high end when in the role of a seller.\(^1\) Collective decision-makers also face difficulties that lead to imprecise discrimination. Policy makers, for instance, must balance the conflicting values and beliefs of individuals in society, which may themselves be perceived coarsely.\(^2\)

When the values of the alternatives are deterministic, greater precision does benefit the agent under reasonable assumptions about behaviour. Suppose, for example, that the choice between two alternatives is random when their realised values cannot be discriminated. Under this assumption, an agent’s choices may get better (and will certainly never get worse) as he becomes more precise. If the values of the alternatives are so far apart that the agent can distinguish between them even at low levels of precision, or if they are so close that the agent cannot distinguish between them even at high levels of precision, then the agent chooses the better alternative with the same probability irrespective of precision. Between these extremes, there are intermediate situations where the agent who chooses at random switches to the better alternative when he becomes more precise.

In the more realistic setting that we consider, the values of the alternatives are noisy. Put differently, not only does the agent have imprecise powers to discriminate between perceived values but the mechanism that generates the perceived values is itself prone to error. Does the logic that increased precision improves choice extend to this setting?

To get some intuition about what might go wrong, consider a simple example where a referee is asked to rank two papers. The true quality of each paper is given by a value in \{1, ..., 5\}. A noisy but perfectly discriminating referee perceives these values with error. Suppose that the true quality of paper “A” is 5 but, with probability 0.1, the referee wrongly perceives it as 2. In turn, paper “B” has true quality 3, which the referee always perceives correctly. Then paper “A” is correctly reported as being the better paper with

\(^1\)See also Tyson [27], who presents various economic applications of a parametric model of stochastic perception of utility differences. In this model, the probability that a given utility difference is perceived decreases exponentially with the size of the difference.

\(^2\)Danan et al. [8] discuss these issues in the context of social decisions. They focus on imprecise beliefs and on the Pareto principle as a means of singling out decisions that are robust to imprecise beliefs.
probability 0.9; and “B” is incorrectly reported as better with probability 0.1.

Now, consider a second referee who cannot discriminate between values 2 and 3. Just like the first referee, he reports “A” to be the better paper with probability 0.9. The difference is that the second referee never reports “B” to be the better paper. Instead, he reports that the papers are indistinguishable with probability 0.1. Paradoxically, by heeding the advice of the more precise referee, the editor faces a larger risk of ranking the lower quality paper above the higher quality paper. We emphasize that this example depends on the interaction of random and imprecise values.

Decision-makers face these two different sources of imperfect information in a variety of choice environments. Consider, for instance, a decision-maker (like a policy maker, investor, journal editor or juror) who relies on advice from experts. Since the experts themselves rely on scientific evidence or technical knowledge (e.g., about climate change) that is genuinely uncertain, their evaluations are noisy. In addition, the experts can often make fine distinctions that they cannot convey precisely to the decision-maker (through reports or classifications of the options). Consequently, the relative values of alternatives implied by expert opinions are perceived coarsely by the decision-maker.

Similar considerations arise in the context of voting. A politician’s value is generally signalled imperfectly (e.g., recent events may cast a disproportionately positive light on a politician with qualities specifically suited to deal with those events). In turn, an ill-informed voter might find it difficult to distinguish between candidates (thinking politicians are “all the same”) while an informed voter might be able to spot crucial differences.

We present a model which captures decision-making situations like these. In our model, precision is a numerical discrimination threshold à la Luce [16]; and we derive comparative statics for marginal changes in precision. While this infinitesimal approach does not change the logic of “harmful” precision from the discrete example above, it simplifies the analysis. In broad strokes, our results may be summarized as follows:

(1) In general, precision may be harmful. In fact, it is only unambiguously beneficial when one alternative is superior to the other in a strong distributional sense.

(2) Under some natural restrictions, whether precision is harmful or beneficial only depends on simple statistics (the mean, median and mode) of the value distributions.

(3) Finally, there are circumscribed, but economically relevant, circumstances (e.g., involving symmetric or identically distributed errors) where precision is broadly beneficial.
2 Overview

2.1 The model

An agent faces the choice between two alternatives $i$ and $j$ whose uncertain values are represented by random variables $u_i$ and $u_j$. Our interpretation is that the variability is due to perception errors rather than taste shocks (this latter case will be examined in section 5.3). What is more, the agent can only discriminate between realisations of $u_i$ and $u_j$ that are sufficiently far apart: a larger value is perceived as such only when it exceeds the lower value by a fixed $\sigma \geq 0$. We interpret $\sigma$ as a discrimination threshold, which measures the agent’s level of imprecision.

When the agent can discriminate between the two values, he chooses the alternative with the higher realized value. Otherwise, he chooses each alternative with equal probability.\(^3\) When the level of imprecision is $\sigma$, the probability of choosing $i$ is then

\[
p(i, \sigma) := \Pr(u_i > u_j + \sigma) + \frac{1}{2} \Pr(\sigma \geq |u_i - u_j|)
\]

\[
= \frac{1}{2} + \frac{1}{2} \left[ \Pr(u_i > u_j + \sigma) - \Pr(u_j > u_i + \sigma) \right].
\]  

(1)

In the sequel, we identify the quality of a decision for a given level of imprecision $\sigma$ with the probability of choosing the “better” alternative (i.e., $p(i, \sigma)$ when $i$ is better).

Formally, our approach grafts a random utility structure onto Luce’s \cite{Luce} deterministic semiorder model. As such, we maintain the assumption—central to the random utility model—that the agent chooses the alternative with the highest (perceived) utility realisation. This differs from the approach recently taken in Natzenz\cite{Natzenz}, where the agent treats the utility realisations as signals used to update a prior.

In the most general case that we consider, we permit any (continuous) distribution and any pattern of correlation between the values of the alternatives. Let $F$ denote the joint cumulative distribution function (henceforth cdf) of the values $u_i, u_j \in \mathbb{R}$ (so that $F(w, z) := \Pr(u_i \leq w, u_j \leq z)$ for all $w, z \in \mathbb{R}$). For simplicity, we assume that there exists a corresponding joint density function $f$ (unless otherwise specified).\(^4\)

For the corresponding value difference of the random variables $u_i - u_j$, we let $f_{u_i-u_j}$ denote the density and $F_{u_i-u_j}$ the cdf, so that (1) can also be written as

\[
p(i, \sigma) = \frac{1}{2} + \frac{1}{2} \left[ F_{u_j-u_i} (\sigma) - F_{u_i-u_j} (\sigma) \right].
\]  

(2)

\(^3\)As explained in footnote 30, our analysis easily extends to other tie-breaking rules. Our approach also covers some other decision procedures. One example is where the agent allocates the residual probability using an exogenous heuristic. Then, $p(i, \sigma) - p(j, \sigma)$ refers to the difference in the choice probabilities allocated via value comparisons (rather than to the total difference in choice probabilities).

\(^4\)Since much of our analysis focuses on unimodal distributions, this assumption is not overly restrictive. Such distributions (are absolutely continuous and) have a density at all points except possibly the mode.
To obtain an explicit formula for the density \( f_{u_i-u_j} \), note that the equality \( u_i - u_j = x \) simply represents the event consisting of all instances where \( u_j \) realises at some \( z \) (resp. \( \hat{z} - x \)) and \( u_i \) realises at \( \hat{z} + x \) (resp. \( \hat{z} \)). Therefore, integrating over these events:

\[
f_{u_i-u_j} (x) = \int_{\mathbb{R}} f (z + x, z) \, dz = \int_{\mathbb{R}} f (z, z - x) \, dz \quad \text{for all } x \in \mathbb{R}. \tag{3}
\]

To remain fairly agnostic about the nature of the agent’s errors (i.e., the relationship between the “true” value of an alternative and realised values), we consider two plausible scenarios about which alternative is better. In the first scenario, the better alternative is the one that is more likely to be chosen when the agent is “standard” (in the sense that he has perfect precision \( \sigma = 0 \)). In the second, the better alternative is the one that provides a higher expected value when the agent is standard.

**Definition 1.** For alternatives \( i \) and \( j \) with random values \( u_i \) and \( u_j \):

(i) \( i \) is median-better than \( j \) if \( p(i, 0) > \frac{1}{2} \). Equivalently, \( m_{u_i-u_j} > 0 \), where the median \( m_{u_i-u_j} \) is the value \( m \) that solves \( \int_{-\infty}^{m} f_{u_i-u_j} (z) \, dz = \frac{1}{2} = \int_{m}^{\infty} f_{u_i-u_j} (z) \, dz \).

(ii) \( i \) is mean-better than \( j \) if \( \mathbb{E}(u_i - u_j) > 0 \). Equivalently, \( \int_{\mathbb{R}} z f_{u_i-u_j} (z) \, dz > 0 \).

