European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

# Decision-based model selection

Arnoud V. den Boer[a,*], Dirk D. Sierag[b]

[a] University of Amsterdam, Korteweg-de Vries Institute for Mathematics and Amsterdam Business School, The Netherlands
[b] Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands

ARTICLE INFO

*Article history:*
Received 3 September 2018
Accepted 13 August 2020
Available online xxx

*Keywords:*
Analytics
Model selection
Data-driven optimization
Decision making under uncertainty

ABSTRACT

A key step in data-driven decision making is the choice of a suitable mathematical model. Complex models that give an accurate description of reality may depend on many parameters that are difficult to estimate; in addition, the optimization problem corresponding to such models may be computationally intractable and only approximately solvable. Simple models with only a few unknown parameters may be misspecified, but also easier to estimate and optimize. With such different models and some initial data at hand, a decision maker would want to know which model produces the best decisions. In this paper we propose a decision-based model-selection method that addresses this question.



## 1. Introduction

### 1.1. Motivation

Model selection is the art of choosing from different mathematical models the one that provides the best description of a certain real-world phenomenon. Many different model selection criteria have been proposed, typically based on statistical or information-theoretic notions related to 'goodness-of-fit' or 'explanatory power', while also (albeit sometimes implicitly) taking into account the number of parameters present in a model.

Mathematical models play a fundamental role in data-driven optimization problems studied in operations research and management science. In these problems one is not primarily interested in obtaining a good *description* of some aspect of reality, but rather in identifying a good *decision* that maximizes a certain objective function. One would therefore expect that the main criterion based upon which one selects a model in a data-driven optimization problem is its ability to produce good decisions.

Perhaps surprisingly, this is not the case. Models are often selected using 'classical' criteria related to obtaining estimates with small statistical distance (such as mean squared error or Kullback-Leibler divergence). But, as illustrated in Fig. 1, small statistical distance need not at all imply that the selected model leads to good decisions (and the Appendix of this paper contains an example showing that the loss of using the 'wrong' model selection criterion can in fact be unbounded). A striking practical example of this phenomenon is given by Feldman, Zhang, Liu, and Zhang (2019), who compare a sophisticated machine learning model to a simple multinomial logit (MNL) choice model, in a product display optimization problem of a large online market place. They conclude:

> Our experiments show that despite the lower prediction power of our MNL-based approach, it generates significantly higher revenue per visit compared to the current machine learning algorithm with the same set of features.

In addition, Besbes and Zeevi (2015) and Cooper, Homem-de Mello, and Kleywegt (2015) have shown that misspecified models may sometimes lead to good or even better decisions than a 'correct' model. Thus, in data-driven optimization problems, the value of a model should solely be judged by the quality of the decisions it produces, and not by, e.g., 'goodness-of-fit'. This has also been pointed out by Besbes, Phillips, and Zeevi (2010), who write:

> '[...] there has long been a recognition within the decision analysis literature that the value of quantitative modeling should be judged primarily by the quality of the decisions they support (see, for example, Nickerson & Boyd, 1980). However, there has been a lack of methodologies for evaluating the adequacy of a particular model from this vantage point.'

In decision problems, model selection is typically between complex, 'realistic' models and 'simple' or simplified models. A complex model, that takes into account many factors that are thought to be relevant and important for the problem at hand, typically depends on many unknown parameters that may be difficult to estimate accurately (especially if only limited data is available). In addition, determining the corresponding optimal decision may be computationally intractable, such that heuristics or simulations have to be used to find an (approximately) optimal solution.

* Corresponding author.
*E-mail address:* boer@uva.nl (A.V. den Boer).

https://doi.org/10.1016/j.ejor.2020.08.025
0377-2217/

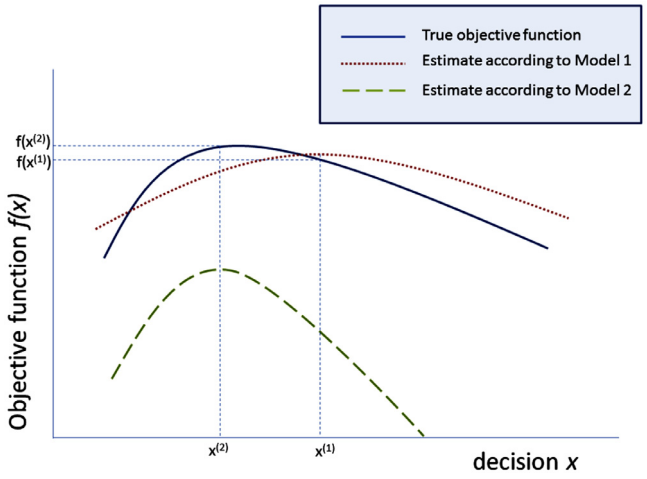


**Fig. 1.** Although the objective function estimated by Model 1 is closer to the truth than that of Model 2 (measured, e.g., by their $L^2$ distance), the optimal decision $x^{(2)}$ corresponding to Model 2 yields a higher objective $f(x^{(2)})$ than the optimal decision $x^{(1)}$ corresponding to Model 1.

A simple model that neglects important factors may be misspecified, but it may also involve fewer unknowns that need to be estimated, and the associated optimization problem may be exactly solvable.

With two such models at hand, an important question is whether the *modeling* error of the simple model outweighs the larger *estimation* and *optimization* errors associated with the complex model. A large variety of model-selection methods exists (discussed in more detail in the next section) based on statistical or information-theoretic criteria. Although these criteria may perform well when one wants to derive qualitative insights or make predictions, they are generally not tailored to their use in generic *optimization problems*, and thus may select a model based on the 'wrong' criterion as illustrated in Fig. 1. This motivates the current study, in which we integrate model selection with data-driven decision problems, by proposing a concrete and generic decision-based model-selection method.

