Vilkkumaa, Eeva; Liesiö, Juuso

What Causes Post-Decision Disappointment? Estimating the Contributions of Systematic and Selection Biases

Published in:
European Journal of Operational Research

DOI:
10.1016/j.ejor.2021.04.018

Published: 16/01/2022

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Vilkkumaa, E., & Liesiö, J. (2022). What Causes Post-Decision Disappointment? Estimating the Contributions of Systematic and Selection Biases. *European Journal of Operational Research*, *296*(2), 587-600. https://doi.org/10.1016/j.ejor.2021.04.018

European Journal of Operational Research 296 (2022) 587–600

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

What causes post-decision disappointment? Estimating the contributions of systematic and selection biases

Eeva Vilkkumaa*, Juuso Liesiö

Department of Information and Service Management, Aalto University School of Business, P.O. Box 21210, Aalto 00076, Finland

ARTICLE INFO

Article history:
Received 3 September 2020
Accepted 7 April 2021
Available online 18 April 2021

Keywords:
Behavioural OR
Decision-making
Post-decision disappointment
Systematic bias
Selection bias

ABSTRACT

Many empirical studies suggest that the realized values of highest-ranked decision alternatives tend to be systematically lower than estimated, causing the decision-maker to experience post-decision disappointment. This systematic overestimation of value has been explained by a systematic bias in the alternatives' estimated values resulting from, e.g., a behavioural disposition towards overoptimism or even strategic misrepresentation. Nevertheless, even if these estimates are unbiased, the value of the selected alternative is likely to be overestimated due to selection bias. In this paper, we build models for measuring the shares of systematic and selection biases in generating post-decision disappointment, and develop approaches for estimating these models from data which contains the estimated values of all alternatives but the realized values of selected alternatives only. Results obtained from applying these models to real data on 5610 transportation infrastructure projects suggest that out of the total cost overrun of \$2.77 billion, only ca. 24% can be attributed to systematic bias.

1. Introduction

Businesses and public organizations frequently need to decide between several decision alternatives; for instance, which R&D projects to launch (Baker, Bosetti, & Salo, 2020), which supplier to work with (Kellner, Lienland, & Utz, 2019), or which contractor to select for carrying out a public infrastructure project (Gupta, Snir, & Chen, 2015). These types of decision processes typically involve three phases: (i) identifying the set of feasible decision alternatives, (ii) estimating the values (e.g., expected utilities or certainty equivalents) of these alternatives, and (iii) selecting the alternative that has the highest estimated value.

Because of uncertainties, the value estimates are never perfectly accurate, whereby the *ex post* realized values of the alternatives rarely coincide with their *ex ante* estimated values. In particular, many empirical studies suggest that the realized values of selected alternatives tend to be systematically *lower* than estimated, causing the decision-maker to experience *post-decision disappointment* (Bell, 1985; Harrison & March, 1984; Smith & Winkler, 2006). For instance, revenue forecasts of capital investment projects in companies have been shown to be overstated in roughly 80% of cases (Brous, Hiedemann, & Schultz, 2009; Pruitt & Gitman, 1987). On the other hand, cost overruns have been shown to occur in 90% of large transportation infrastructure projects worldwide with an average overrun of 27.6% (Flyvbjerg, 2009). Cost overruns are typical in many other decision contexts as well, including the selection of road construction projects (Bajari, Houghton, & Tadelis, 2011; Odeck, 2004) and software projects (Jø rgensen, 2013), defense procurement (Terasawa, Quirk, & Womer, 1989), and plant investments in chemical process industries (Merrow, Phillips, & Myers, 1981).

The systematic overestimation of value – or, analogously, the systematic underestimation of cost – has typically been attributed to a *systematic bias* in the alternatives' estimated values or costs. A great deal of research has been devoted to examining the possible sources for such systematic bias, particularly in the fields of Behavioural Economics (Tversky & Kahneman, 1974), Behavioural Finance (Shleifer, 2000; Subrahmanyam, 2008), and Behavioural Operational Research (Franco & Hämäläinen, 2016). First, there may be motivational reasons for producing overly optimistic estimates for the preferred alternative. In competitive bidding situations, for instance, project promoters gain a strategic advantage by underestimating the costs of their project proposal (Bajari et al., 2011; Flyvbjerg, 2009; Montibeller & von Winterfeldt, 2015). On the other hand, even without outside motivation for deliberate misrepresentation, biased estimates are likely due to a behavioural disposition toward overoptimism or overconfidence (Flyvbjerg, 2009; Kahneman & Lovallo, 1993; Lovallo & Kahneman, 2003; Montibeller &

* Corresponding author.
E-mail addresses: eeva.vilkkumaa@aalto.fi (E. Vilkkumaa), juuso.liesio@aalto.fi (J. Liesiö).

https://doi.org/10.1016/j.ejor.2021.04.018
0377-2217/

von Winterfeldt, 2015). For instance, overconfident CEOs have been found to overinvest in the firm's projects due to a systematic overestimation of the projects' returns, even when such overinvestment exposes them to personal risk (Malmendier & Tate, 2005).

When the performance of decision alternatives is assessed on multiple criteria, systematic biases in the alternatives overall value estimates (computed as weighted sums of the criterion-specific assessments, for instance) may arise from scaling bias (assessments are affected by the choice of measurement scales), gain-loss bias (assessments are affected by whether they are framed as gains vs. losses), proxy bias (criteria for which proxy measurement scales are used are overweighted), or even by the vividness with which different criteria are presented (Ferretti & Geneletti, 2020; Montibeller & von Winterfeldt, 2015; Morton & Fasolo, 2009). Moreover, biases associated with different stages of multicriteria decision analytic methods can accumulate in ways that result in significant overall estimation errors (Lahtinen & Hämäläinen, 2016; Lahtinen, Hämäläinen, & Jenytin, 2020).

Nevertheless, even if there is no systematic bias in the alternatives' value estimates but only the alternative with the highest value estimate is selected, the value of this selected alternative is likely to be overestimated due to *selection bias* (Begg & Bratvold, 2008; Kettunen & Salo, 2017; Quirk & Terasawa, 1986; Smith & Winkler, 2006; Vilkkumaa, Liesiö, & Salo, 2014). This is because although the estimation errors among *all* proposed alternatives might cancel each other out, the value of the *selected* alternative – i.e., the one with the highest value estimate – is more likely to have been overestimated rather than underestimated. In other words, even in the absence of systematic estimation bias, post-decision disappointment is likely due to a *selection process* that favors alternatives with overestimated values.

This leads to the main research question of this paper: what are the relative contributions of systematic and selection biases in generating post-decision disappointment? The need to distinguish between these biases has been recognized in the literature (Jø rgensen, 2013; Quirk & Terasawa, 1986; Smith & Winkler, 2006) but, to our knowledge, no results have been presented on their relative shares in real-life applications. The lack of such results is probably explained by both methodological challenges as well as difficulties in obtaining suitable data. First, suitable data would need to include both the estimated and realized values of decision alternatives over multiple decision settings. Few firms keep careful records documenting the estimated values of their decision alternatives and, even if they did, such data is likely to be confidential. Second, if such data could be obtained, it would not (and could not) contain information about the true values of those alternatives that were never implemented. Such incomplete data would rule out the use of standard regression techniques for estimating the relationship between the alternatives' true and estimated values through which the distinction between the two biases could be made.

In this paper we develop models for estimating the relative shares of systematic and selection biases based on incomplete empirical data. In particular, we build models for the alternatives' true and estimated values and develop Expectation Maximization algorithms (Dempster, Laird, & Rubin, 1977) for estimating the parameters of these models from data that does not contain the true values of those alternatives that were not selected. We demonstrate the use of these algorithms through multiple simulation examples, and illustrate how the use of standard regression techniques on data consisting of the true and estimated values of selected alternatives alone – as is often done in practice– is likely to lead to overestimating systematic bias and, consequently, underestimating selection bias.

We apply the developed models to a real data set consisting of 5610 decisions on the selection of transportation infrastructure projects during 2000–2014, each with 1–33 decision alternatives. To our knowledge, this is the first contribution to estimating the relative magnitudes of systematic bias and selection bias based on data from real-life decision settings. Our results suggest that out of the cumulative post-decision disappointment of \$2.77 billion, only ca. 24% can be attributed to systematic bias. An important implication of this finding is that measures taken to reduce systematic bias through debiasing techniques (Montibeller & von Winterfeldt, 2015; Morton & Fasolo, 2009) or even financial or professional penalties (Flyvbjerg, Holm, & Buhl, 2002) may be relatively ineffective in mitigating overall post-decision disappointment.

The rest of this paper is structured as follows. In Section 2 we present the model for post-decision surprise, define systematic and selection biases that contribute to expected post-decision disappointment, and review earlier results on eliminating such disappointment through Bayesian modeling of estimation uncertainties. In Section 3, we present the EM algorithm and examples thereof. The application of the models to transportation infrastructure data is presented in Section 4. In Section 5 we conclude by discussing the limitations of our models and the implications of our results.

2. Modeling post-decision surprise

2.1. Basic concepts and notation

Suppose[1] that a decision-maker is considering n decision alternatives indexed by $j = 1, \ldots, n$ with true values $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^T$. These true values, which are assumed to be a realization of a vector-valued random variable $\boldsymbol{M} \sim f(\boldsymbol{\mu})$, represent the values that the alternatives would yield if implemented. If the decision-maker knew $\boldsymbol{\mu}$ beforehand, she would select the alternative with the highest true value, the index of which is

$$j(\boldsymbol{\mu}) = \operatorname*{argmax}_j \mu_j. \tag{1}$$

Yet, the decision-maker does not know the true values, but needs to make the decision based on uncertain estimates $\boldsymbol{v} = [v_1, \ldots, v_n]^T$ thereof. These value estimates $\boldsymbol{v}$ are a realization of a vector-valued random variable $(\boldsymbol{V}|\boldsymbol{M} = \boldsymbol{\mu}) \sim f(\boldsymbol{v}|\boldsymbol{\mu})$. The index of the alternative selected based on such estimates is

$$j(\boldsymbol{v}) = \operatorname*{argmax}_j v_j. \tag{2}$$

The ex post observed true value of the selected alternative is $\mu_{j(\boldsymbol{v})}$, which is a realization of $M_{j(\boldsymbol{v})}$. Fig. 1 summarizes the decision setting and notation through an influence diagram.

Because the value estimates are uncertain, the true value of the selected alternative $\mu_{j(\boldsymbol{v})}$ rarely coincides with its estimated value $v_{j(\boldsymbol{v})}$. The difference $\mu_{j(\boldsymbol{v})} - v_{j(\boldsymbol{v})}$ between the true and estimated value of the selected alternative is called *post-decision surprise* (Harrison & March, 1984). Assume first that the value estimates are *conditionally unbiased* so that

$$\mathbb{E}[V_j|\boldsymbol{M} = \boldsymbol{\mu}] = \mu_j \text{ for all } j = 1, \ldots, n. \tag{3}$$

In this case, the expected post-decision surprise of any randomly selected alternative j is zero, i.e., $\mathbb{E}[M_j - V_j] = 0$. Yet, the expected post-decision surprise $\mathbb{E}[M_{j(\boldsymbol{V})} - V_{j(\boldsymbol{V})}]$ from selecting the alternative with the *highest value estimate* is non-positive, whereby the decision-maker is expected to experience *post-decision disappointment* (Harrison & March, 1984; Vilkkumaa et al., 2014). Smith and Winkler (2006) have named this phenomenon the *optimizer's curse*.

The expected post-decision disappointment is illustrated in Fig. 2, which shows the distribution of the maximum value estimate $V_{j(\boldsymbol{V})} = \max\{V_1, \ldots, V_n\}$ among n decision alternatives when

[1] A list of notation is presented in Appendix A.

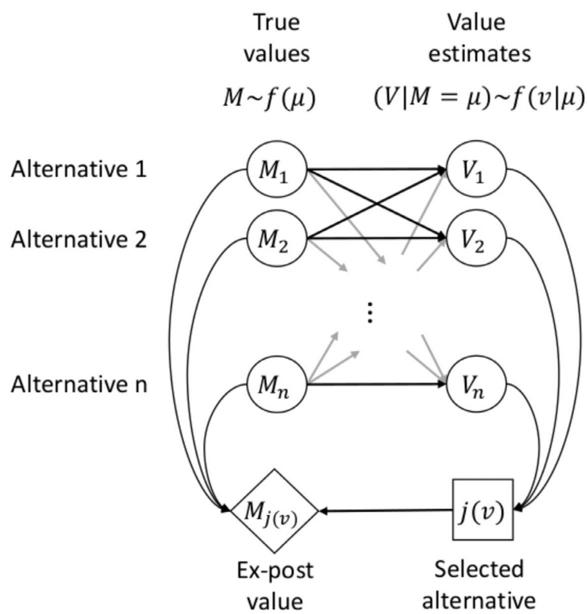


Fig. 1. Influence diagram describing the decision setting.

each value estimate is conditionally unbiased and follows the standard normal distribution $(V_j|\boldsymbol{M}=\boldsymbol{\mu}) \sim \mathrm{N}(0, 1^2)$ so that the true value of each alternative is $\mu_j = \mathbb{E}[V_j|\boldsymbol{M}=\boldsymbol{\mu}] = 0$ (solid lines). The table shows the expected post-decision disappointment (i.e., the expected value of the maximum value estimate minus the true value zero) for different values of n. In this setting, the expected post-decision disappointment increases with the number of alternatives n. For instance, the expected disappointment $\mathbb{E}[V_{j(\boldsymbol{V})} - M_{j(\boldsymbol{V})}]$ is 56% of the standard deviation of the value estimates when there are two alternatives, and 174% of the standard deviation when the number of alternatives is 15. Here, post-decision disappointment does not result from a systematic bias in the alternatives' value estimates, but from consistently selecting the alternative with the highest value estimate. Indeed, the expected disappointment is zero only if there is a single decision alternative so that no selection decision needs to be made.