By substituting weak inequalities into these definitions, one obtains weak versions of each betterness notion. When alternative \( i \) is (weakly) better than alternative \( j \) according to either notion of betterness, we simply say that \( i \) is (weakly) better than \( j \).

In closing, we emphasize that median-betterness is based on choice probabilities, which are observable in principle (by identifying probabilities with choice frequencies) as long as the standard case of perfect precision can be identified. In contrast, mean-betterness is defined in terms of value distributions, which are not directly observable.

### 2.2 Some examples

To illustrate that the quality of decisions need not increase with precision in our model, we first consider a simple example where the probability is concentrated at just two points:

**Example 1. (Discrete distribution)** Suppose that the value pair \( u = (u_i, u_j) \) realises at \((10, 1)\) with probability \( \frac{3}{4} \), and at \((1, 2)\) with probability \( \frac{1}{4} \). In this case, \( i \) is the better alternative.\(^7\) When \( \sigma = 1 \), the worse alternative \( j \) is chosen with probability \( \frac{1}{2} \) \( \Pr (1 \geq |u_i - u_j|) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \). When the level of imprecision decreases to \( \sigma = 1 - \epsilon \) for arbitrarily small \( \epsilon > 0 \), the probability of choosing \( j \) increases to \( \Pr (u_2 > u_1 + 1 - \epsilon) = \frac{1}{4} \) (since \( j \) is now chosen outright when \((1, 2)\) realises). Accordingly, increased precision harms the agent when \( \sigma = 1 \).

---

\(^5\)This is the classical view (see e.g. Quandt [23]) that “long run” frequencies reflect preference.

\(^6\)When the mean is undefined (as it is in the case of Cauchy distributions, for instance), we say that \( i \) is mean-better than \( j \) if the Cauchy principal value \( \lim_{x \to \infty} \int_{-\infty}^{x} z f_{u_i-u_j} (z) \, dz \) is strictly positive.

\(^7\)Simple computation shows that \( p(i, 0) = 3/4 > 1/2 \) and \( \mathbb{E}(u_i - u_j) = \mathbb{E}(u_i) - \mathbb{E}(u_j) = 26/4 > 0 \).
In this example, a decrease in precision at \( \sigma = 1 - \varepsilon \) makes the value difference imperceptible in the event where the worse alternative is chosen. Since the change in precision is too small to obscure the value difference in the event where the better alternative is chosen, the overall effect is to increase the probability of choosing the better alternative.

For ease of exposition, we assumed the value realisations to be correlated. However, correlation actually plays no role in the effect. Indeed, it is easy to modify the example so that the values are independent but higher precision remains harmful.\(^8\)

By no means is Example 1 the end of the story. For some distributions commonly used in applications, it turns out that precision \textit{unambiguously} improves decisions:

**Example 2. (Logit distribution)** Suppose that \( u_i = \hat{u}_i + \varepsilon_i \) and \( u_j = \hat{u}_j + \varepsilon_j \) where \( \hat{u}_i, \hat{u}_j \in \mathbb{R} \) and the random errors \( \varepsilon_i, \varepsilon_j \) are i.i.d. Gumbel with location \( \nu = 0 \) and scale \( c = 1.\(^9\) Then, \( u_i - u_j \) is logistic with location \( \nu = \hat{u}_i - \hat{u}_j \) and scale \( c = 1 \). (While it is well-known, we derive this fact in Appendix A.) For a given level of imprecision \( \sigma \), it then follows that \( i \) “beats” \( j \) with probability

\[
\Pr \left( u_i > u_j + \sigma \right) = \frac{e^{\hat{u}_i}}{e^{\hat{u}_i + \sigma} + e^{\hat{u}_j}}.
\]

The formula for \( \Pr \left( u_j > u_i + \sigma \right) \) is symmetric. From equation (2), it then follows that

\[
p(i, \sigma) = \frac{1}{2} + \frac{1}{2} \left( \frac{e^{\hat{u}_i}}{e^{\hat{u}_i + \sigma} + e^{\hat{u}_j}} - \frac{e^{\hat{u}_j}}{e^{\hat{u}_i + \sigma} + e^{\hat{u}_j}} \right).
\]

Evaluating at \( \sigma = 0 \) shows that \( i \) is the median-better alternative if and only if \( \hat{u}_i > \hat{u}_j \).

Since the mean of a logistic distribution corresponds to its location, the same condition describes the circumstances where \( i \) is the mean-better alternative. From the formula for \( p(i, \sigma) \), the marginal effect of imprecision is then

\[
\frac{\partial p(i, \sigma)}{\partial \sigma} = \frac{1}{2} \left( \frac{e^{\hat{u}_i + \hat{u}_j + \sigma}}{(e^{\hat{u}_i + \sigma} + e^{\hat{u}_j})^2} - \frac{e^{\hat{u}_i + \hat{u}_j + \sigma}}{(e^{\hat{u}_j + \sigma} + e^{\hat{u}_i})^2} \right).
\]

This shows that a marginal decrease in precision has (one of) two possible effects. For an agent who is already imprecise (\( \hat{\sigma} > 0 \)), it reduces the quality of decision-making:

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \hat{\sigma}} < 0 \iff \hat{u}_i > \hat{u}_j.
\]

In contrast, it has no effect for an agent who is already perfectly precise (\( \hat{\sigma} = 0 \):

\[
\frac{\partial p(i, 0)}{\partial \sigma} = 0.
\]

\(^8\)To illustrate, suppose \( u_i = (10, 1) \) and \( u_j = (1, 2) \) realise independently with probabilities \( (3/4, 1/4) \). Just as in the example, \( i \) is better (since \( p(i, 0) = 17/32 > 1/2 \); and \( E(u_i - u_j) = 131/16 > 0 \) and precision harms at \( \sigma = 1 - \varepsilon \) (since \( p(i, 1) = \Pr (u_i = 10) + [\Pr (u_i = 1 \& u_j = 1) + \Pr (u_i = 1 \& u_j = 2)]/2 > \Pr (u_i = 10) + \Pr (u_i = 1 \& u_j = 1)/2 = p(i, 1 - \varepsilon) \)).

\(^9\)For a Gumbel with location \( \mu \) and scale \( c \), \( F(z) = e^{-e^{-\frac{z}{c}}} \) and \( f(z) = \frac{1}{c} e^{-e^{-\frac{z}{c}}} e^{-\frac{z}{c}} \).
Our third example shows how a minor change alters these conclusions dramatically:

**Example 3. (Scaled errors)** As in Example 2, suppose that that \( u_i = \hat{u}_i + \varepsilon_i \) and \( u_j = \hat{u}_j + \varepsilon_j \) and that the random errors \( \varepsilon_i, \varepsilon_j \) are i.i.d. Gumbel with location zero. The only difference is that \( \varepsilon_i \) is now scaled by a factor \( c > 1^{10} \).

The critical difference from Example 2 is the fact that \( u_i - u_j \) is not logistic when \( c > 1 \). While the distribution lacks a simple closed form expression, it is not difficult to show that the effect of the scaling factor is to skew the logistic distribution in Example 2 towards the right. (In Appendix A, we derive an integral representation for the cdf, from which the choice probabilities can be computed using equation (2).)

This has two implications for our analysis: first, it pushes the mean above the median; and, as a related matter, it fattens the right tail of the distribution relative to the left. The first change has the potential to drive a wedge between our two notions of betterness. In turn, the second change creates the possibility that precision has the opposite effect for a very imprecise agent (\( \sigma \to \infty \)) as it does for a very precise agent (\( \sigma \to 0 \)).

To illustrate these observations, let us suppose that \( 0 = \hat{u}_i < \hat{u}_j = 1/2 \) and \( c = 2 \). The density for this parametrization is shown in Figure 1 below.

![Figure 1: Plot of \( f_{u_i-u_j} \) for \( u_i \sim Gumbel(0,2) \) and \( u_j \sim Gumbel(1/2,1) \)](image)

By numerical calculation, it is not difficult to establish the following facts:

\(^{10}\)Scaling a random variable \( \varepsilon \) by \( c \in \mathbb{R}_+^+ \) gives a random variable \( \varepsilon' \) that is distributed like \( c\varepsilon \). When \( \varepsilon \) is Gumbel with location zero and unit scale, it follows that the cdf of \( \varepsilon' \) is \( F(z) = e^{-e^{-z/c}} \).
(i) While \( j \) is the better alternative before scaling (since the median and mean value differences \( u_i - u_j \) are both \(-1/2\)), this is no longer true after \( u_i \) is scaled. While the median and mean both increase, \( m_{u_i - u_j} \approx -0.211 \) remains negative while \( E(u_i - u_j) \approx 0.078 \) becomes positive. So, alternative \( i \) is both mean-better and median-worse than \( j \).