### 1.2. Literature

*Model selection.* The rich field of model selection has produced a wide variety of tools and techniques to select a model, such as Akaike Information Criterion (AIC; Akaike, 1973), Bayesian information criterion (BIC; Schwarz, 1978), deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), Mallows' $C_p$ (Mallows, 1973), the Minimum Description Length principle (Rissanen, 1978), Bayesian model selection and model averaging based on Bayes factors (Jeffreys, 1935; 1961), cross-validation (Geisser, 1975; Stone, 1974), and many more. For reviews and in-depth discussions of these methods we refer to the books and survey papers by Arlot and Celisse (2010), Burnham and Anderson (2002), Claeskens and Hjort (2008), Grünwald (2007), Kass and Raftery (1995), Lahiri (2001), Wasserman (2000), and Zucchini (2000).

These model selection methods are based on statistical or information-theoretic criteria, and generally are designed with the aim of identifying models with small statistical distance to the underlying ground truth or a good fit on future (test) data coming from the same source. If the goal of the decision maker is to use models to derive qualitative insights or make predictions, these criteria may perform quite well. However, these criteria are not necessarily aligned with the goal of selecting a model that produces good decisions. This observation, that a model selection method should be aligned with the purpose for which the model is used, is also made by Claeskens and Hjort (2003):

> 'The idea of finding a single satisfactory statistical model for one's data, perhaps aided by the model information criteria discussed previously, is a central one in statistics, and carries with it considerable intellectual and conceptual appeal. The chosen model is fitted to data and is seen as the statistician's best approximation to the real data generating mechanism used by nature, and secures a coherent view of statistical analysis of a dataset. In this article we carefully extricate ourselves from this classic point of view; that a single model should be used to explain all aspects of data or to predict all types of future data points seems to us a little too constrained. Our view is that such a "best model" should depend on the parameter under focus.'

Claeskens and Hjort (2003) proceed by proposing a method, the Focused Information Criterion (FIC), aimed at selecting a model that gives good precision for estimating a certain parameter of interest. The idea of FIC is to estimate the mean squared error of the parameter of interest for each available model, and then select a model for which this estimate is minimal. The method proposed in the present paper is different from FIC, but it is inspired by the same philosophy that model selectors should be aligned with the purpose for which the models are used (in our case: producing good decisions).

*Statistical learning theory.* Statistical learning theory (Bousquet, Boucheron, & Lugosi, 2004; Hastie, Tibshirani, & Friedman, 2009; Vapnik, 1998; 2000) addresses questions that are closely related to model selection. The main goal in this field is to construct a prediction function $\hat{f}: \mathcal{X} \to \mathcal{Y}$ that provides a good description of the relation between an input random variable $X$ with support $\mathcal{X}$ and an output random variable $Y$ with support $\mathcal{Y}$. The joint distribution of $(X, Y)$ is unknown, but data consisting of i.i.d. realizations $(x_i, y_i)_{1 \le i \le n}$ of $(X, Y)$ is available. The quality of a predictor $\hat{f}$ is measured by the risk $R(\hat{f}) := \mathbb{E}[L(Y, \hat{f}(X))]$, where the so-called loss function $L: \mathcal{Y}^2 \to [0, \infty)$ quantifies the error between $Y$ and the predicted $\hat{f}(X)$. As described in Bousquet et al. (2004) and Guyon, Saffari, Dror, and Cawley (2010), the main methods to determine a good predictor (empirical risk minimization, structural risk minimization, regularization methods) are based on the idea of minimizing the empirical risk $\sum_{i=1}^{n} L(y_i, \hat{f}(x_i))$ over some class $\mathcal{G} \subset \mathcal{Y}^{\mathcal{X}}$ of predictors, possibly augmented with a term that penalizes the 'model complexity' of $\hat{f}$. Selecting $\mathcal{G}$ can be seen as a model selection problem.

The framework is quite general - covering, for example, classification, regression, and density estimation problems - but is in several regards different from the setting considered in this paper. First, we do not make the assumption that $x_1, \ldots, x_n$ are i.i.d. realizations from a random $X$. Second, as outlined in Section 2.1, we are solely interested in estimating a *maximizer* of the function $f(x) := \mathbb{E}[\mathfrak{r}(X, Y) \mid X = x]$ for some known $\mathfrak{r}: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, instead of the 'full' relation between $X$ and $Y$. For many problems, e.g. the assortment optimization problem considered in Section 4.1, it is unclear if (at all) it is possible to put this into the framework described above.

*Bayesian model averaging.* Bayesian model selection techniques share the same drawbacks as frequentists' approaches, in that they decouple model selection from a particular optimization problem at hand. Bayesian model averaging (Kass & Raftery, 1995; Wasserman, 2000), however, is an approach that can connect optimization problems to the availability of different models. The main idea of Bayesian model averaging is not to select a single model from available alternatives, but to maintain a probability

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

distribution that each of the given models is correct, and use this probability distribution in all further derivations.

In data-driven optimization problems, such an approach could look as follows. Let $M_0, \ldots, M_K$ be $K+1$ models that may have generated a given data set $D$, let each model $M_k$ have an associated parameter $\theta_k$ living in a space $\Theta_k \subset \Theta$, let $\mathcal{X}$ be a space of feasible decisions, and let $r : \mathcal{X} \times \Theta \to \mathbb{R}$ be a reward function. Let $p(\theta_k|M_k)$ be a prior on the parameter $\theta_k \in \Theta_k$, and let $p(M_k)$ be a prior on the probability that model $M_k$ is correct, for $k = 0, \ldots, K$. A fully Bayesian approach to maximize the reward, in the spirit of Bayesian model averaging and thus without first selecting a model, consists of maximizing the function

$$x \mapsto \sum_{k=0}^{K+1} \int_{\theta_k \in \Theta_k} r(x, \theta_k) p(\theta_k \mid D, M_k) p(M_k \mid D) \mathrm{d}\theta_k, \quad (x \in \mathcal{X}), \quad (1)$$

where $p(\theta_k \mid D, M_k) = p(D \mid \theta_k, M_k) p(\theta_k \mid M_k)/p(D \mid M_k)$ is the posterior of $\theta_k$, $p(D|\theta_k, M_k)$ is the likelihood of the data given parameter value $\theta_k$ and model $M_k$, $p(D \mid M_k) = \int_{\theta_k \in \Theta_k} p(D \mid \theta_k, M_k) p(\theta_k \mid M_k) \mathrm{d}\theta_k$ is the evidence for model $M_k$, and $p(M_k \mid D) = p(D \mid M_k) p(M_k) / \sum_{l=0}^{K} p(D \mid M_l) p(M_l)$ is the posterior probability that model $M_k$ is correct, given data $D$.