Consider now the general case in which there may be a systematic bias in the value estimates. The dashed lines in Fig. 2 show the distributions of the maximum value estimate among n decision alternatives when the true value of each alternative is zero, but the distribution of each value estimate is $(V_j|\boldsymbol{M}=\boldsymbol{0}) \sim \mathrm{N}(0.5, 1^2)$ so that $\mathbb{E}[V_j|\boldsymbol{M}=\boldsymbol{0}] = 0.5 \neq 0$. Compared to the case with unbiased estimates (solid lines), the distributions of the maximum value estimates are shifted toward the right by 0.5. Consequently, the expected disappointment increases by 0.5 for each value of n. An Excel file for generating the illustrations shown in Fig. 2 is provided as supplementary material for this paper.

Systematic bias in the value estimates can be eliminated by adjusting each estimate in a suitable way. Technically, such an adjustment corresponds to defining a transformation $d(\cdot)$ such that the transformed estimate $d(V_j)$ is conditionally unbiased. For instance, the systematic bias in the value estimates of Fig. 2 could be eliminated through transformation $d(V_j) = V_j - 0.5$.

Definition 1. Let μ_j and $(V_j|\boldsymbol{M}=\boldsymbol{\mu})$ be the true and estimated values of alternative j. The *debiasing transformation* is a mapping $d : \mathbb{R} \to \mathbb{R}$ that satisfies

$$\mathbb{E}[d(V_j)|\boldsymbol{M}=\boldsymbol{\mu}] = \mu_j.$$

The alternatives' debiased value estimates form a vector-valued random variable denoted by $\tilde{\boldsymbol{V}} = [\tilde{V}_1, \ldots, \tilde{V}_n]^T$, where $\tilde{V}_j = d(V_j)$. The realization of this variable is denoted by $\tilde{\boldsymbol{v}} = [\tilde{v}_1, \ldots, \tilde{v}_n]^T$. The index of the alternative selected based on debiased estimates is

$$j(\tilde{v}) = \operatorname*{argmax}_j \tilde{v}_j. \tag{4}$$

In the example of Fig. 2, for instance, the distributions of the maximum debiased estimates $\tilde{V}_{j(\tilde{V})} = \max\{\tilde{V}_1, \ldots, \tilde{V}_n\}$ for different values of n would correspond to those illustrated by the solid lines.

When the alternatives' true values and value estimates are random variables $\boldsymbol{M}$ and $\boldsymbol{V}$, the indices of the alternatives with the highest true value, value estimate, and debiased estimate are also random variables. We denote these random variables by J^*, J, and $\tilde{J}$, respectively:

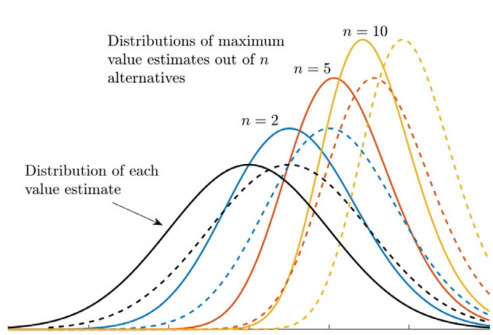
$$J^* = j(\boldsymbol{M}) = \operatorname*{argmax}_j M_j, \tag{5}$$

$$J = j(\boldsymbol{V}) = \operatorname*{argmax}_j V_j, \tag{6}$$

$$\tilde{J} = j(\tilde{\boldsymbol{V}}) = \operatorname*{argmax}_j \tilde{V}_j. \tag{7}$$

2.2. Total, systematic, and selection biases

We define *total bias* as the expected difference between the true and estimated value of the selected alternative, i.e., the expected post-decision surprise. This total bias consists of two parts: the part explained by the systematic bias in the value estimates (*systematic bias*), and the part that remains after the estimates have been debiased and is, therefore, explained entirely by the selection process (*selection bias*). In the example of Fig. 2, the total bias corresponding to biased value estimates and $n = 4$ alternatives is -1.53, of which systematic bias constitutes -0.5 and selection bias



Number n of alternatives	Expected disappointment (unbiased estimates)	Expected disappointment (biased estimates)
1	0.00	0.50
2	0.56	1.06
3	0.85	1.35
4	1.03	1.53
5	1.16	1.66
6	1.27	1.77
7	1.35	1.85
8	1.42	1.92
9	1.49	1.99
10	1.54	2.04
15	1.74	2.24

Fig. 2. The distribution of the maximum of n value estimates, and the expected disappointment resulting from selecting the alternative with the maximum value estimate when the estimates are (i) conditionally unbiased $(V_j|\boldsymbol{M}=\boldsymbol{\mu}) \sim \mathrm{N}(0, 1^2)$ (solid line) or (ii) conditionally biased $(V_j|\boldsymbol{M}=\boldsymbol{\mu}) \sim \mathrm{N}(0.5, 1^2)$ (dashed line).

the remaining -1.03. These concepts are formalized in the following Definition.

Definition 2. Total bias (*TB*), selection bias (*SeB*), and systematic bias (*SyB*) are defined as

$$TB = \mathbb{E}[M_J - V_J], \tag{8}$$

$$SeB = \mathbb{E}[M_{\tilde{J}} - \tilde{V}_{\tilde{J}}], \tag{9}$$

$$SyB = TB - SeB = \mathbb{E}[(M_J - V_J) - (M_{\tilde{J}} - \tilde{V}_{\tilde{J}})], \tag{10}$$

where the indices of the alternatives with the highest value estimate J and the highest debiased value estimate $\tilde{J}$ are given by (6) and (7), respectively.

Even if the value estimates are unbiased, it can be shown that the value of the selected alternative is expected to be lower than or equal to its estimated value, which implies a non-positive selection bias. Moreover, if there is a chance of selecting the 'wrong' alternative, selection bias is strictly negative – as was the case in the example of Fig. 2 with normally distributed value estimates. These results, which hold for any distributions $f(\boldsymbol{\mu})$ and $f(\boldsymbol{v}|\boldsymbol{\mu})$ of the alternatives' true and estimated values, are formalized in Proposition 1.

Proposition 1. *SeB* ≤ 0. *Furthermore, if* $\mathbb{P}(J^* \neq \tilde{J}) > 0$, *then SeB* < 0.

All proofs are in Appendix B. Proposition 1 is a generalization of Proposition 1 by Smith and Winkler (2006) to cases in which the value estimates are not necessarily conditionally unbiased.

In settings with a single decision alternative no selection decision needs to be made, whereby selection bias is zero. In this case total bias is entirely explained by a systematic bias in the estimated values of the alternatives. This is illustrated in the example of Fig. 2 where, in the case of biased value estimates, the expected difference between the true and estimated value of the selected alternative given $n = 1$ is equal to the systematic bias in the value estimates, i.e., -0.5.

Proposition 2. *If there is a single decision alternative, then SeB* $= 0$, *whereby TB* $=$ *SyB.*

Unlike selection bias, systematic bias can be either positive or negative. If, for instance, the values of all decision alternatives are systematically underestimated, then systematic bias is positive and can in fact cancel out the negative selection bias. Interestingly, systematic bias can also be positive when the values of all decision alternatives are systematically overestimated, as long as the expected difference $\mathbb{E}[M_J - M_{\tilde{J}}]$ between the true values of the alternatives selected based on biased and debiased estimates is larger than the expected difference $\mathbb{E}[V_J - \tilde{V}_{\tilde{J}}]$ between their estimated values (cf. Eq. (10)). This is the case, for instance, when there are 20 decision alternatives representing two types: (i) the first ten alternatives have true values with relatively large variability $M_i \sim \mathrm{N}(0, 4^2)$ and value estimates that are heavily biased but otherwise relatively accurate $(V_i|M_i = \mu_i) \sim \mathrm{N}(\mu_i + 2, 1^2)$; and (ii) the other ten have true values with relatively low variability $M_i \sim \mathrm{N}(0, 1^2)$ and value estimates that are only slightly biased but very inaccurate $(V_i|M_i = \mu_i) \sim \mathrm{N}(\mu_i + 0.1, 10^2)$. The distributions of the true values as well as biased and debiased value estimates for both types are illustrated in Fig. 3. Following from the distribution assumptions, it is likely that the alternative that has the highest true value μ_i is of the first type (cf. Fig. 3a). Nevertheless, an alternative of the second type is much more likely to have the highest value estimate due to the large estimation error variance, and this likelihood is further increased when the estimates are debiased (cf. Fig. 3a). In other words, debiasing increases the probability of selecting an alternative of the 'wrong' type. Therefore, the expected difference $\mathbb{E}[M_J - M_{\tilde{J}}]$ between the true values of the alternatives selected based on biased and debiased estimates is positive (ca. 0.40) and, in this case, larger than expected difference $\mathbb{E}[V_J - \tilde{V}_{\tilde{J}}]$ between their estimated values (ca. 0.27).

Cases in which systematic bias is positive (at least to the extent that it would cancel out the effect of negative selection bias) do not represent the empirically observed tendency of the value of the selected alternative to have been overestimated. Therefore, we focus here on situations in which systematic bias is non-positive. This can be shown to be the case if debiasing (i) does not change the selected alternative (i.e., $\mathbb{P}(J = \tilde{J}) = 1$) and (ii) does not make the value estimate of the selected alternative larger (i.e., $\mathbb{P}(\tilde{V}_J > V_J) = 0$). An example of a situation in which these conditions hold is shown in Fig. 2, in which the biased estimates are debiased by subtracting a constant value of 0.5 so that debiasing (i) does not change the ranking of the alternatives and (ii) makes the value estimate of each alternative smaller. Moreover, systematic bias is zero if and only if the estimate of the selected alternative is conditionally unbiased so that $\mathbb{P}(\tilde{V}_J = V_J) = 1$.

Proposition 3. *If* $\mathbb{P}(J = \tilde{J}) = 1$ *and* $\mathbb{P}(\tilde{V}_J > V_J) = 0$, *then SyB* ≤ 0. *Moreover, SyB* $= 0$ *if and only if* $\mathbb{P}(\tilde{V}_J = V_J) = 1$.

Because total bias is the sum of systematic and selection biases, total bias is non-positive under the conditions of Proposition 3. Moreover, if there is a chance of selecting the wrong alternative or if debiasing makes the value estimate of the selected alternative strictly smaller, then total bias is strictly negative.

Corollary 1. *If* $\mathbb{P}(J = \tilde{J}) = 1$ *and* $\mathbb{P}(\tilde{V}_J > V_J) = 0$, *then TB* ≤ 0. *Moreover, if* $\mathbb{P}(J^* \neq J) > 0$ *or* $\mathbb{P}(\tilde{V}_J \geq V_J) = 0$, *then TB* < 0.

Fig. 4 shows the total, systematic, and selection biases for different numbers of alternatives n with true values $\mu_j = 0$ and estimated values $(V_j|\boldsymbol{M} = \boldsymbol{\mu}) \sim \mathrm{N}(\mu_j + 0.5, 1^2)$ for all $j = 1, \ldots, n$. Systematic bias is constant -0.5, but selection bias (and, therefore, total bias) becomes more pronounced when the number of decision alternatives increases. Consequently, the larger the number of alternatives, the higher the share of total bias that is explained by selection bias. For instance, when there is only one alternative, the share of selection bias out of total bias is zero, but when the number of alternatives is 10, this share is $1.54/2.04 \approx 75\%$.

2.3. Computation of posterior means for the alternatives' true values

Biases in the estimated values of the selected alternatives can be overcome by modeling estimation uncertainties through standard Bayesian methods (see, e.g., Harrison & March, 1984; Smith & Winkler, 2006). In particular, assuming that the prior distribution $f(\boldsymbol{\mu})$ and the likelihood distribution $f(\boldsymbol{v}|\boldsymbol{\mu})$ have been estimated using data from past decision settings, the posterior distribution $f(\boldsymbol{\mu}|\boldsymbol{v})$ for the alternatives' true values given the value estimates can be obtained through Bayes' rule: $f(\boldsymbol{\mu}|\boldsymbol{v}) \propto f(\boldsymbol{v}|\boldsymbol{\mu})f(\boldsymbol{\mu})$. Then, rather than taking the estimates $\boldsymbol{v}$ at face value, decisions are based on the alternatives' posterior means $\mathbb{E}[M_j|\boldsymbol{V} = \boldsymbol{v}]$.