(ii) At the same time, the impact of precision changes sharply at \( \bar{\sigma} \approx 4.244 \). Below this cutoff level, more precision decreases the probability \( p(i, \sigma) \) of selecting alternative \( i \); and, above this level, it has the opposite effect.

From these observations, it follows that increased precision: (1) harms the mean-better alternative \( i \) and helps the median-better alternative \( j \) for sufficiently small levels of \( \sigma \); and (2) has the opposite effect for sufficiently large levels of \( \sigma \). From the standpoint of a policy maker (charged with choosing a desired level of precision), these conclusions are striking. They show that the marginal benefit of precision does not necessarily bear any systematic relationship to the agent’s (initially positive) level of imprecision and may depend critically on the relevant notion of betterness.

Taken together, our three examples beg the question: what feature of the value distributions ensure the “intuitive” effect that precision improves the quality of decisions? After deriving the fundamental condition that governs the marginal impact of precision, we focus on independent and unimodal value distributions, restrictions which are satisfied in many common applications of the random utility model. While a discrepancy between precision and quality persists under these restrictions, they allow us to characterise the relationship in terms of primitive features of the value distributions. In turn, this dependence makes it easy to do comparative statics.

### 3 General Analysis

We start the analysis by considering the most general case. From (2), it follows that:

\[
p(i, \sigma) = \frac{1}{2} + \frac{1}{\sqrt{2}} \left( \int_R \int_{z+\sigma} f(w, z) dw dz - \int_R \int_{w+\sigma} f(w, z) dz dw \right).
\]

Differentiating this expression and evaluating at the level of imprecision \( \sigma = \hat{\sigma} \) yields

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \sigma} = \frac{1}{\sqrt{2}} \left( \int_R f(w, w + \hat{\sigma}) dw - \int_R f(z + \hat{\sigma}, z) dz \right). \tag{4}
\]

It then follows that

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \sigma} < 0 \iff \int_R f(w, w + \hat{\sigma}) dw < \int_R f(z + \hat{\sigma}, z) dz. \tag{5}
\]

Using (3), condition (5) can then be written more compactly as

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \sigma} < 0 \iff f_{u_i - u_j} (-\hat{\sigma}) < f_{u_i - u_j} (\hat{\sigma}). \tag{\star}
\]
Condition (⋆) is the fundamental inequality governing the effect of precision. It shows that the marginal impact at \( \hat{\sigma} \) depends on local comparisons of the value difference \( u_i - u_j \). The only events that matter are those where the value difference exactly matches the level of imprecision \( \hat{\sigma} \). In these “threshold events”, the probability of choosing alternative \( i \) changes by one half, either positively when the agent stops perceiving \( j \) as better (at \( u_j - u_i = \hat{\sigma} \)); or negatively when he stops perceiving \( i \) as better (at \( u_i - u_j = \hat{\sigma} \)). Overall, the marginal impact on the probability of choosing alternative \( i \) is one-half times the probability difference between these threshold events.\(^{11}\)

Condition (⋆) has several notable consequences. The first is that the impact of precision depends on the cross-correlation in the sense of signal-processing between \( u_i \) and \( u_j \) when \( u_i \) is displaced by a “lag” or “lead” of \( \hat{\sigma} \) (and it is unrelated to the statistical correlation between \( u_i \) and \( u_j \)). As it turns out, this is the source of a systematic disconnect between condition (⋆) and the measures of quality described in Definition 1. While, as noted, the effects of increased precision are driven by local features of the value distributions, the quality of the alternatives depend on global features of these distributions. The following calibration result puts this point in the starkest possible terms:

**Proposition 1.** For every level of imprecision \( \hat{\sigma} > 0 \) and for all parameter values \( \hat{m}, \hat{\mu}, \hat{p} \in \mathbb{R}_+ \), there exists a density \( f_{u_i - u_j} \) such that \( \mathbb{E}(u_i - u_j) \geq \hat{\mu}, m_{u_i - u_j} \geq \hat{m} \) and \( \frac{\partial p(i, \hat{\sigma})}{\partial \sigma} \geq \hat{p} \).

By condition (⋆), the sign of the derivative \( \frac{\partial p(i, \hat{\sigma})}{\partial \sigma} \) is pinned down by the density of the value difference at exactly two points. So, \( f_{u_i - u_j} \) is effectively unconstrained.

Proposition 1 shows that precision may harm unboundedly at a given level of imprecision: regardless of how much better alternative \( i \) is than alternative \( j \), there is some distribution of value differences for which the marginal decrease in the probability of choosing \( i \) exceeds a given threshold \( \hat{p} \).

A second and complementary implication of condition (⋆) is that precision cannot always harm for a given distribution of value differences:

**Proposition 2.** If \( i \) is mean-better or median-better, then \( \frac{\partial p(i, \hat{\sigma})}{\partial \sigma} < 0 \) for some \( \hat{\sigma} > 0 \).

By condition (⋆), \( \frac{\partial p(i, \hat{\sigma})}{\partial \sigma} \geq 0 \) for all \( \hat{\sigma} > 0 \) implies \( f_{u_i - u_j}(-z) \geq f_{u_i - u_j}(z) \) for all \( z > 0 \). By integrating over these inequalities, we make the following observations:

(i) \( \int_{-\infty}^{0} f_{u_i - u_j}(z) \, dz \geq \int_{0}^{\infty} f_{u_i - u_j}(z) \, dz \); and
(ii) \( \int_{\mathbb{R}} f_{u_i - u_j}(z) \, dz = \int_{0}^{\infty} z \left[ f_{u_i - u_j}(z) - f_{u_i - u_j}(-z) \right] \, dz \leq 0 \).

Observation (i) shows that \( j \) is (weakly) median-better, while observation (ii) says that \( j \) is (weakly) mean-better. By contraposition, they imply the desired result.

\(^{11}\)For discrete increases in \( \sigma \), the set of threshold events includes all those events where the value difference is (strictly) greater than the initial level of \( \sigma \) but (weakly) less than the new level of \( \sigma \).
The basis of observations (i) and (ii) is another consequence of condition \((\ast)\) which is important enough to highlight separately. In particular, since \(f_{u_i-u_j}(z) = f_{u_j-u_i}(-z)\), condition \((\ast)\) implies the following:

**Proposition 3.** \(\frac{\partial p(i, \hat{\sigma})}{\partial \sigma} \leq 0\) for almost every level of imprecision \(\hat{\sigma} \geq 0\) if and only if \(f_{u_i-u_j}(z) \geq f_{u_j-u_i}(z)\) for almost every value difference \(z \geq 0\).

To interpret this result, suppose that \(i\) is the better alternative. Then, for precision to have an unambiguously beneficial impact, \(u_i\) must display a strong form of distributional dominance over \(u_j\). Not only must \(u_i\) beat \(u_j\) “on average” in the sense that

\[
\int_{\mathbb{R}} zf_{u_i-u_j}(z) \, dz \geq \int_{\mathbb{R}} zf_{u_j-u_i}(z) \, dz
\]

but \(u_i\) must also beat \(u_j\) “point-wise” in the sense that, for almost every \(z \in \mathbb{R}\),

\[
z f_{u_i-u_j}(z) \geq z f_{u_j-u_i}(z).
\]

Clearly, this type of point-wise dominance is stronger than first-order stochastic dominance between the value difference distributions \(u_i-u_j\) and \(u_j-u_i\). In fact, it implies that \(f_{u_i-u_j}\) and \(f_{u_j-u_i}\) cross exactly once at \(z = 0\), which in turn entails that \(u_i-u_j\) first-order stochastically dominates \(u_j-u_i\). From another angle, point-wise dominance may also be viewed as a strong form of skewness of \(f_{u_i-u_j}\) relative to zero.\(^{12,13}\)

To make further progress in clarifying what drives the connection between quality and precision in applications, we will focus on some significant (but economically relevant) restrictions on value distributions. Even under these restrictions, the lack of connection between quality and precision persists.

### 4 Unimodality and independence

Our aim in this section is to express more concretely, in terms of familiar summary statistics of the value distributions, the relation between quality and increased precision analysed at a general level in the previous section. With this in mind, we consider two restrictions on value distributions: unimodality and independence. A real-valued random variable \(X\) with cdf \(F\) is (strictly) unimodal around \(\nu \in \mathbb{R}\) (and \(\nu\) is a mode of \(X\)) if \(F\) is (strictly) convex on \((-\infty, \nu)\) and (strictly) concave on \((\nu, \infty)\). Since we assume that \(F\) can be associated with a density \(f\), this is equivalent to \(f\) (increasing) non-decreasing on

\(^{12}\)We thank Ian Jewitt for making this observation (in private communication).