Apart from the computational difficulties that solving (1) could involve (which could introduce further optimization errors), a main difference between this and our approach is that we do not (implicitly) assume that each of the available models is 'correct' with some (positive) probability. Even if it is known beforehand that a certain model $M_k$ is incorrect and could never have generated the data, i.e. $p(M_k) = 0$, it still could produce better decisions than a correctly specified model. This aspect is not captured in this approach.

Somewhat related is the literature on inconsistent (Bayesian) inference with misspecified models (see Grünwald and van Ommen, 2014; Watson and Holmes, 2016, and the references therein), which blends Bayesian methods with statistical learning theory. Similar to the statistical learning theory literature discussed above, these papers differ, among other things, from our framework by assuming i.i.d. decisions and a different structural form of the loss function.

*Bridging model selection and data-driven optimization.* Several recent studies in the operations research and management science literature consider aspects of model selection in conjunction with data-driven optimization problems. For example, Besbes et al. (2010) design and analyze a hypothesis test to discriminate between models based on the quality of decisions they produce; Chu, Shanthikumar, and Shen (2008), Lim, Shanthikumar, and Shen (2006), Liyanage and Shanthikumar (2005), and Ramamurthy, Shanthikumar, and Shen (2012) argue for integrating estimation, optimization, and model uncertainty in data-driven optimization problems; and Besbes and Zeevi (2015), Cachon and Kök (2007), Cooper and Li (2012), Cooper, Homem-de Mello, and Kleywegt (2006), Cooper et al. (2015), and Lee, Homem-de Mello, and Kleywegt (2012) study the quality of decisions under misspecified models in pricing, revenue management, and inventory optimization problems.

The notion that model selection methods should be aligned with their purpose of producing good decisions is implicitly present in some recent studies. Bastani and Bayati (2016), for example, consider a linear multi-armed bandit problem with high-dimensional covariates, and adaptively tune the regularization parameter of the LASSO estimator (Tibshirani, 1996) in order to achieve optimal asymptotic reward. Since this regularization parameter is a measure of model complexity, their method can be seen as an example of adapting a model selection method to the purpose of generating good decisions. A similar idea appears in Vahn, El Karoui, and Lim (2014), who enhance a data-driven portfolio optimization problem by a regularization parameter that bounds the sample variance of the estimated objective function. The regularization parameter is optimized by a variant of $k$-fold cross-validation, where the validation step is based on a performance metric relevant to investment problems. Although this paper is not directly about model selection, it can again be seen as an example where model selection techniques are tuned in order to maximize the objective function of a data-driven optimization problem.

Kao and Van Roy (2014) (cf. Kao & Van Roy, 2013) consider a quadratic optimization problem, the solution of which depends on an unknown covariance matrix $\Sigma$ of a Gaussian random variable. The authors discuss various regularized maximum likelihood estimators with regularization parameter tuned via cross-validation. In addition, they propose to estimate $\Sigma$ by maximizing the in-sample performance of the objective function, subject to a lower-bound on the posterior probability of $\Sigma$ to mitigate overfitting. Thus, the estimator of the unknown parameter is adapted to take the decision objective into account. A similar idea is considered by Kao, Van Roy, and Yan (2009), who estimate an unknown parameter of a quadratic function by a convex combination of ordinary least squares and empirical loss minimization, and who choose this convex combination while taking into account the goal of maximizing the objective function.

### 1.3. Contributions

This paper proposes a model-selection method that evaluates models based on the quality of the decisions they produce. The key idea of the approach, named DBMS after decision-based model selection, is to use a resampling procedure to estimate which of the decisions suggested by different models gives the highest reward. The method is applicable to a wide class of data-driven decision problems, is not computationally intensive, and does not depend on some hyper-parameter that is difficult to tune. Conceptually, it connects the fields of model selection and data-driven optimization. Our numerical results are encouraging, while also suggesting that there still is room for further improvement.

The main practical insight for managers or practitioners who work with models and data is that one does not have to confine oneself to using either simple and misspecified or complex and intractable models: one can (and in fact: should) use both, together with a method such as DBMS that predicts which model produces the best decision given the data set at hand.

Our numerical results also reveal that it is in general quite hard to conclude which model selection method is 'the best'. In the assortment optimization problem considered in Section 4.1, DBMS is much better than AIC when model $M^{(1)}$ is clearly misspecified, while AIC is better than DBMS when model $M^{(1)}$ is (near)correct. In the newsvendor optimization problem considered in Section 4.2, a similar conclusion holds: DBMS is better than cross-validation when model $M^{(1)}$ is clearly misspecified, while cross-validation is better than DBMS when $M^{(1)}$ is correctly specified. From a practical point of view, conducting numerical simulations or real-life experiments to evaluate the performance of different models (as in Feldman et al., 2019) might be insightful. From a theoretical point of view, it would be useful to derive worst-case performance bounds for different model selection methods, accompanied by lower bounds on the best achievable performance of any model selection method. Our analysis in Section 3.1 suggests that a general analysis of this kind might be technically challenging. However, in particular problem instances (such as newsvendor optimization), deriving informative upper and lower bounds on the performance of model selection methods might be feasible. This is left as an interesting direction for future research.

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

### 1.4. Organization of the paper

The rest of this paper is organized as follows. Section 2 describes the formal decision-making framework that we consider, and contains our decision-based model-selection criterion DBMS. In Section 3 we explain the intuition behind DBMS, discuss a few alternatives, prove that DBMS is reward-consistent, and comment on several other aspects of DBMS. Section 4 contains two numerical illustrations, on an assortment optimization problem and on the newsvendor problem, and Section 5 ends the paper with a few concluding remarks. The supplementary material in the appendix shows, by means of an example, that unbounded losses may be incurred when model selection is not based on optimizing the objective function.