Let us denote the posterior means by vector $\hat{\boldsymbol{v}} = [\hat{v}_1, \ldots, \hat{v}_n]^T$, $\hat{v}_j = \mathbb{E}[M_j|\boldsymbol{V} = \boldsymbol{v}]$, where $\hat{\boldsymbol{v}}$ is a realization of a vector-valued random variable $\hat{\boldsymbol{V}} = [\hat{V}_1, \ldots, \hat{V}_n]^T$, $\hat{V}_j = \mathbb{E}[M_j|\boldsymbol{V}]$. The index of the alternative selected based on $\hat{\boldsymbol{v}}$ is

$$j(\hat{\boldsymbol{v}}) = \operatorname*{argmax}_j \hat{v}_j. \tag{11}$$

Without making any specific assumptions about distributions $f(\boldsymbol{\mu})$ and $f(\boldsymbol{v}|\boldsymbol{\mu})$, it can be shown that the expected difference between the true value and the posterior mean of the selected alternative is zero. Thus, the use of posterior means $\hat{\boldsymbol{v}}$ as a basis for

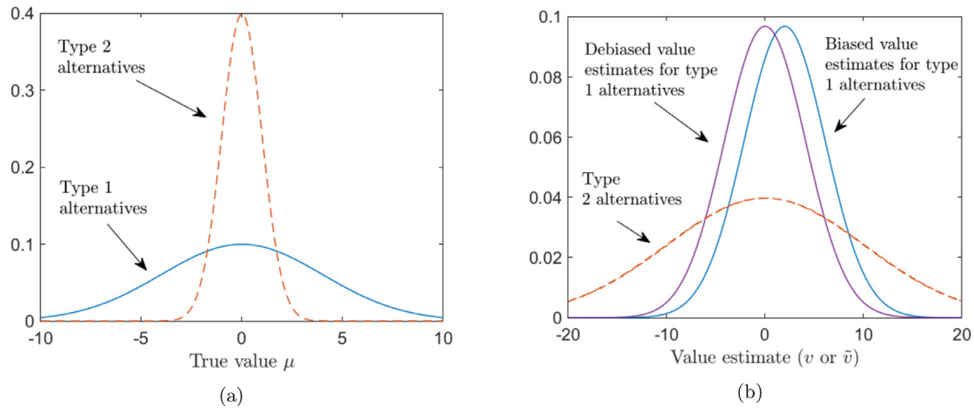


Fig. 3. The distributions of the true values M_i, value estimates V_i and debiased value estimates $\tilde{V}_i$ for two types of alternatives. Type 1 (solid line): $M_i \sim \mathrm{N}(0, 4^2)$, $V_i \sim \mathrm{N}(2, 4^2 + 1^2)$, $\tilde{V}_i = V_i - 2$. Type 2 (dashed line): $M_i \sim \mathrm{N}(0, 1^2)$, $V_i \sim \mathrm{N}(0.1, 1^2 + 10^2)$, $\tilde{V}_i = V_i - 0.1$. For type 2 alternatives, the distribution of $\tilde{V}_i$ is indistinguishable from that of V_i.

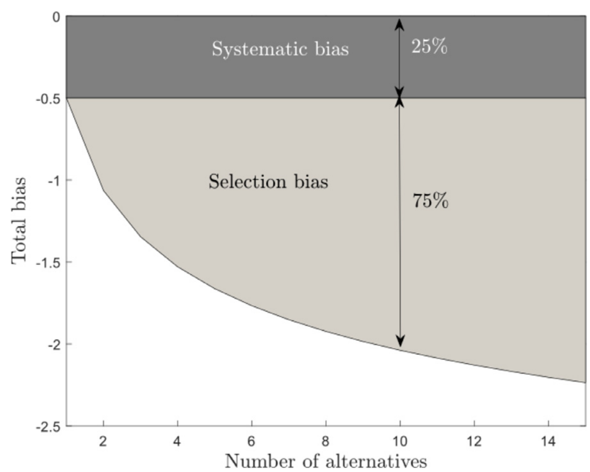


Fig. 4. Total, systematic, and selection biases for different numbers of alternatives $n = 1, \dots, 15$, when $\mu_j = 0$ and $(V_j|\boldsymbol{M} = \boldsymbol{\mu}) \sim \mathrm{N}(\mu_j + 0.5, 1^2)$ for all $j = 1, \dots, n$.

selection eliminates total bias (also in cases where systematic bias is positive). This well-known result is formalized in the following Proposition using our notation. Similar propositions have been presented in the literature under slightly different assumptions (e.g., conditionally unbiased estimates; Smith & Winkler, 2006).

Proposition 4. *Let $\hat{\boldsymbol{V}} = [\hat{V}_1, \dots, \hat{V}_n]$, where $\hat{V}_j = \mathbb{E}[M_j|V]$, and let random variable $\hat{J} = j(\hat{\boldsymbol{V}})$ be the index of the alternative with the highest expected value, obtained through (11) with random $\hat{\boldsymbol{V}}$. Then, $\mathbb{E}[M_{\hat{J}} - \hat{V}_{\hat{J}}|\boldsymbol{V} = \boldsymbol{v}] = 0$ for all $\boldsymbol{v}$, and hence $\mathbb{E}[M_{\hat{J}} - \hat{V}_{\hat{J}}] = 0$.*

For some distribution families it is possible to obtain a closed-form representation for the posterior distribution and its mean. Assume, for instance, that the alternatives' true values are independent and identically distributed random variables $M_j \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ and their value estimates are $(V_j|\boldsymbol{M} = \boldsymbol{\mu}) \sim \mathrm{N}(\mu_j + \eta, \tau^2)$ for all $j = 1, \dots, n$. Then, the posterior distribution for the true value of alternative j is also normal, and the posterior mean $\mathbb{E}[M_j|\boldsymbol{V} = \boldsymbol{v}]$ is a convex combination of the prior mean $\overline{\mu}$ and the debiased value estimate $v_j - \eta$:

$$(M_j|\boldsymbol{V} = \boldsymbol{v}) \sim \mathrm{N}\left(\underbrace{\frac{\tau^2}{\sigma^2 + \tau^2}\overline{\mu} + \frac{\sigma^2}{\sigma^2 + \tau^2}(v_j - \eta)}_{=\mathbb{E}[M_j|\boldsymbol{V}=\boldsymbol{v}]}, \frac{\sigma^2\tau^2}{\sigma^2 + \tau^2}\right). \tag{12}$$

3. Estimating biases from incomplete data

To estimate total, systematic, and selection biases, one must (i) identify the shape of the joint distribution $f(\boldsymbol{\mu}, \boldsymbol{v}; \theta)$ for the alternatives' true and estimated values and (ii) estimate the parameters θ of this distribution using the observed true and estimated values of decision alternatives over multiple decision settings. We consider m decision settings, each with n_i decision alternatives, $i = 1, \dots, m$. The alternatives' true values in decision setting i are denoted by $\boldsymbol{\mu}^i = [\mu^i_1, \dots, \mu^i_{n_i}]$, and their estimated values by $\boldsymbol{v}^i = [v^i_1, \dots, v^i_{n_i}]$. While the estimated values v^i_j can be observed for each alternative $j = 1, \dots, n_i$ in all decision settings $i = 1, \dots, m$, the true values μ^i_j can only be observed for the selected alternatives $j(\boldsymbol{v}^i)$, $i = 1, \dots, m$. For this reason, standard regression techniques cannot be applied in estimating θ.

A popular approach for dealing with missing data is the Expectation Maximization (EM) algorithm (Dempster et al. 1977). In EM algorithms, the parameter values θ are obtained iteratively. In particular, given current estimates for the model parameters and the observed data (here, the estimated values of all alternatives and true values of the selected alternatives over all decision settings), the expected value of the log-likelihood function is computed by taking expectations over the missing data (here, the unobserved true values of the alternatives that were not selected). Then, new estimates for the parameters θ are computed by maximizing the expected value of the log-likelihood function computed in the previous step. These two steps are referred to as the *Expectation step* (E-step) and *Maximization step* (M-step), and they are repeated until some termination criterion is reached (e.g., the change in the estimated values of the parameters is below some tolerance value). A more formal description of the EM algorithm is given in Appendix C.

The EM algorithm works well for large sample sizes. The implementation of the algorithm is particularly straightforward for problems in which the expected value of the log-likelihood function in the E-step and the solution to the maximization problem in the M-step exist in closed form (Little & Rubin, 2014). This is the case when, for instance, the joint distribution of the alternatives' true and estimated values belongs to a regular exponential family, such as normal or log-normal distributions. Furthermore, within this family of distributions, the EM algorithm has a guaranteed convergence (Wu, 1983). Yet, if the share of missing observations is large (i.e., if the number of alternatives across decision settings is large), then convergence may be very slow. Furthermore, if the distribution $f(\boldsymbol{\mu}, \boldsymbol{v}; \theta)$ is such that the solution to the maximization problem in the M-step does not exist in closed form, then

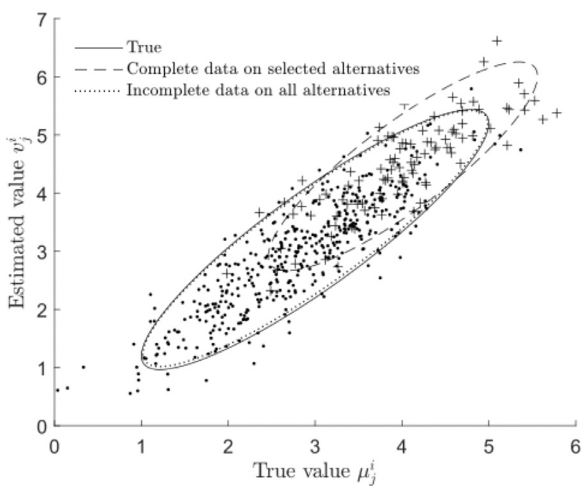


Fig. 5. True distribution (solid line) and estimated joint distributions for the alternatives' true and estimated values based on complete data on the selected alternatives alone using standard regression techniques (dashed line) and incomplete data on all alternatives using the EM algorithm (dotted line). Crosses denote selected alternatives and dots unselected alternatives.

the application of the algorithm becomes complicated. Extensions of the EM algorithm for tackling these two problems are discussed by Gelman et al. (2013) and Little and Rubin (2014). If the sample size is small, it is advisable to attach a prior distribution to the unknown parameters θ and, in the M-step, maximize the logarithm of the posterior distribution of these parameters instead of the likelihood function. This Bayesian variant of the EM algorithm is discussed in more detail by Schafer (1997).

Below we present three examples illustrating model estimation by the EM algorithm when the alternatives' true values and value estimates follow a bivariate normal distribution. In particular, we demonstrate the differences between the estimated parameter values and share of systematic bias when estimation is based on applying (i) the EM algorithm to incomplete data on all decision alternatives or (ii) standard regression techniques to complete data on selected alternatives alone. In each example and the application presented in Section 4, the EM algorithm was implemented by using Matlab on a standard laptop (2.60 GHz, 8GB memory). A detailed description of the EM algorithm for each example is given in Appendix D.

3.1. Example 1: identical decision settings

In this example we consider $m = 100$ simulated decision settings in which the number of alternatives n_i in a given setting i is uniformly distributed between 3 and 10. The alternatives' true values $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ are independent and identically distributed random variables and the value estimates are $(V_j^i|\boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$ for all $j = 1, \ldots, n_i,\ i = 1, \ldots, m$ with parameter values $\theta = [\overline{\mu}, \eta, \sigma^2, \tau^2] = [3, 0.2, 1^2, 0.5^2]$. The cumulative post-decision disappointment taken over all 100 decision settings is ca. 45 units. Based on the model parameters, $(100 \times \eta)/45 = (100 \times 0.2)/45 \approx 44\%$ of this disappointment is due to systematic bias.

Fig. 5 illustrates the simulated data as well as the true and estimated joint distributions for the alternatives' true and estimated values. In this figure, selected alternatives are marked with crosses and unselected alternatives with dots. The distributions estimated based on (i) complete data on selected alternatives using standard regression techniques and (ii) incomplete data on all alternatives using the EM algorithm are represented by dashed and dotted lines, respectively. The estimated parameter values corresponding to these distributions are $\theta_{SR} = [4.01, 0.45, 0.61^2, 0.45^2]$ and $\theta_{EM} = [3.00, 0.22, 0.99^2, 0.48^2]$. The distribution corresponding to the true parameter values $\theta = [3, 0.2, 1^2, 0.5^2]$ is represented by a solid line.

The results exemplify that using data on selected alternatives only leads to overemphasizing systematic bias. In fact, by using standard regression techniques on selected alternatives alone, 100% of the cumulative post-decision disappointment of 45 units is attributed to systematic bias ($100 \times \eta_{SR} = 100 \times 0.45 = 45$), although the true share of systematic bias was only 44%. The use of the EM algorithm on incomplete data on all alternatives results only in a slight overestimation of the share of systematic bias: $(100 \times \eta_{EM})/45 = (100 \times 0.22)/45 \approx 49\%$.

3.2. Example 2: Alternative types differ in systematic error

Next, consider a case in which the alternatives' true values $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ are independent and identically distributed random variables for all $j = 1, \ldots, n_i,\ i = 1, \ldots, m$, but the distribution of the value estimates is different for different *types* of alternatives j: $(V_j^i|\boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta_j, \tau_j^2)$. These alternative types may correspond to, for instance, different business units within a company repeatedly competing against one another for resources. In this case η_j and τ_j would represent the systematic error and the standard deviation of the random error in the value estimates provided by unit j.