\(^{13}\)Macgillivray [18] considers a similar notion of skewness relative to the mean \(\mu_X\) of a unimodal random variable \(X\) (see the next section for definitions). He shows that the skewness measure \(\text{E}(X-\mu_X)^3\) is strictly positive if the difference \(f_X(\mu_X + z) - f_X(\mu_X - z)\) changes signs exactly once for \(z \geq 0\).
\(-\infty, \nu\) and (decreasing) non-increasing on \((\nu, \infty)\). When this condition holds for some \(\nu\), we refer to the density \(f\) as unimodal.

It follows from the definition that a strictly unimodal distribution has a single mode. More generally, for a unimodal distribution, the set of modes forms a closed interval with minimal and maximal modes denoted by \(\nu^{\text{min}}\) and \(\nu^{\text{max}}\). For convenience, we also denote the central mode of a unimodal distribution by \(\bar{\nu} := (\nu^{\text{max}} + \nu^{\text{min}})/2\).

### 4.1 Background results

Our characterisation results in the next section require the distribution of the value difference \(u_i - u_j\) to be unimodal. However, for economic applications, it is more practical to express this requirement in terms of assumptions on the primitive distributions \(u_i\) and \(u_j\). In what follows, we adapt some classical results from statistics that allow us to make a connection between the two types of conditions. While many of these results are quite well-known in statistics, they do not seem to be widely known in economics.

The starting point of analysis is a classical result due to Ibragimov [15]. Recall that the convolution \(h \ast g\) of functions \(g, h : \mathbb{R} \to \mathbb{R}\) is defined by
\[
[h \ast g](x) := \int_{\mathbb{R}} h(z) g(x - z) \, dz \quad \text{for all } x \in \mathbb{R}.
\]

It is clear from the definition that convolution has some nice algebraic properties—including that it is commutative, associative and commutes with translation.\(^{14}\)

Ibragimov showed that unimodality (of two functions \(g\) and \(h\)) is preserved under convolution provided that one of the two functions is log-concave.\(^{15}\) Recall that a function \(f\) is log-concave if and only if its log-transformation is concave.\(^{16}\) This condition, which implies unimodality, is satisfied by the densities of many distributions commonly used in applications—including the normal, Gumbel, uniform, exponential, logistic, Chi-squared (with scale parameter \(c \geq 2\)), Gamma (with scale parameter \(c \geq 1\)) and Laplace distributions (see e.g., Table 1 in Bagnoli and Bergstrom [2]).

If \(f_i\) and \(f_j\) denote the (independent) densities of \(u_i\) and \(u_j\), then (3) simplifies to
\[
f_{u_i - u_j}(x) = \int_{\mathbb{R}} f_i(z) f_j(z - x) \, dz \quad \text{for all } x \in \mathbb{R}.
\]\(^{6}\)

Since it will be useful in the sequel, we denote the reflection of a function \(f : \mathbb{R}^n \to \mathbb{R}^n\) through the vertical origin by \(\bar{f}\). Formally, \(\bar{f}(y) := f(-y)\) for all \(y \in \mathbb{R}^n\). If \(f = \bar{f}\), then \(f\) is said to be symmetric around the origin (or zero in the case where \(n = 1\)).\(^{17}\)

---

\(^{14}\)Formally: \(h \ast g = g \ast h, h \ast (g \ast f) = (h \ast g) \ast f\) and \((\tau_c h) \ast g = \tau_c (h \ast g)\). Recall that the translation of a function \(f : \mathbb{R}^n \to \mathbb{R}^n\) by a vector \(c \in \mathbb{R}^n\) is given by \([\tau_c f](x) := f(x + c)\) for all \(x \in \mathbb{R}\).

\(^{15}\)Previously, it had been wrongly conjectured that unimodality was closed under convolution.

\(^{16}\)In other words, \(f(\lambda x + (1 - \lambda) y) \geq [f(x)]^{\lambda} [f(y)]^{1-\lambda}\) for all \(\lambda \in [0, 1]\).

\(^{17}\)Symmetry of \(f\) around \(c \in \mathbb{R}^n\) amounts to the symmetry of the translation \(\tau_{-c} f\) around the origin.
By using this notation and the definition of a convolution, (6) simplifies to

\[ f_{u_i - u_j}(x) = [f_i \ast \bar{f}_j](x). \]  \hspace{1cm} (7)

Since it is clear that reflection preserves the relevant properties of \( f_j \) (i.e., unimodality or log-concavity), equation (7) allows us to restate Ibragimov’s result as follows:

**Lemma 1. (Ibragimov)** If \( u_i \) and \( u_j \) are unimodal, independently distributed random variables and at least one of the two is log-concave, then \( f_{u_i - u_j} \) is unimodal.

Before Ibragimov, Wintner [28] showed that unimodality is preserved under convolution provided that both of the functions are symmetric.\(^{18}\) Since reflection preserves the symmetry of \( f_j \), this result and (7) establish the first part of the next lemma.

**Lemma 2. (Wintner)** If \( u_i \) and \( u_j \) are symmetric, unimodal and independently distributed random variables, then \( f_{u_i - u_j} \) is unimodal. What is more, if \( u_i \) and \( u_j \) are symmetric around \( \bar{\nu}_i \) and \( \bar{\nu}_j \) respectively, then \( f_{u_i - u_j} \) is symmetric around \( \bar{\nu}_i - \bar{\nu}_j \).

For the second part of the lemma, suppose \( \bar{\nu}_i = 0 = \bar{\nu}_j \). (This is without loss since one can “de-mode” \( u_i \) and \( u_j \) when the central modes are non-zero.\(^{19}\) Then,

\[ f_{u_i - u_j}(x) = [f_i \ast \bar{f}_j](x) = [\bar{f}_i \ast f_j](x) = [f_j \ast \bar{f}_i](x) = [\bar{f}_i \ast f_j](x) = \bar{f}_{u_i - u_j}(x) \]

where: the second equality follows by symmetry; the third by commutativity; the fourth by the fact that reflection commutes with convolution;\(^{20}\) and the last from the definition of \( f_{u_i - u_j} \). Since this equality holds for all \( x \in \mathbb{R} \), \( f_{u_i - u_j} \) is symmetric around zero.

A third important result about unimodality is Hodges and Lehmann’s [14] observation that the convolution of a unimodal function \( f \) with its reflection \( \bar{f} \) is unimodal.\(^{21}\) (In the statistics literature, the convolution \( f \ast \bar{f} \) is known as the symmetrization of \( f \).) Using (7), it is possible to translate their result into our framework as follows.\(^{22}\)

**Lemma 3. (Hodges and Lehmann)** If \( u_i \) and \( u_j \) are unimodal i.i.d. random variables, then \( f_{u_i - u_j} \) is unimodal and symmetric around zero.

### 4.2 Characterising results

We are now ready to demonstrate the central conclusion of this section: as long as the value distributions \( u_i \) and \( u_j \) satisfy the assumptions associated with (any of) the three lemmas

\(^{18}\)For a concise and recent treatment of this result, see Purkayastha [22] (Theorem 2.1).

\(^{19}\)In particular, \((\tau_{-\bar{\nu}_i} f_i) \ast (\tau_{-\bar{\nu}_j} \bar{f}_j) = (\tau_{-\bar{\nu}_i} f_i) \ast (\tau_{-\bar{\nu}_j} \bar{f}_j) = (\tau_{-\bar{\nu}_i} \cdot \tau_{-\bar{\nu}_j})[f_i \ast \bar{f}_j] = \tau_{-\bar{\nu}_i-\bar{\nu}_j}[f_i \ast \bar{f}_j].\)

\(^{20}\)The fact that reflection commutes with convolution means that \( h \ast g = \bar{h} \ast \bar{g} \).

\(^{21}\)See also Purkayastha [22] (Theorem 2.2) or Dharmadhikari and Joag-Dev [10] (Theorem 1.8).

\(^{22}\)Technically, the Hodges and Lehmann result only establishes the unimodality of \( f_{u_i - u_j} \) around zero. However, symmetry follows directly from the same kind of argument given after Lemma 2.
from the previous section, the marginal impact of precision can be described succinctly in
terms of the modal values of the value difference distribution. This simplifies considerably
the task of checking condition (\(\star\)) for an infinite number of values of \(\sigma\).

To start, we define a critical upper bound on the level of imprecision:

**Definition 2.** For a unimodal value difference distribution \(f_{u_i-u_j}\), the level of imprecision
\(\hat{\sigma} \geq 0\) is non-confounding if \(\hat{\sigma} < \max (|\nu^{\min}|, |\nu^{\max}|)\).

If the level of imprecision is non-confounding, then the agent can perceive which alternative is better at the extreme modes. (If \(f_{u_i-u_j}\) realises at \(\nu > 0\), then \(\hat{u}_i = \hat{u}_j + \nu > \hat{u}_j + \hat{\sigma}\).

So, the agent perceives \(\hat{u}_i > \hat{u}_j\). Similar reasoning applies when \(\nu < 0\).)