## 2. Decision-based model selection

### 2.1. Mathematical framework of decision-making

Consider a decision maker who tries to determine a *decision* or *action* $x$ in an *action space* $\mathcal{X}$ that maximizes her expected reward $\mathbb{E}[\mathfrak{r}(x, Y(x))]$, where $\{Y(x) \mid x \in \mathcal{X}\}$ is a collection of (possibly multivariate) random variables with common support $\mathcal{Y}$, and $\mathfrak{r} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is a known function called the *reward function*. The distributions of $Y(x)$ $(x \in \mathcal{X})$ are unknown to the decision maker, but a data set $d_0 = (x_1, y_1, \ldots, x_n, y_n)$, $n \in \mathbb{N}$, consisting of previously used actions $x_i \in \mathcal{X}$ and realizations $y_i$ of $Y(x_i)$ $(i = 1, \ldots, n)$ is available. To determine her data-driven decision, the decision maker uses a *model*, an *estimator*, and an *optimization algorithm*.

A model is a set of the form

$$M = \{F_{x,\theta} \in \mathbf{F} \mid x \in \mathcal{X}, \theta \in \Theta\},$$

where $\mathbf{F}$ is the set of cumulative distribution functions (cdfs) on $\mathcal{Y}$, and $\Theta$ is a non-empty and possibly infinite-dimensional set. A model is called correctly specified if there is a unique $\theta^* \in \Theta$ such that, for all $x \in \mathcal{X}$, $F_{x,\theta^*}$ is the cdf of $Y(x)$; $\theta^*$ is then called the true parameter. An estimator is a function $\tau : (\mathcal{X} \times \mathcal{Y})^n \to \Theta$ that maps data to parameter values. An optimization algorithm is a function $\chi : \Theta \to \mathcal{X}$ that maps parameter values to decisions. In data-driven decision problems, $\chi(\theta)$ typically maximizes $\int_{\mathcal{Y}} \mathfrak{r}(x, y)\mathrm{d}F_{x,\theta}(y)$ with respect to $x \in \mathcal{X}$, for all $\theta \in \Theta$; in this case $\chi$ is called *exact*. In many optimization problems, however, maximizing $\int_{\mathcal{Y}} \mathfrak{r}(x, y)\mathrm{d}F_{x,\theta}(y)$ is intractable, and $\chi$ is an heuristic or approximate optimal solution. If a single model $M$ with corresponding estimator $\tau$ and optimization algorithm $\chi$ is at hand, then the decision maker uses action $\chi(\tau(d_0))$.

### 2.2. Decision-based model-selection criterion

We consider the case where *multiple* models $M^{(0)}, M^{(1)}, \ldots, M^{(K)}$ are available $(K \in \mathbb{N})$, each of the form

$$M^{(k)} = \{F^{(k)}_{x,\theta} \in \mathbf{F} \mid x \in \mathcal{X}, \theta \in \Theta^{(k)}\}, \quad k = 0, 1, \ldots, K,$$

and each with corresponding estimator $\tau^{(k)}$ and optimization algorithm $\chi^{(k)}$. Model $M^{(0)}$ is called the 'true model' and is correctly specified with true (but unknown) parameter $\theta^*$ (For a discussion about this assumption, see Section 3.4). The other models are considered simplifications, and may be misspecified. The decision maker knows that model $M^{(0)}$ is correctly specified.

Let $x^{(k)} := \chi^{(k)}(\tau^{(k)}(d_0))$ denote the decision suggested by model $k$ $(k = 0, 1, \ldots, K)$. The decision maker needs to determine the model $k$ for which $x^{(k)}$ gives the highest expected reward; i.e. she needs to estimate

$$\arg \max_{k \in \{0,1,\ldots,K\}} r(x^{(k)}, \theta^*), \tag{2}$$

where we write

$$r(x, \theta) := \int_{\mathcal{Y}} \mathfrak{r}(x, y)\mathrm{d}F^{(0)}_{x,\theta}(y)$$

for the expected reward under model $M^{(0)}$ as function of $x$ and $\theta$.

Observe that simply replacing $\theta^*$ by $\tau^{(0)}(d_0)$ in (2) is not informative: if $\chi^{(0)}$ is exact, then we have $r(x^{(0)}, \tau^{(0)}(d_0)) \geq r(x^{(k)}, \tau^{(0)}(d_0))$ *by definition*, for all $k = 1, \ldots, K$.

Our idea is to estimate (2) by a resampling procedure, as follows: construct a new data set $d_r = (x_1, y^r_1, \ldots, x_n, y^r_n)$ - the sub-/superscript 'r' refers to 'resampled' - with the same covariates $x_i$ as in $d_0$, but with the observations $y_i$ replaced by random samples $y^r_i$ drawn according to their estimated cdf $F^{(0)}_{x_i,\tau^{(0)}(d_0)}$ $(i = 1, \ldots, n)$. We subsequently estimate (2) by replacing the true parameter $\theta^*$ by its estimate *based on the resampled data*:

$$\arg \max_{k \in \{0,1,\ldots,K\}} r(x^{(k)}, \tau^{(0)}(d_r)). \tag{DBMS}$$

In case of a tie, we select the maximizer with the smallest $k$.

## 3. Discussion and analysis

### 3.1. Intuition behind DBMS

In this section we provide an intuition behind DBMS. To this end, we introduce some notation: we write $Y(x, \theta)$ for the random variable with cdf $F^{(0)}_{x,\theta}$, and $D(\theta) = (x_1, Y(x_1, \theta), \ldots, x_n, Y(x_n, \theta))$ for the random data vector as function of $\theta$. We write $\theta_k = \tau^{(k)}(D(\theta^*))$ for the parameter estimate under model $M^{(k)}$ (viewed as a random variable), and $\theta_r = \tau^{(0)}(D(\theta_0))$ for the parameter estimate based on the resampled data set. Note that we can regard the initial data set $d_0$ as a realization of $D(\theta^*)$, and the resampled data set $d_r$ as a realization of $D(\theta_0)$.