Fig. 6 shows the box plots of η_j and τ_j from 1000 simulation rounds, when $\overline{\mu} = 3,\ \sigma = 1$, and data on each round is obtained from $m = 500$ simulated decision settings such that each setting i has $n_i = 6$ alternatives, each of different type. The true values of η_j and τ_j are marked in Fig. 6 with black squares, and their average estimated values obtained by using data on selected alternatives alone are marked with asterisks.

Based on Fig. 6, the average estimated values of η_j and τ_j using the EM algorithm (marked with horizontal lines) are very close to their true values (black squares). As illustrated by the gaps between the asterisks and the black squares, using standard regression techniques on selected alternatives only leads to overestimating systematic bias η_j for each alternative type. The larger the standard deviation τ_j of the random estimation error, the more systematic bias would be overestimated. For instance, the systematic bias in the value estimates of alternative type 1 would be considered to be the third largest among the six types, although these estimates are in fact unbiased. On the other hand, estimation uncertainty τ_j would be underestimated for all alternative types. A similar result is obtained for parameters $\overline{\mu}$ and σ of the prior distribution: Using the EM algorithm, the average estimated values of these parameters coincide with their true values (3.00 and 1.00, respectively), whereas the use of data on selected alternatives alone would lead to overestimating the mean true value $\overline{\mu}$ (4.13) and underestimating the standard deviation of the true value σ (0.73).

Fig. 7 shows the histogram of the estimated share of systematic bias out of cumulative post-decision disappointment over 500 decision settings, when parameters are estimated based on incomplete data on all alternatives by using the EM algorithm. The average estimated share 32.9% is close to the true share 31.8% with 95% confidence interval [21.7%, 44.1%]. As in the previous example, 100% of cumulative post-decision disappointment would be attributed to systematic bias, if parameter estimation was based on the observed true and estimated values of selected alternatives alone.

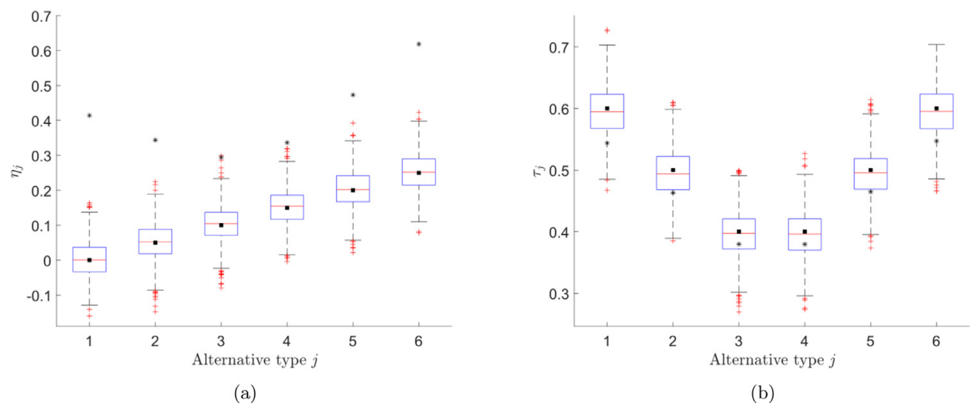


Fig. 6. Box plots for estimated values of parameters (a) η_j and (b) τ_j for alternative types $j = 1, \dots, 6$ using the EM algorithm. True parameter values marked with black squares, and average estimates using data on selected alternatives alone are marked with asterisks.

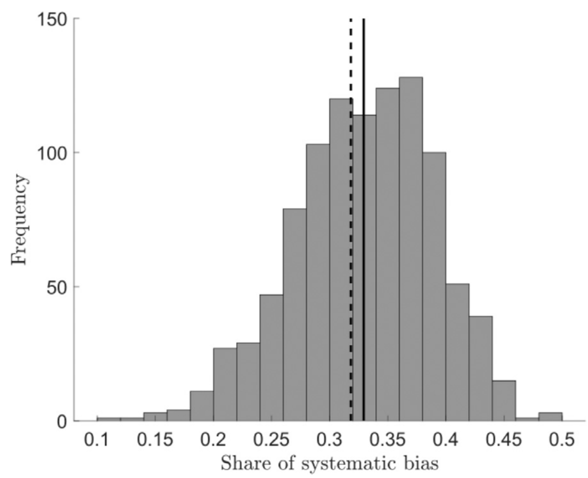


Fig. 7. Histogram of the estimated share of systematic bias. Expected true share (31.8%) marked with dashed line. Mean estimated share marked with solid line, 95% confidence interval 32.9 ± 11.2%.

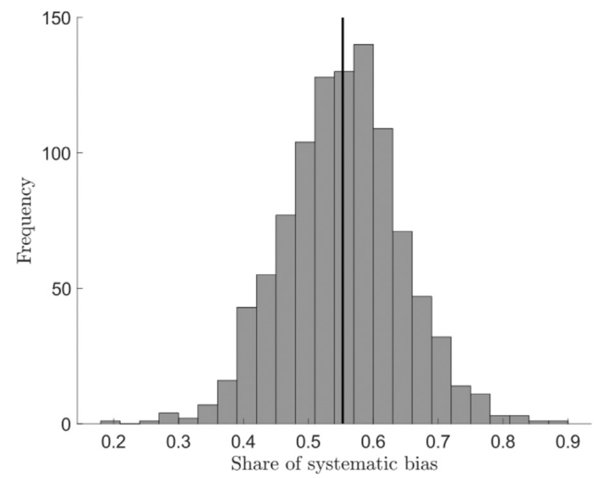


Fig. 8. Histogram of estimated share of systematic bias out of total bias. True share 55.4%; 95% confidence interval of estimated share 55.3 ± 17.8%.

3.3. Example 3: decision settings differ in average value and variability

The last example considers a case in which both systematic bias η and estimation uncertainty τ can be assumed equal for all alternatives j across all decision settings i in that $(V_j^i | M_j^i = \mu_j^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$. Yet, each decision setting is different in terms of average value and variability among the alternatives, the true values of which follow distributions $M_j^i \sim \mathrm{N}(\overline{\mu}_i, \sigma_i^2)$. Such decision settings may correspond to, for instance, a government agency awarding contracts for public works projects of different sizes and different levels of uncertainty. In this case it is not possible to distinguish between estimation error variance τ^2 and the variance in the alternatives' true values σ_i^2, unless there are some decision settings with only a single decision alternative. If sufficient data on single-alternative decision settings can be obtained, then η and τ can be estimated from these data using standard Maximum Likelihood estimation techniques, after which the EM algorithm can be used to estimate $\overline{\mu}_i$ and σ_i for $i = 1, \dots, m$.

To study the performance of the EM algorithm under the above distribution assumptions, we generated 1000 simulated data sets with 1000 decision settings each. In all decision settings, systematic bias was $\eta = 0.2$ and estimation uncertainty $\tau = 0.5$, whereas the setting-specific mean values $\overline{\mu}_i$ and standard deviations σ_i were generated from distributions $\mathrm{N}(3, 0.1^2)$ and $\mathrm{N}(1, 0.1^2)$, respectively. The number of alternatives n_i in each decision setting was uniformly distributed between 1 and 5 so that the average number of settings with a single alternative was 200 on each simulation round.

Fig. 8 illustrates the histogram of the estimated share of systematic bias out of the cumulative post-decision disappointment taken over all 1000 decision settings, when this share is estimated based on those ∼ 200 decision settings on each simulation round with a single decision alternative. The average estimated share 55.3% is very close to the true expected share 55.4% with a 95% confidence interval of [37.5%, 73.1%]. Again, if estimated based on the observed true and estimated values of selected alternatives alone, the share of systematic bias out of cumulative post-decision disappointment would be 100%.

4. Application to the procurement of highway construction projects

4.1. Problem description

In this section we apply our model to real data on highway construction projects procured by California's Department of Transportation (Caltrans). Contracts for such projects are summarized by a list of input items (e.g., asphalt or concrete), the required quantities of which are estimated by Caltrans engineers. Caltrans selects contractors to carry out the projects by sealed-bid unit-price auctions. In such auctions, a contractor submits a list of itemized unit

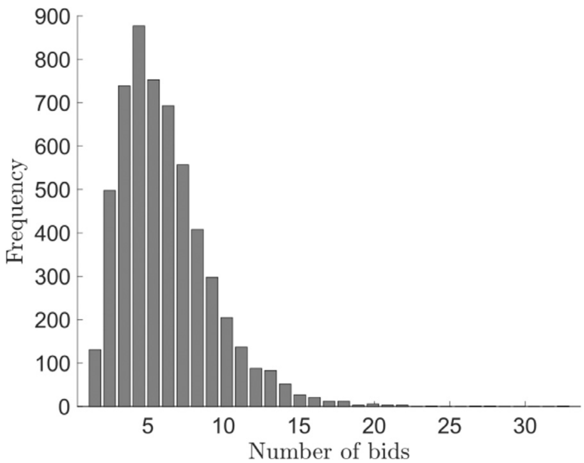


Fig. 9. Histogram of the number of bids in 5610 projects.

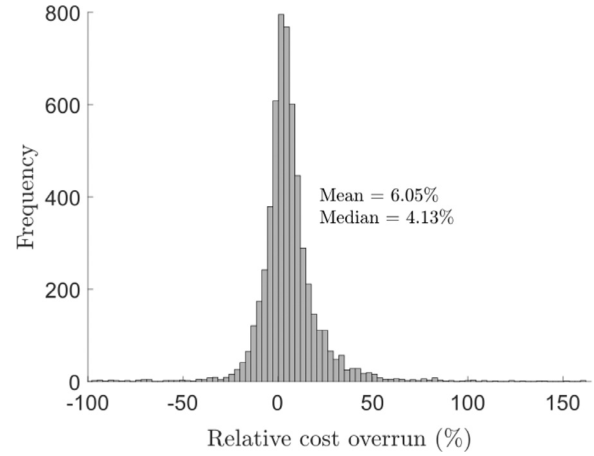


Fig. 10. Relative cost overruns $(\mu_j^i - \nu_j^i)/\nu_j^i$ in 5607 Caltrans projects. Three values left out from the figure correspond to overruns 218%, 494%, and 658%.

prices which, when multiplied by the items' estimated quantities, constitute the contractor's bid. From Caltrans' perspective, this bid can be seen as the estimated cost at which the contractor can carry out the project. Caltrans then selects the contractor (i.e., decision alternative) with the lowest bid (i.e., estimated cost ν_j^i), unless the contractor is not appropriately bonded or the bid is judged to be highly unbalanced. The true cost μ_j^i of project i from Caltrans' perspective is revealed by the final payment by Caltrans to the selected contractor after the project has been completed.

Data on such decision settings – including bids, awarded contracts, and final payment forms – can be found on Caltrans website[2] Our data set consists of 5610 decision settings between years 2000–2014. The total number of bids in these settings was 32,619. The histogram of the number of bids n_i per project (i.e., decision setting) is shown in Fig. 9. The number of bids ranges from 1 to 33 with mean 5.8 and median 5. For 131 projects there was only one bid.

The sum of the projects' estimated costs (i.e., the winning bids) was \$24.88 billion. However, the sum of the true project costs was \$27.65 billion, implying a cumulative post-decision disappointment of \$2.77 billion. Thus, the total relative cost overrun was $2.77/24.88 = 11.13\%$. The histogram of relative cost overruns across all projects is shown in Fig. 10 with mean 6.05% and median 4.13%.

There are several reasons why the projects' true costs μ_j^i do not coincide with the estimated costs ν_j^i. First, the *ex ante* estimated and *ex post* realized item quantities never perfectly agree. Also, during the project, the engineer and contractor may discover that some activities need to be added to the initial plan, resulting in a change order. Additional costs may also be incurred by the use of unanticipated materials. Finally, payment deductions can be made by Caltrans if work is not completed in time or fails to meet specifications (Bajari et al., 2011).

The above reasons can be a source of systematic bias that could explain the positive average and total cost overrun. For instance, there is evidence that contractors in unit-price auctions, anticipating changes in item quantities, set higher (lower) unit prices for those items for which they expect quantity overruns (underruns). In this way, contractors can increase their expected profits without increasing their bid (Gupta et al., 2015). Also, project scope may be systematically underestimated by Caltrans engineers, resulting in repeated change orders that increase costs. Nevertheless, at least a part of the cost overruns is likely to have resulted from selection bias, because in 86% of projects the contractor with the lowest bid was selected, and in 97% of projects the bid of the selected contractor was among the two lowest.

[2] Our data set includes only those settings found at http://www.dot.ca.gov/hq/esc/oe/contract_awards_services.html which include information on all bids, awarded contracts, and final payments. Furthermore, we have excluded bids marked as 'irregular' on the website.