In the next result we show that, when the distribution of the value difference is uni-
modal and symmetric (as guaranteed by the conditions in Lemmas 2 and 3), the marginal impact of precision is determined by the sign of the central mode \(\nu\). When the distribution of the value difference is unimodal but not necessarily symmetric (which is guaranteed by the condition in Lemma 1), this is only true for non-confounding levels of imprecision.

**Proposition 4.** Suppose that \(f_{u_i-u_j}\) is unimodal with central mode \(\nu \geq 0\). Then:

(i) \(\frac{\partial p(i,\nu)}{\partial \sigma} \leq 0\) for every non-confounding level of imprecision \(\hat{\sigma}\); and

(ii) if \(f_{u_i-u_j}\) is symmetric, then \(\frac{\partial p(i,\nu)}{\partial \sigma} \leq 0\) for every level of imprecision \(\hat{\sigma}\).

What is more: if \(\nu > 0\), then there exists some non-confounding \(\hat{\sigma}\) such that \(\frac{\partial p(i,\nu)}{\partial \sigma} < 0\).

**Proof:** (i) Since \(\nu \geq 0\), \(\nu^{\max} \geq \nu^{\min}\). Fix a non-confounding \(\hat{\sigma}\). Since \(\hat{\sigma} \in [0,\nu^{\max}]\) and \(f_{u_i-u_j}\) is unimodal, \(f_{u_i-u_j}(-\hat{\sigma}) \leq f_{u_i-u_j}(\hat{\sigma})\). So, \(\partial p(i,\hat{\sigma})/\partial \sigma \leq 0\) by (\(\star\)).

(ii) If \(\nu \in [0,\nu^{\max}]\), then the argument in (i) implies \(\partial p(i,\hat{\sigma})/\partial \sigma \leq 0\). So, suppose \(\hat{\sigma} \geq \nu^{\max}\). By symmetry around \(\nu\), \(f_{u_i-u_j}(-\hat{\sigma}) = f_{u_i-u_j}(\hat{\sigma} + 2\nu)\). Since \(\nu \geq 0\), \(f_{u_i-u_j}(-\hat{\sigma}) = f_{u_i-u_j}(\hat{\sigma} + 2\nu) \leq f_{u_i-u_j}(\hat{\sigma})\) by unimodality. So, \(\partial p(i,\hat{\sigma})/\partial \sigma \leq 0\) by (\(\star\)).

For the last part of the statement, note that \(\nu > 0\) implies \(\nu^{\max} > \nu^{\min}\). Pick some \(\sigma^* \in (|\nu^{\min}|, |\nu^{\max}|)\). Since \(f_{u_i-u_j}\) is unimodal, \(f_{u_i-u_j}(\sigma^*) > f_{u_i-u_j}(-\sigma^*)\) which, in turn, gives \(\partial p(i,\sigma^*)/\partial \sigma < 0\) by (\(\star\)).

Proposition 4 does not ensure that the sign of the central mode completely determines
the sign of the marginal impact. While \(\partial p(i,\hat{\sigma})/\partial \sigma = 0\) holds (for all relevant levels of imprecision) when \(\nu = 0\), the same equality may also hold for certain levels of imprecision
when \(\nu > 0\). However, if the value difference is strictly unimodal, a tighter characterisation obtains. In that case, the sign of the central mode is determinative:

**Corollary 1.** Suppose that \(f_{u_i-u_j}\) is strictly unimodal with mode \(\nu\). Then:

(i) for every non-confounding level of imprecision \(\hat{\sigma} > 0\), \(\frac{\partial p(i,\hat{\sigma})}{\partial \sigma} < 0 \iff \nu > 0\); and

(ii) if \(f_{u_i-u_j}\) is symmetric, then the same holds for every level of imprecision \(\hat{\sigma} > 0\).
Just like the results in Section 3, the results in this section draw a sharp distinction between the quality of an alternative and the marginal effect of precision. Whereas the mean or the median of the value difference determine quality, the mode determines the effect of precision.\footnote{Complicating matters is the fact that unimodality imposes no restrictions on the ordering of the mean, the median, and the mode (contrary to the common misconception that the median must fall between the other two measures of centrality).} Complicating matters is the fact that unimodality imposes no restrictions on the ordering of the mean, the median, and the mode when combined with unimodality and independence, guaranteeing that the better alternative is chosen with higher probability as precision increases. The first type of condition (see Propositions 5 and 6) relates to the way the noise is distributed but imposes no limits on the amount of noise. In contrast, the second type of condition (see Proposition 7) limits the total amount of noise, but otherwise places no restrictions on its distribution.

The first result establishes that precision increases the chance of selecting the better alternative when the errors in the value distributions are identically distributed.

**Proposition 5. (Identical errors)** Suppose \( u_i := \hat{u}_i + \varepsilon_i \) and \( u_j := \hat{u}_j + \varepsilon_j \) where \( \hat{u}_i, \hat{u}_j \in \mathbb{R} \) and \( \varepsilon_i, \varepsilon_j \) are mean-zero, i.i.d., unimodal random variables. Then: \( \frac{\partial p(i,\hat{\sigma})}{\partial \sigma} \leq 0 \) for every level of imprecision \( \hat{\sigma} \iff i \) is weakly better.

**Proof:** By Lemma 3, \( \varepsilon_i - \varepsilon_j \) is unimodal and symmetric around zero. From the i.i.d. and mean-zero assumptions, it follows that the mean and median of \( \varepsilon_i - \varepsilon_j \) are also zero. So, \( u_i - u_j \) is unimodal and symmetric around the mean, median and modal value \( \nu = \hat{u}_i - \hat{u}_j \). The result then follows by Proposition 4.

This result covers all of the standard i.i.d. specifications used to model random utility, including logit (Gumbel) and probit (normal) errors. For the latter, it turns out that precision cannot harm even when the errors are not identical. More generally, as shown by the next result, the same holds true for all symmetric error distributions.

**Proposition 6. (Symmetric errors)** Suppose that \( u_i := \hat{u}_i + \varepsilon_i \) and \( u_j := \hat{u}_j + \varepsilon_j \) where \( \hat{u}_i, \hat{u}_j \in \mathbb{R} \) and \( \varepsilon_i, \varepsilon_j \) are independent random variables that are unimodal and symmetric around zero. Then: \( \frac{\partial p(i,\hat{\sigma})}{\partial \sigma} \leq 0 \) for every level of imprecision \( \hat{\sigma} \iff i \) is weakly better.

---

\footnote{In Example 3 \( f_i \) and \( f_j \) are both log-concave, so that by Lemma 1 the value difference density is unimodal, and we can apply the results of this section. For the scale \( c \) sufficiently close to 1, it can be shown that \( \nu_{u_i - u_j} < m_{u_i - u_j} \). Then Corollary 1 implies the harmful effect of precision noted there.}

\footnote{See Abadir [1] for some examples that the mean, the median, and the mode can fall in any order.}
Proof: By Lemma 2, the difference $\varepsilon_i - \varepsilon_j$ is unimodal and symmetric around zero. Then, the argument in the proof of Proposition 5 applies verbatim (with the Cauchy principal value standing in for the mean whenever the latter is undefined).

Among symmetric distributions, normal distributions have a “special” feature: the difference of independent normals is normal even when their standard deviations (or “scale” parameters) differ. More generally, all stable distributions that are symmetric around zero have this closure property. This class contains families of distributions indexed by a stability parameter $\alpha \in (0, 2]$. It includes normal distributions ($\alpha = 2$), Cauchy distributions ($\alpha = 1$) and a continuum of fat-tailed distributions ($1 < \alpha < 2$) between the two (see Mandelbrot [19] for some early applications of stable distributions in economics). Even when their scale parameters differ, the difference of two independent stable distributions (with stability parameter $\alpha$) is stable (with parameter $\alpha$).

We emphasize that Proposition 6 does not require anything like the requirement that the error distributions belong to the same family (or even the same “stability class”) of distributions. The result is much more general: as long as the error distributions are independent, unimodal and symmetric, increased precision cannot harm.

Our third result relies on a known bound for the relative locations of median, mean and mode in unimodal distributions. This bound depends on the variance of the distribution which, in our framework, can be interpreted as the total amount of noise in the perception of values. Provided that the total noise is sufficiently low and the level of precision is high enough to be non-confounding, the quality of decisions increases with precision.

Proposition 7. (Bounded noise) Suppose that $u_i$ and $u_j$ are independent and unimodal random variables and at least one of the two is log-concave. Then, there exists some $K > 0$ such that if $\text{Var}(u_i) + \text{Var}(u_j) < K$: $rac{\partial p_i}{\partial \sigma} \leq 0$ for every non-confounding level of imprecision $\hat{\sigma} \iff i$ is better.