We first explain in Section 3.1.1 why misspecified models may yield better decisions than the correct model. Next, in Section 3.1.2, we study several structural properties of DBMS by means of an example that involves two models. In this example, we show how the probability that the misspecified model outperforms the correct model, the corresponding expected performance gain, and the probability that DBMS selects the misspecified model are related to the variance of the estimator under $M^{(0)}$ and to the expected gain or loss under the misspecified model. We also obtain an explicit expression for the performance of DBMS. In Section 3.1.3 we discuss the difficulties of extending these insights to more general decision problems.

#### 3.1.1. Better decisions by a misspecified model.

For ease of exposition we assume that there are only two models under consideration: a correctly specified model $M^{(0)}$ and a possibly misspecified model $M^{(1)}$. In addition suppose that $\mathcal{X}$, $\Theta^{(0)}$ and $\Theta^{(1)}$ are metric spaces, the function $x \mapsto r(x, \theta^*)$ is globally Lipschitz continuous on $\mathcal{X}$ with unique maximizer $x^* \in \mathcal{X}$, the algorithm $\chi^{(0)}$ is exact, the function $\chi^{(1)} : \Theta^{(1)} \to \mathcal{X}$ is globally Lipschitz continuous, and there are constants $c_0 > 0$ and $\omega > 0$ such that estimation error and performance loss of model $M^{(0)}$ are related in the following way:

$$r(\chi^{(0)}(\theta^*), \theta^*) - r(\chi^{(0)}(\theta), \theta^*) \geq c_0 ||\theta^* - \theta||^{\omega} \text{ for all } \theta \in \Theta^{(0)}.$$

We define the *regret* under model $M^{(k)}$, denoted by $\mathrm{Regret}^{(k)}$, as the expected performance loss caused by using decision $x^{(k)}$ instead of the optimal decision $x^*$. It follows that the regret under model $M^{(0)}$ satisfies

$$\begin{aligned} \mathrm{Regret}^{(0)} &= \mathbb{E}[r(x^*, \theta^*) - r(x^{(0)}, \theta^*)] \\ &\geq c_0\, \mathbb{E}[||\theta^* - \theta_0||^{\omega}], \end{aligned} \tag{3}$$

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

where the expectation is taken with respect to the distribution of the initial data $D(\theta^*)$. Define the misspecification loss of model $M^{(1)}$ by

$$\Delta^{(1)} := \min_{\theta\in\Theta^{(1)}} \left\{ r(x^*,\theta^*) - r(\chi^{(1)}(\theta),\theta^*) \right\},$$

and, for ease of exposition, suppose that there is a unique $\tilde{\theta} \in \Theta^{(1)}$ where this minimum is achieved. The Lipschitz conditions on $r$ and $\chi^{(1)}$ imply that there is constant $c_1 > 0$ such that

$$\begin{aligned}\text{Regret}^{(1)} &= \mathbb{E}[r(x^*,\theta^*) - r(x^{(1)},\theta^*)] \\ &= \Delta^{(1)} + \mathbb{E}[r(\chi^{(1)}(\tilde{\theta}),\theta^*) - r(\chi^{(1)}(\theta_1),\theta^*)] \\ &\le \Delta^{(1)} + c_1\mathbb{E}[||\tilde{\theta} - \theta_1||]. \end{aligned} \tag{4}$$

By comparing (3) and (4) it becomes clear why a misspecified model may yield better decisions than a correctly specified model. In particular, if the estimation error of model $M^{(0)}$, measured by the expression on the righthandside of (3), is larger than the sum of the misspecification error and estimation error of model $M^{(1)}$, measured by the two respective terms in Eq. (4), then $x^{(1)}$ has lower regret than $x^{(0)}$ and thus model $M^{(1)}$ is preferable from a decision-making perspective, even if this model is misspecified. It is worth emphasizing that this discussion assumes that $\chi^{(0)}$ is exact. Optimization errors in the algorithm corresponding to model $M^{(0)}$ can be another source of why a misspecified model performs better than a well-specified model.

*3.1.2. Properties of DBMS in an example.*

We now explain key properties of DBMS by means of an example. Suppose that $\mathcal{X} = \mathbb{R}$, $\Theta^{(0)} = \mathbb{R}$, and $r(x,\theta) = -(x-\theta)^2$ for all $(x,\theta) \in \mathbb{R}^2$. Data of the form $(x_i, y_i)$, $i = 1,\dots,n$, $n \in \mathbb{N}$, is available, where $x_1,\dots,x_n \in \mathbb{R}$ are not all zero, $y_i = \theta^* x_i + \epsilon_i$ $(i = 1,\dots,n)$, $\theta^* \in \mathbb{R}$ is the true but unknown parameter, and $\epsilon_1,\dots,\epsilon_n$ are i.i.d. standard normally distributed random variables. The decision maker again considers two models. In model $M^{(0)}$, she (correctly) assumes that $y_i \sim N(\theta x_i, 1)$, for some $\theta \in \mathbb{R}$ and all $i = 1,\dots,n$; she estimates the unknown parameter by ordinary least squares, i.e. $\theta_0 = (\sum_{i=1}^n x_i^2)^{-1}\sum_{i=1}^n x_i y_i$, and uses the exact algorithm $\chi^{(0)}(\theta) = \theta$ for all $\theta \in \mathbb{R}$; that is, the decision $x^{(0)}$ corresponding to model $M^{(0)}$ is given by $x^{(0)} := \theta_0$. In the simplified model $M^{(1)}$, the decision maker simply assumes $\theta^* = \theta_1$ for some fixed $\theta_1 \in \mathbb{R}$, with corresponding decision $x^{(1)} = \theta_1$, and $\Theta^{(1)} := \{\theta_1\}$. (This represents the situation that $\theta_1$ has a much smaller variance than $\theta_0$).