4.2. Modeling the projects' true and estimated costs

Fig. 11a shows the histogram of the standardized logarithms of the contractors' bids across all 4981 projects with at least three bids. Together with the Kolmogorov-Smirnov test for standard normal distribution ($p = 3.1 \cdot 10^{-20}$), this figure suggests that the bids $\boldsymbol{\nu}^i$ for each project i can be modeled as realizations of independent and identically distributed log-normal random variables. Assuming that the distribution of the bids serves as a reasonable proxy for the distribution of the true costs $\boldsymbol{\mu}^i$, we model the true costs as realizations of random variables

$$\boldsymbol{M}^i = [M_1^i, \dots, M_{n_i}^i], \quad M_j^i \sim \mathrm{LogN}(\overline{\mu}_i, \sigma_i^2), \; j = 1, \dots, n_i,$$

where parameters $\overline{\mu}_i$ and σ_i^2 are to be estimated from data.

Fig. 11b shows the histogram of the differences between the logarithms of the alternatives' estimated and true costs $(\ln \nu_j^i - \ln \mu_j^i)$ across those 131 projects with a single bid. Based on this figure and the Anderson-Darling normality test ($p = 5 \cdot 10^{-4}$) we model these differences as realizations of independent and identically distributed normal random variables. Assuming that this holds also in those projects i for which there were multiple bids, the bids ν_j^i can be modelled as realizations of random variables

$$(V_j^i | M_j^i = \mu_j^i) \sim \mathrm{LogN}(\ln \mu_j^i + \eta, \tau^2), \; j = 1, \dots, n_i, \tag{13}$$

where parameters η and τ^2 are to be estimated from data.

4.3. Model estimation

The EM algorithm was used to estimate the unknown model parameters η, τ^2, $\overline{\mu}_i$, and σ_i^2 across projects $i = 1, \dots, 5610$. A detailed description of this algorithm is given in Appendix E. The estimated values of η and τ are -0.0347 and 0.1264, respectively. The estimated values of $\overline{\mu}_i$ and σ_i across all projects range within $[11.0, \; 20.9]$ and $[0, \; 3.26]$, respectively.

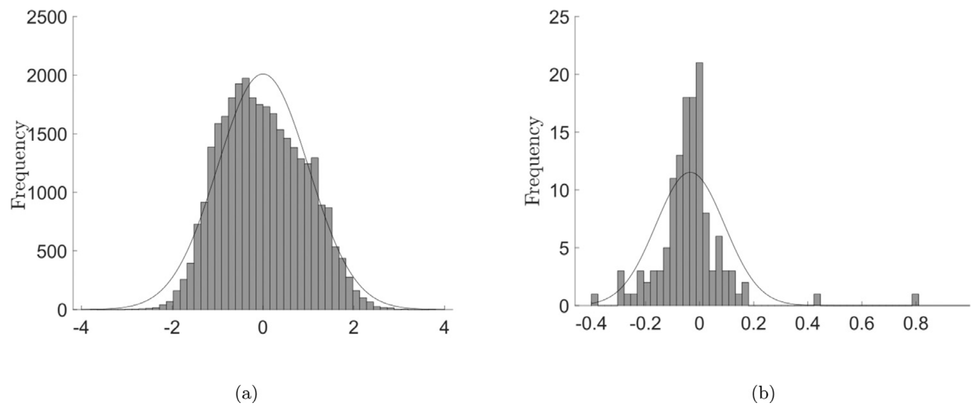


Fig. 11. Histograms of (a) standardized logarithms $(\ln v_j^i - \mathrm{Mean}_j[\ln v_j^i])/\mathrm{Std}_j[\ln v_j^i]$ of the contractors' bids across projects with at least three bids and (b) differences $\ln v_1^i - \ln \mu_1^i$ across 131 projects with a single bid.

Debiased costs estimates $\tilde{v}_j^i$ (cf. Definition 1) are obtained from bids v_j^i through

$$\tilde{v}_j^i = d(v_j^i) = \underbrace{\exp\left(-\eta - \frac{\tau^2}{2}\right)}_{\approx 1.027} v_j^i. \tag{14}$$

It is straightforward to verify that these estimates are conditionally unbiased (i.e., $\mathbb{E}[d(V_j^i)|M_j^i = \mu_j^i] = \mu_j^i$), because each random variable $(V_j^i|M_j^i = \mu_j^i)$ follows the log-normal distribution (13) the mean of which is $\exp(\eta + \frac{\tau^2}{2})\mu_j^i$. In practice, result (14) suggests that to remove systematic bias from the bids, each of them should be increased by 2.7%.

The posterior means $\hat{v}_j^i$ of the true costs can be computed from bids v_j^i through

$$\hat{v}_j^i = \mathbb{E}[M_j^i|V_j^i = v_j^i] = \left[v_j^i \cdot \exp(-\eta)\right]^{\alpha_i} \cdot \left[\exp\left(\overline{\mu}_i + \frac{1}{2}\sigma_i^2\right)\right]^{1-\alpha_i}, \tag{15}$$

where $\alpha_i = \sigma_i^2/(\sigma_i^2 + \tau^2)$. Result (15) follows directly from the properties of the log-normal distribution of $(M_j^i|V_j^i = v_j^i)$ (see Lemma 1 in Appendix B). To have an interpretation for (15), note from (13) that the logarithm of the bid is normally distributed $(\ln V_j^i|M_j^i = \mu_j^i) \sim \mathrm{N}(\ln \mu_j^i + \eta, \tau^2)$, whereby $\ln v_j^i - \eta$ is a conditionally unbiased estimate for $\ln \mu_j^i$. Thus, the posterior mean (15) is a weighted geometric mean of the estimated value of $\mu_j^i = \exp(\ln \mu_j^i)$ – that is, $\exp(\ln v_j^i - \eta) = v_j^i \cdot \exp(-\eta)$ – and its prior expected value $\mathbb{E}[M_j^i] = \exp(\overline{\mu}_i + \sigma_i^2/2)$. The more uncertain the cost estimates (reflected by a large value of τ compared to σ_i), the larger the weight $1 - \alpha_i = \tau^2/(\sigma_i^2 + \tau^2)$ given to prior information.

4.4. Systematic and selection biases

The total true cost of all 5610 projects was \$27.65 billion, exceeding the estimated total cost (i.e., the sum of winning bids) of \$24.88 billion by \$2.77 billion. Table 1 shows how the cumulative post-decision disappointment of \$2.77 billion is divided into shares explained by systematic and selection biases.

If selection had been based on debiased cost estimates $\tilde{v}_j^i$, then the cumulative disappointment would have reduced to \$2.10 billion, whereby our model suggests that $(2.77 - 2.10)/2.77 \approx 24\%$ of the cumulative disappointment is due to systematic bias. Using posterior means $\hat{v}_j^i$, the cumulative disappointment would have reduced further down to \$250 million, implying that $(2.10 - 0.25)/2.77 \approx 67\%$ of this disappointment is due to selection bias. The remaining $0.25/2.77 \approx 9\%$ is unexplained by our model, whereby the share of explained variation is 91%.

Table 1
Shares of systematic and selection biases in generating post-decision disappointment.

Source of disappointment	\$B	%	95% CI
Systematic bias	0.67	24	[0,45]
Selection bias	1.85	67	[49,82]
Unexplained	0.25	9	[3,20]
Total	2.77	100	

To obtain confidence intervals for the shares of systematic and selection biases, we applied basic bootstrapping with 1000 samples (Davison & Hinkley, 1997). In particular, we first (i) generated a bootstrapping sample from the set of 131 projects with a single bid to estimate the values of η and τ and then, by using these parameter values, (ii) estimated the values of $\overline{\mu}_i$ and σ_i for each project $i = 1, \ldots, 5610$. The resulting 95% confidence intervals for the shares of systematic and selection biases are shown in the right-most column of Table 1. These intervals are relatively wide, but nevertheless demonstrate that a significant share of post-decision disappointment can be assumed to have been generated by selection bias.

4.5. Implications for decision support

The use of posterior means $\hat{v}_j^i$ instead of estimated costs v_j^i as a basis for contractor selection would have decreased the relative cumulative disappointment from $2.77/24.88 \approx 11.1\%$ to $0.25/27.39 \approx 0.9\%$, and the mean setting-specific disappointment from $\$2.77 \cdot 10^9/5610 \approx \$494,000$ to $\$0.25 \cdot 10^9/5610 \approx \$45,000$. Yet, the parameters required to compute the posterior means $\hat{v}_j^i$ through Eq. (15) can only be estimated after the true project costs have been observed, whereby they cannot be utilized for decision support at the time of selecting the contractor. In particular, even if parameters η and τ related to systematic bias would have been estimated based on past projects, it is impossible to obtain estimates for the project-specific parameters $\overline{\mu}_i$ and σ_i^2 *ex ante*. These parameters are required to compute the prior mean $\mathbb{E}[M_j^i] = \exp(\overline{\mu}_i + \sigma_i^2/2)$ for the true cost of carrying out project i, as well as the weighting coefficient $\alpha_i = \sigma_i^2/(\sigma_i^2 + \tau^2)$ for the weighted geometric mean in Eq. (15).

One way to circumvent this issue is to use Caltrans' engineers' estimates for the total cost of each project as proxies for the prior means $\mathbb{E}[M_j^i]$. These estimates are obtained by multiplying the en-

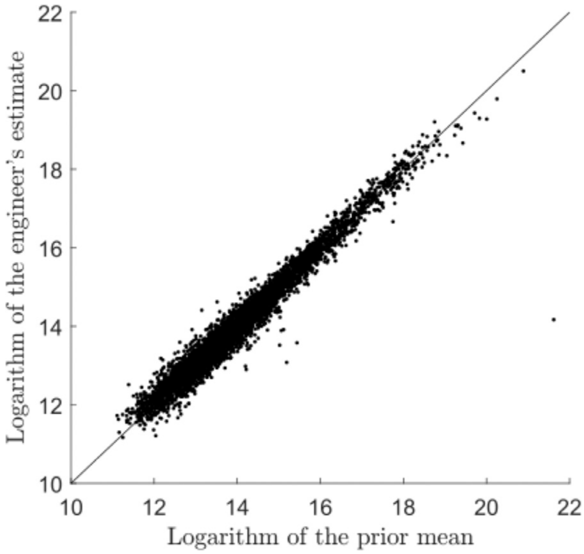


Fig. 12. Scatter plot of the logarithms of the projects' prior mean costs $\mathbb{E}[M^i_j]$ and the engineers' estimates $\overline{\mu}^i_e$ of these costs. Regression line without intercept: $\overline{\mu}^i_e = 1.0000 \cdot \mathbb{E}[M^i_j]$, $R^2 = 0.963$.

gineer's estimates on item quantities required to carry out a given project by the item-specific prices that are published in the Contract Cost Data Book[3] prepared annually by Caltrans. Fig. 12 shows the scatter plot between the logarithms of the project-specific prior mean costs $\mathbb{E}[M^i_j]$ and the logarithms of the engineers' estimates for these costs, denoted by $\overline{\mu}^i_e$. The regression line between these two variables without intercept is $\ln \overline{\mu}^i_e = 1.0000 \cdot \ln \mathbb{E}[M^i_j]$ with $R^2 = 0.963$. Based on the regression model and Fig. 12, the engineer's estimate $\overline{\mu}^i_e$ gives quite accurate information about the average true cost with which the contractors would be able to carry out project i.

To estimate the weighting coefficient α_i in Eq. (15), we note that the logarithm of the bid is normally distributed: $\ln V^i_j \sim \mathrm{N}(\overline{\mu}_i + \eta, \tau^2 + \sigma_i^2)$, $\quad j = 1, \ldots, n_i$. Thus, we can estimate the project-specific cost variability as $\hat{\sigma_i}^2 = \mathrm{Var}[\ln v^i_j] - \tau^2$, where $\mathrm{Var}[\ln v^i_j]$ is the sample variance of the logarithms of the bids for project i. Hence, the weight α_i can be estimated by $\hat{\alpha}_i = \hat{\sigma}_i^2/(\hat{\sigma}_i^2 + \tau^2) = 1 - \tau^2/\mathrm{Var}[\ln v^i_j]$. If the sample variance $\mathrm{Var}[\ln v^i_j]$ of the logarithms of the bids for project i is smaller than τ^2, then the true cost of carrying out this project is assumed to be the same across bidders and to coincide with the engineer's estimate $\overline{\mu}^i_e$ so that $\hat{\alpha}_i = 0$. Finally, if there is a single bid for project i so that sample variance $\mathrm{Var}[\ln v^i_j]$ cannot be computed, then the true cost of contractor j to carry out this project is assumed to be the estimated value $v^i_j \cdot \exp(-\eta)$ of true cost μ^i_j based on bid v^i_j, whereby $\hat{\alpha}_i = 1$. Modifying Eq. (15) accordingly, we define the *engineer's mean* e^i_j for the true cost of contractor j to carry out project i as

$$e^i_j = \left[v^i_j \cdot \exp(-\eta)\right]^{\hat{\alpha}_i} \cdot \left[\overline{\mu}^i_e\right]^{1-\hat{\alpha}_i}, \tag{16}$$

where $\hat{\alpha}_i = \min\{0, 1 - \tau^2/\mathrm{Var}[\ln v^i_j]\}$ for projects with multiple bids and $\hat{\alpha}_i = 1$ for projects with a single bid.

In the Caltrans case, the engineers' means work relatively well in mitigating post-decision disappointment. This is illustrated by Table 2, which shows the absolute cumulative, relative cumulative, and mean setting-specific disappointment based on bids v^i_j, engineers' means e^i_j, and posterior means $\hat{v}^i_j$. While the use of posterior means instead of bids to estimate the projects' true costs would have reduced each measure of disappointment by roughly 90%, even the use of engineers' means would have managed to eliminated close to 80% of the disappointment.