Proof: Let $X := u_j - u_i$. Define $K := (\nu_{\text{max}})^2/3$ and suppose $\text{Var}(u_i) + \text{Var}(u_j) < K$. By construction, it follows that $s_X := \sqrt{\text{Var}(u_i) + \text{Var}(u_j)} < \sqrt{K} = |\nu_{\text{max}}| / \sqrt{3}$.

By Lemma 1, $X$ is unimodal. As such, every mode $\nu_X$ of $X$ is related to the median $m_X$, the mean $\mu_X$ and the standard deviation $s_X$ by the inequalities $|m_X - \nu_X| \leq \sqrt{3}s_X$ and $|\mu_X - \nu_X| \leq \sqrt{3}s_X$ (see e.g., Corollary 4 of Basu and DasGupta [3]). Since $s_X < |\nu_{\text{max}}| / \sqrt{3}$, these inequalities imply $|\nu_{\text{max}}| > |m_X - \nu_X|, |\mu_X - \nu_X|$. Under these restrictions, it is not difficult to show that $m_X > 0 \iff \nu_{\text{max}} > 0 \iff \mu_X > 0$.

25A stable random variable $X(\alpha; \beta, c, \nu)$ is defined by four parameters: stability $\alpha \in (0, 2]$; skewness $\beta \in [-1, 1]$; scale (i.e., standard deviation) $c \in (0, \infty)$; and location $\nu \in (-\infty, \infty)$. For $\beta = 0$, the distribution is symmetric and $\nu$ coincides with the mode and median (as well as the mean when $\alpha > 1$).

26In other words, $X(\alpha; 0, c_1, 0) - X(\alpha; 0, c_2, 0) = X(\alpha; 0, (c_1^\alpha + c_2^\alpha)^{1/\alpha}, 0)$. 

15
To complete the proof, we establish that $\nu_{\text{max}} > 0 \iff \partial p (i, \hat{\sigma}) / \partial \sigma \leq 0$ for every non-confounding $\hat{\sigma}$. If $\nu_{\text{max}} > 0$, then $|\nu_{\text{max}}| > |m_X - \nu_X|$ implies $\nu_{\text{max}} > |\nu_{\text{min}}|$. If $\hat{\sigma}$ is non-confounding, then $f_{u_i - u_j} (-\hat{\sigma}) \leq f_{u_i - u_j} (\hat{\sigma})$ by unimodality; and $\partial p (i, \hat{\sigma}) / \partial \sigma \leq 0$ by $(\ast)$. Conversely, if $\partial p (i, \hat{\sigma}) / \partial \sigma \leq 0$ for every non-confounding $\hat{\sigma}$, then $\nu_{\text{max}} \geq \bar{\nu} \geq 0$ by Proposition 4. So, $\nu_{\text{max}} > 0$. Otherwise, $\text{Var}(u_i) + \text{Var}(u_j) < \nu_{\text{max}} = 0$.

Propositions 5 to 7 identify distributional restrictions that are sufficient for increased precision to be unambiguously positive. While such restrictions are fairly common in economic models, they are not always appropriate. There are natural situations (e.g., Example 3) where the assumption of identically or symmetrically distributed errors are violated. Similarly, the assumption that the agent perceives the values with sufficiently low noise may be unwarranted (e.g., when he is unfamiliar with the alternatives).

5 Extensions

5.1 Beyond unimodality and independence

Even without unimodality or independence, there remains a link between the mode of the value difference and the impact of precision. In particular, Proposition 4 imposes no assumptions on the individual value distributions; and it may be applied even when they are dependent or multi-modal. The difficulty (already noted in Section 4.2) is that the mode of the value difference need not bear any relationship to the median or the mean. Despite this, it is still possible to make some general statements about the impact of precision on the quality of decision-making.

Our first result in this setting extends Proposition 6 by dispensing with independence. It establishes that precision cannot harm for a broad class of distributions introduced by Ghosh [13] (see also Dharmadhikari and Jogdeo [9]). Formally, a real-valued random vector $X := (X_1, X_2)$ is linear unimodal (around the origin) if every linear combination of the random variables $aX_1 + bX_2$ for $a, b \in \mathbb{R}$ is unimodal (around zero).

**Proposition 8. (Dependence)** Suppose that $u_i := \hat{u}_i + \varepsilon_i$ and $u_j := \hat{u}_j + \varepsilon_j$ where $\hat{u}_i, \hat{u}_j \in \mathbb{R}$ and $(\varepsilon_i, \varepsilon_j)$ is linear unimodal and symmetric around the origin. Then, for every level of imprecision $\hat{\sigma}$: $\frac{\partial p (i, \hat{\sigma})}{\partial \sigma} \leq 0 \iff i$ is weakly better.

**Proof:** By definition, $\varepsilon_i - \varepsilon_j$ is unimodal around zero. To see that it is also symmetric, let $g$ denote the density of $(\varepsilon_i, \varepsilon_j)$ and $g_{\varepsilon_i - \varepsilon_j}$ the density of $\varepsilon_i - \varepsilon_j$. Then, for all $x \in \mathbb{R}$,

$$g_{\varepsilon_i - \varepsilon_j} (x) = \int g (z + x, z) \, dz = \int g (-z - x, -z) \, dz = \int g (z - x, z) \, dz = \tilde{g}_{\varepsilon_i - \varepsilon_j} (x)$$

\[27\text{While Ghosh defines the class for } n\text{-dimensional random vectors, we only require two dimensions.}\]
where: the first and last equalities follow by equation (3); the second by symmetry; and the third by the change of variables \( z \to -z \). Since \( \varepsilon_i - \varepsilon_j \) is unimodal and symmetric around zero, the argument in the proof of Proposition 6 then applies verbatim. 

The error distributions covered by this result may be viewed as generalizations of bi-variate normals. As noted after Proposition 6, the difference of two independent normals is normal. In fact, this is true even when the random variables are dependent—provided that they are jointly normal. Linear unimodal and symmetric distributions have a similar closure property: the difference of two such random variables is unimodal and symmetric around zero. Not only does this show that increased precision cannot harm for bi-variate normal errors but it shows, more generally, that the same is true for every symmetric bi-variate distribution which is stable or log-concave.

In the other direction, we can dispense with unimodality and retain a “local” version of Proposition 5. As in that result, our assumption of i.i.d. errors plays an important role—since it ensures that the better alternative is the one with the higher mean.

**Proposition 9. (Multi-modality)** Suppose that \( u_i := \hat{u}_i + \varepsilon_i \) and \( u_j := \hat{u}_j + \varepsilon_j \) where \( \hat{u}_i, \hat{u}_j \in \mathbb{R} \) and \( \varepsilon_i, \varepsilon_j \) are mean-zero, i.i.d. random variables with continuous densities. Then, there exists an open interval \( B(|\hat{u}_i - \hat{u}_j|) \) around \( |\hat{u}_i - \hat{u}_j| \) such that, for every level of imprecision \( \hat{\sigma} \in B(|\hat{u}_i - \hat{u}_j|) \): \( \frac{\partial p(i, \hat{\sigma})}{\partial \hat{\sigma}} < 0 \iff i \) is better.

**Proof:**28 Let \( d := \hat{u}_i - \hat{u}_j \). Since \( f_i(x) = f_j(x - d) \) for all \( x \in \mathbb{R} \), one can use equation (4) and a change of variable \( z \to z + d \) to obtain the following identity:

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \hat{\sigma}} = \frac{1}{2} \left[ \int_{\mathbb{R}} f_j(z) f_j(z + \hat{\sigma} + d) \, dz - \int_{\mathbb{R}} f_j(z + \hat{\sigma} - d) f_j(z) \, dz \right].
\] (8)

Since \( \int_{\mathbb{R}} [f_j(z)]^2 \, dz = \int_{\mathbb{R}} [f_j(z + c)]^2 \, dz \) holds for all \( c \in \mathbb{R} \),

\[
\int_{\mathbb{R}} f_j(z) f_j(z + c) \, dz = \int_{\mathbb{R}} [f_j(z)]^2 \, dz - \frac{1}{2} A(c)
\]

where \( A(c) := \int_{\mathbb{R}} [f_j(z) - f_j(z + c)]^2 \, dz \). By substituting into equation (8), one obtains

\[
\frac{\partial p(i, \hat{\sigma})}{\partial \hat{\sigma}} = \frac{1}{4} [A(\hat{\sigma} - d) - A(\hat{\sigma} + d)].
\]

Suppose that \( i \) is better (or, equivalently, that \( \hat{u}_i > \hat{u}_j \)). Then,

\[
\frac{\partial p(i, d)}{\partial \hat{\sigma}} = \frac{1}{4} [A(0) - A(2d)] = -\frac{1}{4} \int_{\mathbb{R}} [f_j(z) - f_j(z + 2d)]^2 \, dz.
\] (9)

Notice that \( f_j \) is not a density if \( f_j(z) = f_j(z + 2d) \) for all \( z \in \mathbb{R} \). Otherwise,

\[
\int_{\mathbb{R}} f_j(z) \, dz = \sum_{k \in \mathbb{Z}} \int_{k \cdot 2d}^{(k+1) \cdot 2d} f_j(z) \, dz = \lim_{k \to \infty} 2k \left[ \int_{0}^{2d} f_j(z) \, dz \right] \neq 1.
\]

From this last observation, it follows that \( f_j(z) \neq f_j(z + 2d) \) for some \( z \in \mathbb{R} \). So, \( \frac{\partial p(i, d)}{\partial \hat{\sigma}} < 0 \) by equation (9). The desired result then follows by continuity of \( f_j \).  