Let $v := (\sum_{i=1}^n x_i^2)^{-1}$, and observe that $\theta_0$ is normally distributed with mean $\theta^*$ and variance $v$. In what follows, we treat $v$ as a variable and show how $v$ influences the performance of both models $M^{(0)}$ and $M^{(1)}$, and the performance of DBMS. To that end, define the sets

$$S_0 := \{\theta \in \Theta^{(0)} : r(x^{(0)},\theta) \ge r(x^{(1)},\theta)\},$$
$$S_1 := \{\theta \in \Theta^{(0)} : r(x^{(0)},\theta) < r(x^{(1)},\theta)\}.$$

Observe that these sets are random, since they depend (via $\theta_0$) on $\epsilon_1,\dots,\epsilon_t$. In addition, let

$$\Delta_0 := \mathbb{E}\left[\max_{x\in\mathcal{X}} r(x,\theta^*) - r(x^{(0)},\theta^*)\right],$$
$$\Delta_1 := \mathbb{E}\left[\max_{x\in\mathcal{X}} r(x,\theta^*) - r(x^{(1)},\theta^*)\right],$$

be the regret corresponding to model $M^{(0)}$ and model $M^{(1)}$, and let

$$\Delta(v) := \mathbb{E}[r(x^{(1)},\theta^*) - r(x^{(0)},\theta^*)]$$

be the expected gain of using model $M^{(1)}$ instead of model $M^{(0)}$, as function of $v$. In what follows, we write $\mathbb{P}_v(\cdot)$ and $\mathbb{E}_v[\cdot]$ to denote probabilities and expectations that depend on $v$. In the example we consider, $\Delta_0 = v$ and $\Delta_1 = -(\theta_1 - \theta^*)^2$.

In this section we prove four propositions with structural properties and the performance of DBMS in this problem. Our first result shows that both the probability that $M^{(1)}$ outperforms $M^{(0)}$, and the corresponding expected performance gain, are increasing in $v$, but decreasing in $\Delta_1$.

**Proposition 1.** *Both $\mathbb{P}_v(\theta^* \in S_1)$ and $\Delta(v)$ are increasing in $v$ but decreasing in $\Delta_1$.*

**Proof.** Since $\Delta_1 = (x^{(1)} - \theta^*)^2$, we have

$$\begin{aligned}\mathbb{P}_v(\theta^* \in S_1) &= \mathbb{P}_v\big(r(x^{(0)},\theta^*) < r(x^{(1)},\theta^*)\big) \\ &= \mathbb{P}_v\big((\theta_0 - \theta^*)^2 > \Delta_1\big) \\ &= 1 - \int_{\theta^*-\sqrt{\Delta_1}}^{\theta^*+\sqrt{\Delta_1}} \frac{1}{\sqrt{2\pi v}} \exp\left(\frac{-(y-\theta^*)^2}{2v}\right) dy \\ &= 1 - \int_{(\theta^*-\sqrt{\Delta_1})/v}^{(\theta^*+\sqrt{\Delta_1})/v} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(y-\theta^*)^2}{2}\right) dy,\end{aligned}$$

and

$$\begin{aligned}\Delta(v) &= \mathbb{E}_v[r(x^{(1)},\theta^*) - r(x^{(0)},\theta^*)] \\ &= -\Delta_1 + \mathbb{E}_v[(\theta_0 - \theta^*)^2] \\ &= -\Delta_1 + v.\end{aligned}$$

It follows that both $\mathbb{P}_v(\theta^* \in S_1)$ and $\Delta(v)$ are increasing in $v$ but decreasing in $\Delta_1$. □

The resampled data sets that DBMS constructs is of the form $d_r = (x_i, y_i^r)_{1\le i\le n}$, where $y_1^r,\dots,y_n^r \sim N(\theta_0, 1)$. As a result, conditionally on $\theta_0$, the estimate $\theta_r = (\sum_{i=1}^n x_i^2)^{-1}\sum_{i=1}^n x_i y_i^r$ based on resampled data is normally distributed with mean $\theta_0$ and variance $v$.

Our next result shows that the probability that DBMS selects model $M^{(1)}$ is increasing (and in fact differentiable) in $v$.

**Proposition 2.** *The function $(0,\infty) \ni v \mapsto \mathbb{P}_v(\theta_r \in S_1)$ is increasing and differentiable in $v$.*

**Proof.** Let $\Phi$ and $\varphi$ denote the cdf and pdf of the standard normal distribution. For all $t > x^{(1)}$ it holds that

$$\begin{aligned}\mathbb{P}_v(\theta_r \in S_1 | \theta_0 = t) &= \mathbb{P}_v\big(-(t-\theta_r)^2 < -(x^{(1)} - \theta_r)^2 \mid \theta_0 = t\big) \\ &= \mathbb{P}_v\big(t^2 - (x^{(1)})^2 > 2(t - x^{(1)})\theta_r \mid \theta_0 = t\big) \\ &= \mathbb{P}_v\big(\theta_r < (t + x^{(1)})/2 \mid \theta_0 = t\big) \\ &= \Phi((x^{(1)} - t)/(2\sqrt{v})),\end{aligned} \tag{5}$$

since $(\theta_r - t)/\sqrt{v}$ conditional on $\theta_0 = t$ is standard normally distributed, and for all $t < x^{(1)}$,

$$\begin{aligned}\mathbb{P}_v(\theta_r \in S_1 | \theta_0 = t) &= \mathbb{P}_v\big(-(t-\theta_r)^2 < -(x^{(1)} - \theta_r)^2 \mid \theta_0 = t\big) \\ &= \mathbb{P}_v\big(t^2 - (x^{(1)})^2 > 2(t - x^{(1)})\theta_r \mid \theta_0 = t\big) \\ &= \mathbb{P}_v\big(\theta_r > (t + x^{(1)})/2 \mid \theta_0 = t\big) \\ &= 1 - \Phi((x^{(1)} - t)/(2\sqrt{v})).\end{aligned} \tag{6}$$