Table 2
Post-decision disappointment based on bids v^i_j, engineers' means e^i_j, and posterior means $\hat{v}^i_j$ across projects $i \in \{1, \ldots, 5610\}$.

Measure of disappointment	Bids v^i_j	Engineers' means e^i_j	Posterior means $\hat{v}^i_j$
Cumulative ($B)	2.77	0.63	0.25
Relative cumulative (%)	11.1	2.3	0.9
Mean ($k)	494	113	45

5. Discussion and conclusions

To date, two kinds of explanations have been proposed for the empirically observed disappointment in the *ex post* realized values of selected decision alternatives: (i) a systematic bias in the alternatives' value estimates resulting from unjustified optimism or even deliberate misrepresentation (Bajari et al., 2011; Flyvbjerg, 2009), and (ii) selection bias resulting from the higher probability of the values of those alternatives with the highest value estimates to have been overestimated (Quirk & Terasawa, 1986; Smith & Winkler, 2006; Vilkkumaa et al., 2014). This paper presents the first contribution to understanding how both systematic and selection biases generate post-decision disappointment in real-life applications. In particular, the paper presents models through which the relative magnitudes of systematic and selection biases can be estimated using data that contains the estimated values of all alternatives, but the realized values of only those alternatives that were selected.

The focus of this paper has been on studying the average contributions of systematic and selection biases in generating post-decision disappointment from the decision-maker's point of view in recurring decision settings that are at least to some extent comparable. For this reason, our models for capturing the relationship between the decision alternatives' true and estimated values do not distinguish between different sources of systematic bias, although several such sources admittedly exist; including the evaluator's personal disposition towards overoptimism (Lovallo & Kahneman, 2003; Montibeller & von Winterfeldt, 2015), organization-specific incentives and consequences for strategic misrepresentation (Oliva & Watson, 2009), or game-theoretic considerations related to auction-based decision settings (Gupta et al., 2015). Nor do we consider the effects of selection bias from the point of view of the proposers of the decision alternatives, such as the so-called winner's curse suffered by the winning bidder in auction-based selection processes (Capen, Clapp, & Campbell, 1971; Thaler, 1992), or behavioural explanations for non-optimal bidding behavior in first-price sealed-bid auctions (Wang & Guo, 2017).

Another limitation of the models used in our simulation examples as well as the Caltrans case is the assumption that both the alternatives' true and estimated values are independent. Yet, Smith and Winkler (2006) show that the covariance structures between the true and estimated values significantly affect the magnitude of selection bias. In particular, selection bias increases in the correlation between the alternatives' true values and decreases in the correlation between the value estimates. The estimation of covariance structures is, however, impossible in cases where a single alternative is selected in each decision setting. Data from situations in which a portfolio of several alternatives was selected could be used to estimate covariance between the alternatives' true and estimated values (Kettunen & Salo, 2017). Moreover, estimates for the

[3] http://www.dot.ca.gov/hq/esc/oe/awards/

shares of systematic and selection biases obtained from such data would be more accurate, because the share of missing observations would be lower compared to cases in which a single alternative was selected. The acquisition of data from portfolio selection processes and the development of methods to use these data for estimating the covariance structures between the alternatives' true and estimated values provide interesting research topics for the future. Decisions on investments to financial assets such as stocks would offer a particularly suitable context for testing these kinds of methods, because (i) the true and estimated values of the assets are likely to be correlated and (ii) the true values of those assets to which investments were not made can be observed *ex post*.

The implications of this paper for managers and decision-makers are threefold. First, it is important to raise awareness of the role of selection bias in generating post-decision disappointment. Otherwise, the average negative gap between the realized and estimated values of selected alternatives may be entirely attributed to systematic bias in the value estimates. To mitigate systematic bias, measures such as financial, professional, or even criminal penalties have been suggested on forecasters who consistently provide biased estimates (Flyvbjerg et al., 2002). If, however, only a small share of post-decision disappointment is caused by systematic bias (e.g., 24% in the case of highway construction projects procured by Caltrans), such measures are likely to be ineffective at best and demoralizing at worst. The use of penalties can be particularly troublesome in the context of multicriteria decision problems: post-decision disappointment is probably most likely with respect to those criteria on which performance assessments are the least accurate, leaving those responsible for producing such assessments in a disadvantaged position. A thorough examination of the mechanisms of post-decision disappointment in the multicriteria context – including the potential of multicriteria performance measures to alleviate overall disappointment – provides a fruitful avenue for future research.

Second, careful records of past decision settings should be kept and utilized in decision-making to obtain more realistic expectations about the decision alternatives. To be able to distinguish between systematic and selection biases, data should be collected not only on the estimated and realized values of selected alternatives, but also on the estimated values of those alternatives that were not implemented. With the help of models developed here, such data can be used to determine (i) how much the mitigation of systematic bias through, e.g., sanctions or optimal contracting arrangements can be expected to reduce post-decision disappointment and, even if systematic bias could be completely eliminated, (ii) how much the estimated value of the selected alternative should be adjusted to avoid such disappointment. A related call for the use of so-called behavioural databases for debiasing efforts has recently been made by Durbach and Montibeller (2019).

Finally, to enable the mitigation of post-decision disappointment *ex ante* in cases where little or no data exists about the performance of similar alternatives in the past, it is advisable to search for alternative means such as expert evaluations to reliably estimate the average future performance of the decision alternatives (Jørgensen, 2007; Oliva & Watson, 2009). Expert judgment is commonly used to adjust statistical forecasts in contexts such as inventory control, production planning, purchasing, supply chain management, and cash flow planning (De Baets & Harvey, 2020; Petropoulos, Fildes, & Goodwin, 2016). When done right, these adjustments may outperform the forecasting accuracy of statistical methods alone (Davydenko & Fildes, 2013; Petropoulos et al., 2016). Encouraging evidence on the possibilities of using expert judgement to adjust other kinds of value estimates besides statistical forecasts is given by our case study on the procurement of highway construction projects: the engineer's estimate for the cost of a project was shown to be highly accurate in estimating the average true cost at which a contractor would be able to carry out the project. Whether similar approaches work in other types of decision contexts as well remains an interesting topic for future work.

Acknowledgments

The authors thank Robert G. Dyson and three anonymous referees for their excellent suggestions to help improve the paper. The authors would also like to thank Peng Xu for his help with collecting data for the application. This research has been supported by the Academy of Finland (grant number 323800) and The Foundation for Economic Education (grant number 16–9442).

Appendix A. Notation

Symbol	Description
$i \in \{1,\dots,m\}$	Index of decision setting
$j \in \{1,\dots,n_{(i)}\}$	Index of a decision alternative (in setting i)
$\boldsymbol{M}^{(i)} = [M_1^{(i)},\dots,M_n^{(i)}]^T$	Random true values/costs of the decision alternatives (in setting i)
$\boldsymbol{\mu}^{(i)} = [\mu_1^{(i)},\dots,\mu_n^{(i)}]^T$	Realized true values/costs of the decision alternatives (in setting i)
$\boldsymbol{V}^{(i)} = [V_1^{(i)},\dots,V_n^{(i)}]^T$	Random value/cost estimates of the decision alternative (in setting i)
$\boldsymbol{v}^{(i)} = [v_1^{(i)},\dots,v_n^{(i)}]^T$	Realized value/cost estimates of the decision alternatives (in setting i)
$\tilde{\boldsymbol{V}}^{(i)} = [\tilde{V}_1^{(i)},\dots,\tilde{V}_n^{(i)}]^T$	Random debiased value/cost estimates of the decision alternatives (in setting i)
$\tilde{\boldsymbol{v}}^{(i)} = [\tilde{v}_1^{(i)},\dots,\tilde{v}_n^{(i)}]^T$	Realized debiased value/cost estimates of the decision alternatives (in setting i)
$\hat{\boldsymbol{V}}^{(i)} = [\hat{V}_1^{(i)},\dots,\hat{V}_n^{(i)}]^T$	Random posterior means of the decision alternatives (in setting i)
$\hat{\boldsymbol{v}}^{(i)} = [\hat{v}_1^{(i)},\dots,\hat{v}_n^{(i)}]^T$	Realized posterior means of the decision alternatives (in setting i)
$f(\boldsymbol{\mu})$	Distribution of the true values
$f(\boldsymbol{v}\mid\boldsymbol{\mu})$	Distribution of the value estimates given the true values
$f(\boldsymbol{\mu},\boldsymbol{v};\boldsymbol{\theta})$	Joint distribution of true and estimated values given parameters $\boldsymbol{\theta}$
$d(\cdot)$	Debiasing transformation
$J^* = j(\boldsymbol{M})$	Random index of the decision alternative with the highest true value
$j(\boldsymbol{\mu})$	Realized index of the decision alternative with the highest true value
$J = j(\boldsymbol{V})$	Random index of the decision alternative with the highest value estimate
$j(\boldsymbol{v})$	Realized index of the decision alternative with the highest value estimate
$\tilde{J} = j(\tilde{\boldsymbol{V}})$	Random index of the decision alternative with the highest debiased value estimate
$j(\tilde{\boldsymbol{v}})$	Realized index of the decision alternative with the highest debiased value estimate
$\hat{J} = j(\hat{\boldsymbol{V}})$	Random index of the decision alternative with the highest posterior mean
$j(\hat{\boldsymbol{v}})$	Realized index of the decision alternative with the highest posterior mean
TB	Total bias
SeB	Selection bias
SyB	Systematic bias
Application-specific notation	
$\mathrm{LogN}(\overline{\mu}_i,\ \sigma_i^2)$	Distribution for the contractors' true costs of carrying out project i
$\mathrm{LogN}(\ln\mu_j^i + \eta, \tau^2)$	Distribution for the bid of contractor j in project i
$\overline{\mu}_e^i$	Engineer's estimate for the total cost of project i
$\hat{\sigma}_i^2$	Estimated cost variability for project i
e_j^i	Engineer's mean for the true cost of contractor j to carry out the project

Appendix B. Proofs

Proof of Proposition 1. Let us fix $\boldsymbol{\mu}$. Then, $j^* = j(\boldsymbol{\mu})$ (the alternative with the highest true value) is fixed and $\tilde{J}$ (the alternative with the highest debiased value estimate) is a random variable. For any $\tilde{\boldsymbol{V}}$, we have

$$\mu_{\tilde{J}} - \tilde{V}_{\tilde{J}} \leq \mu_{j^*} - \tilde{V}_{\tilde{J}} \leq \mu_{j^*} - \tilde{V}_{j^*}, \tag{B.1}$$

where the first inequality follows from the definition of j^*, and the second from the definition of $\tilde{J}$. Taking expectations of (B.1) conditioned on $\boldsymbol{M} = \boldsymbol{\mu}$ and integrating over the uncertainty regarding the value estimates, we have

$$\mathbb{E}[\mu_{\tilde{J}} - \tilde{V}_{\tilde{J}} | \boldsymbol{M} = \boldsymbol{\mu}] \leq [\mu_{j^*} - \tilde{V}_{\tilde{J}} | \boldsymbol{M} = \boldsymbol{\mu}] \leq \mathbb{E}[\mu_{j^*} - \tilde{V}_{j^*} | \boldsymbol{M} = \boldsymbol{\mu}] = 0, \tag{B.2}$$

where the last equality follows from the conditional unbiasedness of the debiased estimates $\tilde{V}$. Because (B.1) holds for all $\boldsymbol{\mu}$, integrating over uncertain $\boldsymbol{M}$ yields $SeB = \mathbb{E}[M_{\tilde{J}} - \tilde{V}_{\tilde{J}}] \leq 0$. If there is no possibility of selecting a non-optimal alternative (i.e., $\mathbb{P}(J^* = \tilde{J}) = 1$), then the inequalities in (B.1) and (B.2) become equalities. If a non-optimal alternative is selected, then the first inequality in (B.1) will be strict. Thus, if there is some chance that this happens (i.e., $\mathbb{P}(J^* \neq \tilde{J}) > 0$), then the first inequality in (B.2) will also be strict so that $SeB = \mathbb{E}[M_{\tilde{J}} - \tilde{V}_{\tilde{J}}] < 0$. □

Proof of Proposition 2. Assume that there is a single decision alternative with value estimate $(V|M = \mu)$. The debiased estimate $\tilde{V}$ of this alternative is such that $\mathbb{E}[\tilde{V}|M = \mu] = \mu$. Thus, $\mathbb{E}[M - \tilde{V}|M = \mu] = \mu - \mathbb{E}[\tilde{V}|M = \mu] = 0$ for all μ. Integrating over uncertain μ and taking expectations, it follows that $SeB = \mathbb{E}[M - \tilde{V}] = 0$. Because $SyB = TB - SeB$, it follows that $TB = SyB$. □

Proof of Proposition 3. Because $\mathbb{P}(J = \tilde{J}) = 1$, $SyB = \mathbb{E}[(M_J - V_J) - (M_J - \tilde{V}_J)] = \mathbb{E}[\tilde{V}_J - V_J] \leq 0$, where the inequality follows from the assumption that $\mathbb{P}(\tilde{V}_J > V_J) = 0$. If $\mathbb{P}(\tilde{V}_J = V_J) = 1$, then $SyB = \mathbb{E}[\tilde{V}_J - V_J] = 0$. □