---

28The proof adapts standard techniques to show that cross-correlation is minimal at zero.
5.2 Beyond two alternatives

It is not entirely straightforward to extend our analysis to more than two alternatives. In this section, we briefly mention two issues that complicate the task.

(1) Impact of precision: As the number of alternatives increases, the range of possible “ties” in the value realisations increases exponentially. This complicates significantly the task of determining the choice probabilities and the marginal impact of precision.

To illustrate, consider the case of three alternatives and suppose (as in the case of two alternatives) that the alternatives tied “at the top” are chosen with uniform probability. Let $R_{S}^\sigma$ denote the probability of the event that $S$ is the top-set of alternatives:

(i) for every $i \in S$, $\left| u_i - u_j \right| \leq \sigma$ for all $j \in S$; and,

(ii) for every $k \notin S$, $u_k + \sigma < u_j$ for some $j \in S$.

With this notation, $R_{12}^\sigma$, $R_{13}^\sigma$ and $R_{123}^\sigma$ reflect the events where alternative 1 is tied at the top and $R_{1}^\sigma$ the event where it wins outright. It follows that $p(1, \sigma)$ can be written as

$$p(1, \sigma) = R_{1}^\sigma + \frac{1}{2} [R_{12}^\sigma + R_{13}^\sigma] + \frac{1}{3} R_{123}^\sigma$$

$$= \frac{1}{3} + \frac{1}{3} [2R_{1}^\sigma - R_{2}^\sigma - R_{3}^\sigma] + \frac{1}{6} [R_{12}^\sigma + R_{13}^\sigma - 2R_{23}^\sigma].$$

(10)

The second formulation (which follows from the first by re-writing $R_{123}^\sigma$ in terms of its complementary probabilities) makes it more clear how $p(1, \sigma)$ is affected by increases in $\sigma$. Letting $R_{S \rightarrow T}^\sigma$ denote the probability of the threshold event that the top-set switches from $S$ to $T$ when $\sigma$ increases, the marginal effect of precision can then be written as

$$\frac{\partial p(1, \sigma)}{\partial \sigma} = \frac{1}{2} [R_{2 \rightarrow 12}^\sigma - R_{1 \rightarrow 12}^\sigma] + \frac{1}{2} [R_{3 \rightarrow 13}^\sigma - R_{1 \rightarrow 13}^\sigma]$$

$$+ \frac{1}{6} [2R_{23 \rightarrow 123}^\sigma - R_{12 \rightarrow 123}^\sigma - R_{13 \rightarrow 123}^\sigma].$$

(11)

This shows that three different trade-offs determine the marginal impact of precision when there are three alternatives. The first two terms are direct analogs of the two-alternative case where the perceived value of one alternative changes (either positively or negatively) relative to one other alternative. In the last term, the perceived value of one alternative changes relative to two other alternatives.\(^{29}\)

More generally, with $n$ alternatives, the marginal impact of precision on $p(i, \sigma)$ involves $2^{n-1} - 1$ different trade-offs. (This is easy to see: for each $k = 1, ..., n - 1$, there are $\binom{n-1}{k}$ trade-offs where the perceived value of one alternative changes relative to $k$ others.)

(2) Median quality: With two alternatives, we believe that there are compelling reasons to use the median value difference as a measure of quality. However, difficulties arise

\(^{29}\)For the interested reader, we derive explicit formulas for equations (10) and (11) in Appendix B.
in trying to generalize this measure to three or more alternatives. At a fundamental level, the issue is that the “univariate” median used for two alternatives does not necessarily identify a highest quality alternative. The following example serves to illustrate.

**Example 4.** For three alternatives, consider a random utility specification that induces the following distribution $\Pi_>$ over the ranking of value realizations

$$
\Pi_> := \begin{bmatrix}
\Pr(\text{u}_1 > \text{u}_2 > \text{u}_3) \\
\Pr(\text{u}_1 > \text{u}_3 > \text{u}_2) \\
\Pr(\text{u}_2 > \text{u}_1 > \text{u}_3) \\
\Pr(\text{u}_2 > \text{u}_3 > \text{u}_1) \\
\Pr(\text{u}_3 > \text{u}_1 > \text{u}_2) \\
\Pr(\text{u}_3 > \text{u}_2 > \text{u}_1)
\end{bmatrix} = \begin{bmatrix}
18/64 \\
7/64 \\
3/64 \\
15/64 \\
11/64 \\
10/64
\end{bmatrix}.
$$

In this case, the (pairwise) medians induce a cyclic quality ranking. Let $p_{ij}(i, 0)$ denote the probability that a perfectly precise ($\sigma = 0$) agent chooses $i$ over $j$. Then:

$$p_{12}(1, 0) = p_{23}(2, 0) = p_{13}(3, 0) = \frac{36}{64} > \frac{28}{64} = p_{12}(2, 0) = p_{23}(3, 0) = p_{13}(1, 0).$$

We emphasize that an independent random utility specification is sufficient to produce the ranking distribution $\Pi_>$, specifically one where each of $\text{u}_1 = (1, 4, 7, 7)$, $\text{u}_2 = (2, 6, 6, 6)$ and $\text{u}_3 = (3, 5, 5, 8)$ realizes with uniform probability. (In the literature, independent distributions that induce such pairwise “cycles” are known as non-transitive dice.)

To resolve this issue, one possibility is to rely on a “multivariate” concept of median to identify the highest quality alternative. Having said this, there are multiple ways to extend the “univariate” concept and it is not clear which is the most appropriate.

A different approach would be to measure quality by the median of a “multivariate” object. One possibility is the median ordering, which is determined by ranking alternatives by the probability that they are chosen from the grand set of alternatives. When there are two alternatives, the (top-ranked alternative according the) median ordering is the “univariate” median. For three or more alternatives, the median ordering is faithful to the idea that a perfectly precise ($\sigma = 0$) agent tends to choose well. To illustrate, consider the agent in Example 4. In that case, the median ordering is $1 > 3 > 2$ since

$$\Pr(\text{u}_1 > \text{u}_2, \text{u}_3) = \frac{25}{64} > \Pr(\text{u}_3 > \text{u}_1, \text{u}_2) = \frac{21}{64} > \Pr(\text{u}_2 > \text{u}_1, \text{u}_3) = \frac{18}{64}.$$ 

The median ordering captures the intuition that alternative 1 is the best among the three and, as a “second-order” concern, that alternative 3 is better than alternative 2.

### 5.3 Taste shocks

Throughout the paper, we have assumed that the random values reflect perception errors. If the errors instead reflect taste shocks, then the realised values are better interpreted as
welfare relevant utilities. In that case, it seems more appropriate to measure the quality of a decision by its expected utility.

To analyse this variation, we generalise the model by supposing that, in case of vague-ness, the agent chooses alternative \( i \) with probability \( \alpha \in (0, 1) \). The parameter \( \alpha \) (which is not necessarily \( \frac{1}{2} \)) expresses the bias of the agent towards alternative \( j \) when he cannot distinguish it from alternative \( i \). Then, the agent’s expected utility of choosing according to his coarse perception may be expressed as follows:

\[
\mathbb{E}[u(\sigma, \alpha)] := \mathbb{E}[u_i | u_i > u_i + \sigma] + \mathbb{E}[u_j | u_j > u_i + \sigma] + \mathbb{E}[\alpha u_i + (1 - \alpha) u_j | \sigma \geq |u_i - u_j|]
\]

Differentiating this expression and evaluating at the level of imprecision \( \hat{\sigma} \) gives:

\[
\frac{\partial \mathbb{E}[u(\hat{\sigma}, \alpha)]}{\partial \sigma} = - \left[ \alpha \hat{\sigma} \int f(z, z - \hat{\sigma}) \, dz + (1 - \alpha) \hat{\sigma} \int f(z, z + \hat{\sigma}) \, dz \right].
\]

Since each of the terms inside the brackets is non-negative, we conclude the following:

**Proposition 10.** For all \( \alpha \in (0, 1) \) and every level of imprecision \( \hat{\sigma} > 0 \):

\[
\frac{\partial \mathbb{E}[u(\hat{\sigma}, \alpha)]}{\partial \sigma} \leq 0.
\]

This result shows that, when measured by its expected utility, the quality of the agent’s decision never decreases with precision. This is true under completely general conditions that do not depend on either the joint distribution \( f \) of utilities or the bias \( \alpha \).