As a result,

$$\begin{aligned}&\mathbb{P}_v(\theta_r \in S_1) \\ &= \int_{-\infty}^{x^{(1)}} \left(1 - \Phi\left(\frac{x^{(1)} - t}{2\sqrt{v}}\right)\right) \cdot \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(t-\theta^*)^2}{2v}\right) dt \\ &\quad + \int_{x^{(1)}}^{\infty} \Phi\left(\frac{x^{(1)} - t}{2\sqrt{v}}\right) \cdot \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(t-\theta^*)^2}{2v}\right) dt \\ &= \int_{-\infty}^{(x^{(1)}-\theta^*)/\sqrt{v}} \left(1 - \Phi\left(\frac{x^{(1)} - \theta^*}{2\sqrt{v}} - y/2\right)\right) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &\quad + \int_{(x^{(1)}-\theta^*)/\sqrt{v}}^{\infty} \Phi\left(\frac{x^{(1)} - \theta^*}{2\sqrt{v}} - y/2\right) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy,\end{aligned}$$

where we used the variable substitution $y = (t-\theta^*)/\sqrt{v}$. Write $c := x^{(1)} - \theta^*$. By Leibniz's rule for differentiation under the integral sign, it follows that $\mathbb{P}_v(\theta_r \in S_1)$ is differentiable in $v$, for $v > 0$, with derivative equal to

$$\begin{aligned}\frac{d}{dv}\mathbb{P}_v(\theta_r \in S_1) &= -\frac{1}{2}cv^{-3/2}(1-\Phi(0))\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{c^2}{2v}\right)dy\\ &+\int_{-\infty}^{c/\sqrt{v}}\frac{d}{dv}\left(1-\Phi\left(\frac{c}{2\sqrt{v}}-y/2\right)\right)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy\\ &+\frac{1}{2}cv^{-3/2}\Phi(0)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{c^2}{2v}\right)\\ &+\int_{c/\sqrt{v}}^{\infty}\frac{d}{dv}\Phi\left(\frac{c}{2\sqrt{v}}-y/2\right)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy\\ &=\int_{-\infty}^{c/\sqrt{v}}\varphi\left(\frac{c}{2\sqrt{v}}-y/2\right)\cdot\frac{c}{4}v^{-3/2}\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy\\ &-\int_{c/\sqrt{v}}^{\infty}\varphi\left(\frac{c}{2\sqrt{v}}-y/2\right)\cdot\frac{c}{4}v^{-3/2}\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy.\end{aligned}$$

By application of the following claim, with $\gamma = c/\sqrt{v}$, $\tilde{\varphi} = \varphi$, and $\tilde{f}(y) = \frac{1}{4}v^{-1}\frac{1}{\sqrt{2\pi}}\exp(-y^2/2)$ for all $y \in \mathbb{R}$, the statement of the proposition follows.

Claim. Let $\tilde{\varphi} : \mathbb{R} \to [0,\infty)$ and $\tilde{f} : \mathbb{R} \to [0,\infty)$ be symmetric unimodal continuous functions with maximum attained at zero. Then, for all $\gamma \in \mathbb{R}$,

$$\gamma\int_{-\infty}^{\gamma}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy \geq \gamma\int_{\gamma}^{\infty}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy.$$

Proof of Claim. Suppose that $\gamma \geq 0$. Then $\tilde{\varphi}(x/2) = \tilde{\varphi}(-x/2)$ and $\tilde{f}(\gamma - x) \geq \tilde{f}(\gamma + x)$ for all $x \geq 0$. By substitution of variables, we obtain

$$\begin{aligned}\gamma\int_{-\infty}^{\gamma}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy &= \gamma\int_{0}^{\infty}\tilde{\varphi}(x/2)\tilde{f}(\gamma-x)dx\\ &\geq \gamma\int_{0}^{\infty}\tilde{\varphi}(-x/2)\tilde{f}(\gamma+x)dx\\ &= \gamma\int_{\gamma}^{\infty}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy.\end{aligned}$$

Now suppose that $\gamma < 0$. Then $\tilde{\varphi}(x/2) = \tilde{\varphi}(-x/2)$ and $\tilde{f}(\gamma - x) \leq \tilde{f}(\gamma + x)$ for all $x \geq 0$. By substitution of variables, we obtain

$$\begin{aligned}\gamma\int_{-\infty}^{\gamma}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy &= \gamma\int_{0}^{\infty}\tilde{\varphi}(x/2)\tilde{f}(\gamma-x)dx\\ &\geq \gamma\int_{0}^{\infty}\tilde{\varphi}(-x/2)\tilde{f}(\gamma+x)dx\\ &= \gamma\int_{\gamma}^{\infty}\tilde{\varphi}((\gamma-y)/2)\tilde{f}(y)dy.\end{aligned}$$

This completes the proof of the claim. □

Propositions 1 and 2 show that both the probability that DBMS selects $M^{(1)}$, as well as the expected corresponding performance gain $\Delta(v)$, are increasing in $v$. Since $\Delta(v) = v - \Delta_1$ is strictly increasing in $v$, we can also define the probability of selecting $M^{(1)}$ as a function of the performance gain $\Delta$, as follows.

$$p(\Delta) := \mathbb{P}_{\Delta-\Delta_1}(\theta_r \in S_1), \qquad \text{for } \Delta \in (-\Delta_1, \infty).$$

Our next result shows that $p(\Delta)$ is increasing and differentiable in $\Delta$, and provides explicit expressions for the limiting probabilities as $\Delta \downarrow -\Delta_1$ or $\Delta \to \infty$.

**Proposition 3.** *The function $p$ is increasing and differentiable on $(-\Delta_1, \infty)$. In addition, if $\Delta_1 = 0$ then $p(\Delta) = \arctan(2)/\pi$ for all $\Delta \in (\Delta_1, \infty)$, and if $\Delta_1 > 0$, then*

$$\lim_{\Delta\downarrow-\Delta_1} p(\Delta) = 0, \text{ and}$$

$$\lim_{\Delta\to\infty} p(\Delta) = \arctan(2)/\pi \approx 0.3524$$

**Proof.** That $p(\cdot)$ is increasing and differentiable on $(-\Delta_1, \infty)$ follows immediately from Proposition 2. To prove the other statements of the proposition, let $Z_1$, $Z_2$ be independent standard normally distributed random variables. For all $t > x^{(1)}$, Eq. (5) implies