Proof of Corollary 1. The result follows directly from Propositions 1 and 3. □

Proof of Proposition 4. This Proposition can be established by using the proof that Smith and Winkler (2006) provide for their second proposition. Although they generally assume that the alternatives' value estimates are conditionally unbiased, their proof does not make use of this assumption. □

Lemma 1. *Let* $M_j^i \sim \mathrm{LogN}(\overline{\mu}_i, \sigma_i^2)$ *and* $(V_j^i | M_j^i = \mu_j^i) \sim \mathrm{LogN}(\ln \mu_j^i + \eta, \tau^2)$. *Then,*

$$(M_j | V_j = v_j) \sim \mathrm{LogN}\left(\frac{\tau_j^2}{\sigma_i^2 + \tau_j^2}\overline{\mu}_j + \frac{\sigma_j^2}{\sigma_j^2 + \tau_j^2}(\ln v_j - \eta_j), \frac{\sigma_j^2 \tau_j^2}{\sigma_j^2 + \tau_j^2}\right).$$

Proof of Lemma 1. Because $\ln M_j \sim \mathrm{N}(\overline{\mu}_j, \sigma_j^2)$ and $(\ln V_j | \ln M_j = \ln \mu_j) \sim \mathrm{N}(\ln \mu_j + \eta_j, \tau_j^2)$, it follows from the properties of bivariate normal distribution that

$$\begin{aligned}
&(\ln M_j | \ln V_j = \ln v_j) \\
&\sim \mathrm{N}\left(\frac{\tau_j^2}{\sigma_i^2 + \tau_j^2}\overline{\mu}_j + \frac{\sigma_j^2}{\sigma_j^2 + \tau_j^2}(\ln v_j - \eta_j), \frac{\sigma_j^2 \tau_j^2}{\sigma_j^2 + \tau_j^2}\right) \\
&\Leftrightarrow (M_j | V_j = v_j) \\
&\sim \mathrm{LogN}\left(\frac{\tau_j^2}{\sigma_i^2 + \tau_j^2}\overline{\mu}_j + \frac{\sigma_j^2}{\sigma_j^2 + \tau_j^2}(\ln v_j - \eta_j), \frac{\sigma_j^2 \tau_j^2}{\sigma_j^2 + \tau_j^2}\right).
\end{aligned} \tag{B.3}$$

Appendix C. Formal description of the EM algorithm

Let use denote by $j^i = j(\boldsymbol{v}^i)$ the index of the selected alternative in decision setting $i = 1, \ldots, m$. Moreover, let us denote by $v = \{\boldsymbol{v}^i\}_{i=1}^m$ the estimated values of all alternatives, by $\mu = \{\mu_{j^i}^i\}_{i=1}^m$ the true values of the selected alternatives, and by $M = \{M_j^i\}_{i=1, j \neq j^i}^m$ the unobserved true values of the alternatives that were not selected across all decision settings $i = 1, \ldots, m$. The EM algorithm proceeds as follows.

1. Let the initial values for parameters be $\theta = \hat{\theta}_1$. Set $k = 1$ and select a tolerance value δ.
2. **Expectation step**: Assuming that $\theta = \hat{\theta}_k$, compute the expected value $g(\theta) := \mathbb{E}[\log L(\theta | \mu, M, v)]$ of the log-likelihood function, where expectation is taken over the missing true values M.
3. **Maximization step**: Using the expected value $g(\theta)$ of the log-likelihood function computed in Step 2, obtain the maximum likelihood estimate $\hat{\theta}_{k+1}$:

   $$\hat{\theta}_{k+1} = \underset{\theta}{\operatorname{argmax}}\, g(\theta).$$

   Set $k \leftarrow k + 1$ and go back to Step 2.
4. Repeat the expectation and maximization steps 2 and 3 until $\|\hat{\theta}_{k+1} - \hat{\theta}_k\| < \delta$.

Appendix D. Estimation of model parameters in Examples 1–3

Let $N = \sum_{i=1}^m n_i$, and let j^i be the index of the selected alternative in decision setting i.

Example 1: The log-likelihood function for $\theta = [\overline{\mu}, \eta, \sigma^2, \tau^2]$ is

$$\begin{aligned}
\mathcal{L}(\theta | \mu, M, v) = &-N \ln(2\pi\sigma\tau) - \\
&-\frac{1}{2}\sum_{i=1}^m \left[\left(\frac{v_{j^i}^i - \mu_{j^i}^i - \eta}{\tau}\right)^2 + \left(\frac{\mu_{j^i}^i - \overline{\mu}}{\sigma}\right)^2 \right. \\
&\left. + \sum_{j \neq j^i}\left[\left(\frac{v_j^i - M_j^i - \eta}{\tau}\right)^2 + \left(\frac{M_j^i - \overline{\mu}}{\sigma}\right)^2\right]\right].
\end{aligned}$$

Expectation step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ and $(V_j^i | \boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$, it follows from the properties of bivariate normal distribution that

$$\begin{aligned}
\widehat{\mu_j^i} &= \mathbb{E}[M_j^i | v_j^i; \theta] = \frac{\tau^2}{\sigma^2 + \tau^2}\overline{\mu} + \frac{\sigma^2}{\sigma^2 + \tau^2}(v_j^i - \eta) \qquad \forall j \neq j^i, \\
\widehat{(\mu_j^i)^2} &= \mathbb{E}[(M_j^i)^2 | v_j^i; \theta] = \frac{\sigma^2 \tau^2}{\sigma^2 + \tau^2} \\
&\quad + \left[\frac{\tau^2}{\sigma^2 + \tau^2}\overline{\mu} + \frac{\sigma^2}{\sigma^2 + \tau^2}(v_j^i - \eta)\right]^2 \quad \forall j \neq j^i.
\end{aligned}$$

Maximization step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ and $(V_j^i | \boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$, the Maximum Likelihood estimates for the model parameters $\theta = [\overline{\mu}, \eta, \sigma^2, \tau^2]$ are

$$\begin{aligned}
\hat{\overline{\mu}}_{k+1} &= \frac{1}{N}\sum_{i=1}^m \sum_{j=1}^{n_i} \mu_j^i, \\
\hat{\sigma}_{k+1}^2 &= \frac{1}{N}\sum_{i=1}^m \sum_{j=1}^{n_i} \left((\mu_j^i)^2 - \hat{\overline{\mu}}_{k+1}^2\right),
\end{aligned} \tag{D.1}$$

$$\begin{aligned}
\hat{\eta}_{k+1} &= \frac{1}{N}\sum_{i=1}^m \sum_{j=1}^{n_i} v_j^i - \hat{\overline{\mu}}_{k+1}, \\
\hat{\tau}_{k+1}^2 &= \frac{1}{N}\sum_{i=1}^m \sum_{j=1}^{n_i} \left((v_j^i)^2 - 2v_j^i \mu_j^i + (\mu_j^i)^2 - \hat{\eta}_{k+1}^2\right).
\end{aligned} \tag{D.2}$$

Initial values: Suitable initial values for σ^2, η, and τ^2 are obtained by applying (D.1) and (D.2) for the selected alternatives alone, and for $\overline{\mu}$ by noting that $\hat{\overline{\mu}}_{k+1} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} v_j^i - \hat{\eta}_{k+1}$ in (D.2):

$$\hat{\overline{\mu}}_1 = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} v_j^i - \hat{\eta}_1, \quad \hat{\sigma}_1^2 = \frac{1}{m}\sum_{i=1}^{m}\left((\mu_{j^i}^i)^2 - \left(\frac{1}{m}\sum_{i=1}^{m}\mu_{j^i}^i\right)^2\right),$$

$$\hat{\eta}_1 = \frac{1}{m}\sum_{i=1}^{m}\left(v_{j^i}^i - \mu_{j^i}^i\right), \quad \hat{\tau}_1^2 = \frac{1}{m}\sum_{i=1}^{m}((v_{j^i}^i - \mu_{j^i}^i)^2 - \hat{\eta}_1^2).$$

Example 2: The log-likelihood function for $\theta = \{\overline{\mu}, \sigma^2\} \cup \{\eta_j, \tau_j^2\}_{j=1}^{n_i}$ is

$$\begin{aligned}\mathcal{L}(\theta|\mu, M, v) &= -N\ln(2\pi\sigma) - m\sum_{j=1}^{n_i}\ln\tau_j \\ &- \frac{1}{2}\sum_{i=1}^{m}\left[\left(\frac{v_{j^i}^i - \mu_{j^i}^i - \eta_{j^i}}{\tau_{j^i}}\right)^2 + \left(\frac{\mu_{j^i}^i - \overline{\mu}}{\sigma}\right)^2\right. \\ &+ \left.\sum_{j\neq j^i}\left[\left(\frac{v_j^i - M_j^i - \eta_j}{\tau_j}\right)^2 + \left(\frac{M_j^i - \overline{\mu}}{\sigma}\right)^2\right]\right].\end{aligned}$$

Expectation step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ and $(V_j^i|\boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta_j, \tau_j^2)$, it follows from the properties of bivariate normal distribution that:

$$\widehat{\mu_j^i} := \mathbb{E}[M_j^i|v_j^i; \theta] = \frac{\tau_j^2}{\sigma^2 + \tau_j^2}\overline{\mu} + \frac{\sigma^2}{\sigma^2 + \tau_j^2}(v_j^i - \eta_j) \qquad \forall j \neq j^i,$$

$$\begin{aligned}\widehat{(\mu_j^i)^2} &:= \mathbb{E}[(M_j^i)^2|v_j^i; \theta] = \frac{\sigma^2\tau_j^2}{\sigma^2 + \tau_j^2} \\ &+ \left[\frac{\tau_j^2}{\sigma^2 + \tau_j^2}\overline{\mu} + \frac{\sigma^2}{\sigma^2 + \tau_j^2}(v_j^i - \eta_j)\right]^2 \forall j \neq j^i.\end{aligned}$$

Maximization step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}, \sigma^2)$ and $(V_j^i|\boldsymbol{M}^i = \boldsymbol{\mu}^i) \sim \mathrm{N}(\mu_j^i + \eta_j, \tau_j^2)$, the maximum likelihood estimates for the model parameters $\theta = \{\overline{\mu}, \sigma^2\} \cup \{\eta_j, \tau_j^2\}_{j=1}^{n_i}$ are:

$$\begin{aligned}\hat{\overline{\mu}}_{k+1} &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\mu_j^i, \\ \hat{\sigma}_{k+1}^2 &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left((\mu_j^i)^2 - \hat{\overline{\mu}}_{k+1}^2\right), \end{aligned} \tag{D.3}$$

$$\begin{aligned}\hat{\eta}_{j,k+1} &= \frac{1}{m}\sum_{i=1}^{m} v_j^i - \hat{\overline{\mu}}_{k+1}, \\ \hat{\tau}_{j,k+1}^2 &= \frac{1}{m}\sum_{i=1}^{m}\left((v_j^i)^2 - 2v_j^i\mu_j^i + (\mu_j^i)^2 - \hat{\eta}_{j,k+1}^2\right).\end{aligned} \tag{D.4}$$

Initial values: Let us denote by $I_j = \{i \in \{1, \ldots m\} | j(\boldsymbol{v}^i) = j\}$ the indices of those decision settings in which alternative type j was selected. Suitable initial values for σ^2, η_j and τ_j^2 are obtained by applying (D.3) and (D.4) for the selected alternatives alone, and for $\overline{\mu}$ by noting that $\hat{\overline{\mu}}_{k+1} = \frac{1}{m}\sum_{i=1}^{m} v_j^i - \hat{\eta}_{j,k+1}$ in (D.4):

$$\hat{\overline{\mu}}_1 = \frac{1}{N}\sum_{i=1}^{m}\left(\sum_{j=1}^{n_i} v_j^i - \hat{\eta}_{j,1}\right),$$

$$\hat{\sigma}_1^2 = \frac{1}{m}\sum_{i=1}^{m}\left((\mu_{j^i}^i)^2 - \left(\frac{1}{m}\sum_{i=1}^{m}\mu_{j^i}^i\right)^2\right),$$

$$\hat{\eta}_{j,1} = \frac{1}{|I_j|}\sum_{i\in I_j}\left(v_{j^i}^i - \mu_{j^i}^i\right),$$

$$\hat{\tau}_{j,1}^2 = \frac{1}{|I_j|}\sum_{i\in I_j}\left(v_{j^i}^i - \mu_{j^i}^i\right)^2 - \hat{\eta}_{j,1}^2.$$

Example 3: Let us denote by $I = \{i \in \{1, \ldots, m\} | n_i = 1\}$ the set of decision settings with a single decision alternative. Then, the maximum likelihood estimates for η and τ^2 are

$$\hat{\eta} = \frac{1}{|I|}\sum_{i\in I}\left(v_1^i - \mu_1^i\right), \qquad \hat{\tau}^2 = \frac{1}{|I|}\sum_{i\in I}\left(v_1^i - \mu_1^i - \hat{\eta}\right)^2.$$