### 6 Concluding remarks

In this paper, we have studied the interplay between two distinct sources of imprecision in judgment: the absolute error in perceiving the value of a given option; and the relative error in comparing the values of two options. We captured the distinction between these two sources of error by modeling the perceived value of each option as a random variable and the perception of value difference by a “just noticeable” threshold \( \sigma \) (which we called the level of imprecision). While increased precision (i.e., a decrease in \( \sigma \)) is unambiguously beneficial when the agent perceives the value of options perfectly, matters are less clear when the agent perceives values imperfectly. Indeed, our results characterise different classes of problems where increased precision may lead to choosing the worse alternative.

While our analysis is primarily theoretical, our results have practical relevance. Consider, for instance, the GRADE international standard used for evaluating scientific evidence in medical practice, which provides a way to categorise the strength of evidence.

---

\[^{30}\]A similar generalisation could be made to our main model in equation (1). It would change the right-hand side of condition (\( \star \)) from \( f_{u_i - u_j} (-\hat{\sigma}) < f_{u_i - u_j} (\hat{\sigma}) \) to \( \alpha f_{u_i - u_j} (-\hat{\sigma}) < (1 - \alpha) f_{u_i - u_j} (\hat{\sigma}) \).
from clinical trials into different certainty ratings.\footnote{See Schünemann et al. [24] for full details on the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) framework.} If the result of a clinical trial is a random variable, then the attribution of a certainty rating corresponds in our framework to the comparison of the trial results to a fixed benchmark; and the judgment on whether or not to adopt the treatment will come down to the clinician’s perception of the difference between the trial results and the certainty rating.

Concretely, suppose that a physician must choose between recommending two treatments (e.g., an exercise regimen and a specific diet) to an overweight patient.\footnote{This example is loosely based on Clark [7], a meta study on the efficacy of various training and/or dietary regimes for weight loss. In this study, differences in effectiveness are reported with reference to various metrics, from body mass to fat free mass and blood levels of certain hormones. The various tables and figures in that paper summarise the difference in distributions of the relevant variable of interest across the studies considered.} Each treatment is a random variable whose realisation reflects the aggregate weight losses of the participants in the associated medical trial. Which treatment will be judged more effective will depend on the physician’s level of confidence (as measured by $\sigma$). According to our results, a more confident physician (with a smaller $\sigma$) may recommend the wrong treatment more often than a less confident physician (with a larger $\sigma$).

As a second illustration, consider the implications of our analysis for labor markets. To fix ideas, suppose that an applicant’s performance at a job interview is a random variable; and that the prospective employer bases the hiring decision on the interview performance relative to a benchmark (which reflects the criteria in the job description). In particular, suppose that the employer definitely hires an applicant whose interview performance exceeds the benchmark by threshold difference $\sigma$ and definitely does not hire one whose performance falls short by $\sigma$.

Our results show that unintended consequences can result when the threshold $\sigma$ varies with observable characteristics of the applicant. To illustrate, suppose that the employer applies a smaller threshold to applicants with a post-secondary degree (reasoning, perhaps, that such applicants should exhibit less variance between perceived performance and true ability). Then, contrary to compensating for higher variance, our results show that an even higher proportion of errors (in either direction) could be made when hiring applicants who lack post-secondary education. More generally, if the level of imprecision $\sigma$ (interpreted as a level of “tolerance”) varies with observable characteristics like age, sex or race, it can result in unintended discrimination.

To close, it is worth noting that our analysis assumed that the agent’s perception error $\sigma$ does not depend on the choice problem. We took this approach because it allowed us to study the comparative statics of imprecise judgment with as few “moving parts”
as possible. Having said this, it would be interesting to see how our conclusions might change when $\sigma$ depends on the set of the alternatives being compared. We leave this investigation to further research.

References


In the special case where $c = 1$, the expressions in (13) and (14) simplify to

$$F_{u_i-u_j}(x) = A(x) \int_{\mathbb{R}_+} e^{-y(A(x)+1)} dy = \frac{A(x)}{A(x)+1}$$  \hspace{1cm} (15)

$$f_{u_i-u_j}(x) = A(x) \int_{\mathbb{R}_+} e^{-y(A(x)+1)} dy - A(x)^2 \int_{\mathbb{R}_+} ye^{-y(A(x)+1)} dy = \frac{A(x)}{(A(x)+1)^2}$$  \hspace{1cm} (16)

(The last equality in (16) follows from integration by parts.) These formulas correspond to the cdf and density of a logistic distribution with location $\hat{u}_i - \hat{u}_j$ and scale 1.

Given (15), the probability that $i$ “beats” $j$ is then given by

$$\Pr(u_i > u_j + \sigma) = 1 - F_{u_i-u_j}(\sigma) = \frac{1}{A(\sigma)+1} = \frac{e^{\hat{u}_i}}{e^{u_j+\sigma} + e^{\hat{u}_i}}.$$  

Using this formula, the analysis then proceeds as in the main text.
A.2 

Example 3

When $c > 1$, simple closed form expressions are lacking and the analysis is much more involved. Nonetheless, we can make the following simple observations:

(i) Since $u_i$ and $u_i$ are independent Gumbel, their mean value difference is

$$E(u_i - u_j) = E(u_i) - E(u_j) = [\hat{u}_i + c\gamma] - [\hat{u}_j + \gamma] = (\hat{u}_i - \hat{u}_j) + (c - 1)\gamma$$

where $\gamma$ denotes the Euler-Mascheroni constant.

(ii) For $c$ sufficiently close to 1, the median of the value difference $m_{u_i - u_j}$ is approximated by the difference of the median values so that

$$m_{u_i - u_j} \approx m_u - m_u = [\hat{u}_i - c\ln 2] - [\hat{u}_j - \ln 2] = (\hat{u}_i - \hat{u}_j) + (c - 1)|\ln 2|.$$ 

From (i) and (ii), it follows that $i$ is both mean-better and median-worse than $j$ when

$$0.36651 \approx |\ln 2| \approx \frac{\hat{u}_j - \hat{u}_i}{c - 1} < \gamma \approx 0.57722.$$ 

B. Three alternatives

To derive an explicit formula for (10), one must compute $R_{ij}^\sigma$ and $R_{ij}^\sigma$ for $i, j \in \{1, 2, 3\}$. Where $f(x, y, z)$ denotes the joint density of $(u_1, u_2, u_3)$, it is straightforward to see that

$$R_{11}^\sigma = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y, z) dy dz dx$$

$$R_{12}^\sigma = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y, z) dz dy dx + \int_{-\infty}^{y} \int_{-\infty}^{x} f(x, y, z) dz dx dy$$

The expressions for $R_{ij}^\sigma$, $R_{ji}^\sigma$, $R_{12}^\sigma$ and $R_{21}^\sigma$ are symmetric. To elaborate, observe that $\{u_1\}$ is the top-set if and only if the value realizations are such that $u_2, u_3 < u_1 - \sigma$, which gives (17). In turn, (18) follows from the observation that $\{u_1, u_2\}$ is the top-set if and only if: (i) $u_3 < u_1 - \sigma \leq u_2 \leq u_1$; or, similarly, (ii) $u_3 < u_2 - \sigma \leq u_1 \leq u_2$.

To derive an explicit formula for (11), one must compute the threshold probabilities $R_{ij}^\sigma$ and $R_{ij}^\sigma$ for $i, j, k \in \{1, 2, 3\}$. By the same kind of reasoning as above, we have

$$R_{21}^\sigma = \int_{-\infty}^{y} f(y - \sigma, y, z) dz dy$$

$$R_{23}^\sigma - R_{12}^\sigma = \int_{-\infty}^{y} f(y - \sigma, y, z) dz dy + \int_{-\infty}^{z} f(z - \sigma, y, z) dy dz$$

The other threshold probabilities are symmetric. For (19), note that the boundary between top-sets $\{u_2\}$ and $\{u_1, u_2\}$ is defined by the requirement that $u_3 < u_1 = u_2 - \sigma$. Similarly, for (18), note that the boundary between top-sets $\{u_2, u_3\}$ and $\{u_1, u_2, u_3\}$ requires: (i) $u_1 = u_2 - \sigma \leq u_3 \leq u_2$; or (ii) $u_1 = u_3 - \sigma \leq u_2 \leq u_3$. 

25
To obtain an explicit formula for $p(i, \sigma)$, one can replace (17), (18) and their analogs into equation (10). One can then check that the result obtained by differentiating $p(i, \sigma)$ with respect to $\sigma$ coincides exactly with the formula given by replacing (19), (20) and their analogs into equation (11). We leave these calculations to the reader.