With fixed $\eta = \hat{\eta}$ and $\tau^2 = \hat{\tau}^2$, the log-likelihood function of parameters $\theta = \{\overline{\mu}_i\}_{i=1}^m \cup \{\sigma_i^2\}_{i=1}^m$ is

$$\begin{aligned}\mathcal{L}(\theta|\mu, M, v) &= -N\ln(2\pi\tau) - \sum_{i=1}^{m} n_i\ln\sigma_i \\ &- \frac{1}{2}\sum_{i=1}^{m}\left[\left(\frac{v_{j^i}^i - \mu_{j^i}^i - \eta}{\tau}\right)^2 + \left(\frac{\mu_{j^i}^i - \overline{\mu}_i}{\sigma_i}\right)^2\right. \\ &+ \left.\sum_{j\neq j^i}\left[\left(\frac{v_j^i - M_j^i - \eta}{\tau}\right)^2 + \left(\frac{M_j^i - \overline{\mu}_i}{\sigma_i}\right)^2\right]\right].\end{aligned}$$

Expectation step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}_i, \sigma_i^2)$ and $(V_j^i|M_j^i = \mu_j^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$, it follows from the properties of bivariate normal distribution that:

$$\widehat{\mu_j^i} := \mathbb{E}[M_j^i|v_j^i; \theta] = \frac{\tau^2}{\sigma_i^2 + \tau^2}\overline{\mu}_i + \frac{\sigma_i^2}{\sigma^2 + \tau^2}(v_j^i - \eta) \quad \forall j \neq j^i,$$

$$\begin{aligned}\widehat{(\mu_j^i)^2} &:= \mathbb{E}[(M_j^i)^2|v_j^i; \theta] = \frac{\sigma_i^2\tau^2}{\sigma_i^2 + \tau^2} \\ &+ \left[\frac{\tau^2}{\sigma_i^2 + \tau^2}\overline{\mu}_i + \frac{\sigma_i^2}{\sigma_i^2 + \tau^2}(v_j^i - \eta)\right]^2 \forall j \neq j^i.\end{aligned} \tag{D.5}$$

Maximization step: Because $M_j^i \sim \mathrm{N}(\overline{\mu}_i, \sigma_i^2)$ and $(V_j^i|M_j^i = \mu_j^i) \sim \mathrm{N}(\mu_j^i + \eta, \tau^2)$, the Maximum Likelihood estimates for parameters $\theta = \{\overline{\mu}_i\}_{i=1}^m \cup \{\sigma_i^2\}_{i=1}^m$ are

$$\hat{\overline{\mu}}_{i,k+1} = \frac{1}{n_i}\sum_{j=1}^{n_i}\mu_j^i, \quad \hat{\sigma}_{i,k+1}^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}\left((\mu_j^i)^2 - 2\mu_j^i\hat{\overline{\mu}}_{i,k+1} + \hat{\overline{\mu}}_{i,k+1}^2\right). \tag{D.6}$$

Initial values: Suitable initial values for $\overline{\mu}_i$ are obtained by combining (D.5) and (D.6), and for σ_i^2 by applying (D.6) for the selected alternatives alone:

$$\hat{\overline{\mu}}_{i,1} = \frac{1}{n_i}\sum_{j=1}^{n_i} v_j^i - \hat{\eta}, \qquad \hat{\sigma}_{i,1}^2 = (\mu_{j^i}^i)^2 - \left(\frac{1}{m}\sum_{i=1}^{m}\mu_{j^i}^i\right)^2.$$

Appendix E. Estimation of model parameters in the Caltrans case

The values of parameters η and τ^2 are estimated from those 131 decision settings in which there was a single alternative. Denoting the indices of such settings by $I = \{i \in \{1, \ldots, 5610\} | n_i = 1\}$, the Maximum Likelihood estimates for η and τ^2 are

$$\hat{\eta} = \frac{1}{|I|}\sum_{i\in I}(\ln v_1^i - \ln\mu_1^i), \qquad \hat{\tau}^2 = \frac{1}{|I|}\sum_{i\in I}(\ln v_1^i - \ln\mu_1^i - \hat{\eta})^2. \tag{E.1}$$

Fixing $\eta = \hat{\eta}$ and $\tau^2 = \hat{\tau}^2$, parameters $\overline{\mu}_i$ and σ_i^2 are estimated using the EM-algorithm. The log-likelihood function for $\theta = \{\overline{\mu}_i, \sigma_i^2\}_{i=1}^m$ is

$$\mathcal{L}(\theta|\mu, M, v) = -N\ln(2\pi\tau) - \sum_{i=1}^{m} n_i\ln\sigma_i -$$

$$- \frac{1}{2}\sum_{i=1}^{m}\left[\left(\frac{\ln v^i_{j^i} - \ln \mu^i_{j^i} - \eta}{\tau}\right)^2 + \left(\frac{\ln \mu^i_{j^i} - \overline{\mu}_i}{\sigma_i}\right)^2 + \sum_{j\neq j^i}\left[\left(\frac{\ln v^i_j - \ln M^i_j - \eta}{\tau}\right)^2 + \left(\frac{\ln M^i_j - \overline{\mu}_i}{\sigma_i}\right)^2\right]\right],$$

where $N = \sum_{i=1}^{m} n_i$.

Expectation step: Because $\ln M^i_j \sim \mathrm{N}(\overline{\mu}_i, \sigma^2)$ and $(\ln V^i_j | \ln M^i_j = \ln \mu^i_j) \sim \mathrm{N}(\ln \mu^i_j + \eta, \tau^2)$, it follows from the properties of bivariate normal distribution that

$$\widehat{\ln \mu^i_j} = \mathbb{E}[\ln M^i_j | \ln V^i_j = \ln v^i_j; \theta] = \frac{\tau^2}{\sigma_i^2 + \tau^2}\overline{\mu}_i + \frac{\sigma_i^2}{\sigma_i^2 + \tau^2}(\ln v^i_j - \eta),$$

$$\widehat{\ln^2 \mu^i_j} = \mathbb{E}[\ln^2 M^i_j | \ln V^i_j = \ln v^i_j; \theta] = \frac{\sigma_i^2 \tau^2}{\sigma_i^2 + \tau^2} + \left(\widehat{\ln \mu^i_j}\right)^2.$$

Maximization step: Because $\ln M^i_j \sim \mathrm{N}(\overline{\mu}_i, \sigma_i^2)$, the Maximum Likelihood estimates for $\overline{\mu}_i$ and σ_i^2 are

$$\hat{\overline{\mu}}_{i,k+1} = \frac{1}{n_i}\sum_{j=1}^{n_i} \ln \mu^i_j,$$

$$\hat{\sigma}^2_{i,k+1} = \frac{1}{n_i}\sum_{j=1}^{n_i}\left(\ln^2 \mu^i_j - 2\ln \mu^i_j \hat{\overline{\mu}}_{i,k+1} + \hat{\overline{\mu}}^2_{i,k+1}\right). \tag{E.2}$$

Initial values: Suitable initial values for $\overline{\mu}_i$ are obtained by combining (E.1) and (E.2), and for σ_i^2 by applying (E.2) for the selected alternatives alone:

$$\hat{\overline{\mu}}_{i,1} = \frac{1}{n_i}\sum_{j=1}^{n_i} \ln v^i_j - \hat{\eta}, \qquad \hat{\sigma}^2_{i,1} = \left(\ln \mu^i_{j^i} - \hat{\overline{\mu}}_{i,1}\right)^2.$$

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.ejor.2021.04.018

References

Bajari, P., Houghton, S., & Tadelis, S. (2011). Bidding for incomplete contracts: An empirical analysis of adaptation costs. *The American Economic Review, 104*, 1288–1319.

Baker, E., Bosetti, V., & Salo, A. (2020). Robust portfolio decision analysis: An application to the energy research and development portfolio problem. *European Journal of Operational Research, 284*, 1107–1120.

Begg, S. H., & Bratvold, R. B. (2008). Systematic prediction errors in oil and gas project and portfolio selection. *Proceedings of the SPE annual technical conference and exhibition*. Society of Petroleum Engineers.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research, 33*, 1–27.

Brous, P., Hiedemann, B., & Schultz, T. (2009). Success rates, optimism bias, and the predictive power of alternative performance measures: A post-audit of one company's real investment decisions. *Journal of Financial Education, 35*, 29–41.

Capen, E. C., Clapp, R. V., & Campbell, W. M. (1971). Competitive bidding in high-risk situations. *Journal of Petroleum Technology, 23*, 641–653.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application (vol. 1)*. Cambridge University Press.

Davydenko, A., & Fildes, R. (2013). Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. *International Journal of Forecasting, 29*, 510–522.

De Baets, S., & Harvey, N. (2020). Using judgment to select and adjust forecasts from statistical models. *European Journal of Operational Research, 284*, 882–895.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological), 39*, 1–38.

Durbach, I. N., & Montibeller, G. (2019). Behavioural analytics: Exploring judgments and choices in large data sets. *Journal of the Operational Research Society, 70*, 255–268.

Ferretti, V., & Geneletti, D. (2020). Does the spatial representation affect criteria weights in environmental decision-making? insights from a behavioral experiment. *Land Use Policy, 97*, 104613.

Flyvbjerg, B. (2009). Survival of the unfittest: Why the worst infrastructure gets built – and what we can do about it. *Oxford Review of Economic Policy, 25*, 344–367.

Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association, 68*, 279–295.

Franco, L. A., & Hämäläinen, R. P. (2016). Behavioural operational research: Returning to the roots of the OR profession. *European Journal of Operational Research, 249*, 791–795.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC Press.

Gupta, D., Snir, E. M., & Chen, Y. (2015). Contractors' and agency decisions and policy implications in a+b bidding. *Production and Operations Management, 24*, 159–177.

Harrison, J. R., & March, J. G. (1984). Decision making and postdecision surprises. *Administrative Science Quarterly, 29*, 26–42.

Jø rgensen, M. (2013). The influence of selection bias on effort overruns in software development projects. *Information and Software Technology, 55*, 1640–1650.

Jørgensen, M. (2007). Forecasting of software development work effort: Evidence on expert judgement and formal models. *International Journal of Forecasting, 23*, 449–462.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: a cognitive perspective on risk taking. *Management Science, 39*, 17–31.

Kellner, F., Lienland, B., & Utz, S. (2019). An a posteriori decision support methodology for solving the multi-criteria supplier selection problem. *European Journal of Operational Research, 272*, 505–522.

Kettunen, J., & Salo, A. (2017). Estimation of downside risks in project portfolio selection. *Production and Operations Management, 26*, 1839–1853.

Lahtinen, T. J., & Hämäläinen, R. P. (2016). Path dependence and biases in the even swaps decision analysis method. *European Journal of Operational Research, 249*, 890–898.

Lahtinen, T. J., Hämäläinen, R. P., & Jenytin, C. (2020). On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis. *European Journal of Operational Research, 282*, 201–210.

Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data (vol. 333)*. John Wiley & Sons..

Lovallo, D., & Kahneman, D. (2003). Delusions of success. *Harvard Business Review, 81*, 56–63.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance, 60*, 2661–2700.

Merrow, E. W., Phillips, K. E., & Myers, C. W. (1981). Understanding cost growth and performance shortfalls in pioneer process plants. (No. RAND/R-2569-DOE). Rand Corp., Santa Monica, CA.

Montibeller, G., & von Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Analysis, 35*, 1230–1251.

Morton, A., & Fasolo, B. (2009). Behavioural decision theory for multi-criteria decision analysis: a guided tour. *Journal of the Operational Research Society, 60*, 268–275.

Odeck, J. (2004). Cost overruns in road construction – what are their sizes and determinants? *Transport Policy, 11*, 43–53.

Oliva, R., & Watson, N. (2009). Managing functional biases in organizational forecasts: a case study of consensus forecasting in supply chain planning. *Production and Operations Management, 18*, 138–151.

Petropoulos, F., Fildes, R., & Goodwin, P. (2016). Do 'big losses' in judgmental adjustments to statistical forecasts affect experts' behaviour? *European Journal of Operational Research, 249*, 842–852.

Pruitt, S. W., & Gitman, L. J. (1987). Capital budgeting forecast biases: Evidence from the fortune 500. *Financial Management, 16*, 46–51.

Quirk, J., & Terasawa, K. (1986). Sample selection and cost underestimation bias in pioneer projects. *Land Economics, 62*, 192–200.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. CRC press.

Shleifer, A. (2000). *Inefficient markets: an introduction to behavioural finance*. Oxford: Oxford University Press.

Smith, J. E., & Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science, 52*, 311–322.

Subrahmanyam, A. (2008). Behavioural finance: a review and synthesis. *European Financial Management, 14*, 12–29.

Terasawa, K., Quirk, J., & Womer, K. (1989). Turbulence, cost escalation and capital intensity bias in defense contracting. *Cost analysis applications of economics and operations research 97–116*. Springer US.

Thaler, R. H. (1992). *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. New York: Free Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

Vilkkumaa, E., Liesiö, J., & Salo, A. (2014). Optimal strategies for selecting project portfolios using uncertain value estimates. *European Journal of Operational Research, 233*, 772–783.

Wang, C., & Guo, P. (2017). Behavioral models for first-price sealed-bid auctions with the one-shot decision theory. *European Journal of Operational Research, 261*, 994–1000.

Wu, C. F. J. (1983). On the convergence properties of the EM algorithm. *The Annals of Statistics, 11*, 95–103.