$$\begin{aligned}\mathbb{P}_v(\theta_r \in S_1 \mid \theta_0 = t) &= \mathbb{P}_v\left(Z_1 \leq (x^{(1)}-t)/2\sqrt{v}\right)\\ &= \mathbb{P}_v\left(Z_1 \geq \left|\frac{t-x^{(1)}}{2\sqrt{v}}\right|\right)\\ &= \mathbb{P}_v\left(Z_1 \geq \left|\frac{t-\theta^*}{2\sqrt{v}} + \frac{\theta^*-x^{(1)}}{2\sqrt{v}}\right|\right),\end{aligned}$$

and for all $t < x^{(1)}$, Eq. (6) implies

$$\begin{aligned}\mathbb{P}_v(\theta_r \in S_1 \mid \theta_0 = t) &= \mathbb{P}_v\left(Z_1 \geq (x^{(1)}-t)/2\sqrt{v}\right)\\ &= \mathbb{P}_v\left(Z_1 \geq \left|\frac{t-x^{(1)}}{2\sqrt{v}}\right|\right)\\ &= \mathbb{P}_v\left(Z_1 \geq \left|\frac{t-\theta^*}{2\sqrt{v}} + \frac{\theta^*-x^{(1)}}{2\sqrt{v}}\right|\right).\end{aligned}$$

Since $(\theta_0 - \theta^*)/\sqrt{v}$ is standard normally distributed, it follows that

$$\mathbb{P}_v(\theta_r \in S_1) = \mathbb{P}_v\left(Z_1 \geq \left|\frac{Z_2 + c_v}{2}\right|\right),$$

where we write

$$c_v := \frac{\theta^* - x^{(1)}}{\sqrt{v}}.$$

Suppose $\Delta_1 = 0$. Then $c_v = 0$ for all $v > 0$, and therefore $\mathbb{P}_v(\theta_r \in S_1) = \mathbb{P}_v(Z_1 \geq |Z_2/2|) = \arctan(2)/\pi$ for all $v > 0$.

Now suppose that $\Delta_1 > 0$. Then $c_v \neq 0$ for all $v > 0$, $\lim_{v\downarrow 0}|c_v| = \infty$ and $\lim_{v\to\infty} c_v = 0$, and hence $\lim_{v\downarrow 0}\mathbb{P}_v(\theta_r \in S_1) = 0$ and $\lim_{v\to\infty}\mathbb{P}_v(\theta_r \in S_1) = \mathbb{P}_v(Z_1 \geq |Z_2/2|) = \arctan(2)/\pi$. □

Proposition 3 shows that DBMS satisfies a key structural property of model selection methods: the probability of selecting the (potentially misspecified) model $M^{(1)}$ is increasing in the resulting performance gain $\Delta$. In addition, when the performance gain is minimal ($\Delta \downarrow -\Delta_1$), DBMS correctly selects model $M^{(0)}$ with probability one. Interestingly, if model $M^{(1)}$ happens to be correctly specified, then the probability of selecting model $M^{(1)}$ is independent of $v$ in this example.

Let $Z_1$, $Z_2$ be independent standard normally distributed random variables. The probability that DBMS selects model $M^{(1)}$ can also be written as

$$\begin{aligned}\mathbb{P}_v(\theta_r \in S_1) &= \mathbb{P}_v(Z_1 \geq |(Z_2 + c_v)/2|) = \mathbb{P}_v(Z_1 \geq |(Z_2 + |c_v|)/2|)\\ &= \mathbb{P}_v\left(2Z_1 \geq |Z_2 + \sqrt{\Delta_1/\Delta_0}|\right),\end{aligned}$$

where the first equality is shown in the proof of Proposition 3, the second equality follows by $Z_2 \stackrel{d}{=} -Z_2$, and the third equality follows by $\Delta_0 = \mathbb{E}_v[(\theta_0 - \theta^*)^2] = v$. Observe that $\mathbb{P}_v(2Z_1 \geq |Z_2 + x|)$ is differentiable in $x$, for $x > 0$, with

$$\begin{aligned}\frac{d}{dx}\mathbb{P}_v(2Z_1 \geq |Z_2 + x|) &= \frac{d}{dx}\int_0^{\infty}\int_{-x-2z_1}^{-x+2z_1}\varphi(z_2)\varphi(z_1)dz_2dz_1\\ &= \int_0^{\infty}\{-\varphi(-x+2z_1) + \varphi(-x-2z_1)\}\varphi(z_1)dz_1\\ &= \int_0^{\infty}\{\varphi(x+2z_1) - \varphi(x-2z_1)\}\varphi(z_1)dz_1\\ &< 0,\end{aligned}$$

where $\varphi$ is the pdf of the standard normal distribution, and where the third equality follows by symmetry of $\varphi$. It follows that the

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

probability that DBMS selects model $M^{(1)}$ is a decreasing function of the ratio $\Delta_1/\Delta_0$.

A similar monotonicity property holds for the probability that model $M^{(1)}$ outperforms $M^{(0)}$. From the proof of Proposition 1 we obtain

$$\begin{aligned}\mathbb{P}_\nu(\theta^* \in S_1) &= \mathbb{P}_\nu\big((\theta_0-\theta^*)^2 > \Delta_1\big)\\ &= \mathbb{P}_\nu\big(Z_1^2 > \Delta_1/\nu^2\big) = \mathbb{P}_\nu\big(Z_1^2 > \Delta_1/\Delta_0^2\big),\end{aligned}$$

which clearly is decreasing in $\Delta_1/\Delta_0^2$. In contrast to $\mathbb{P}_\nu(\theta_r \in S_1)$, $\mathbb{P}_\nu(\theta^* \in S_1)$ is monotone in $\Delta_1/\Delta_0^2$ instead of $\Delta_1/\Delta_0$.

The next proposition gives an exact expression for the performance of DBMS.

**Proposition 4.** *Let* $x^{(\text{DBMS})} := x^{(0)}\mathbf{1}\{\theta_r \in S_0\} + x^{(1)}\mathbf{1}\{\theta_r \in S_1\}$. *Let* $Z$ *be a standard normally distributed random variables, and let* $\Phi$ *denote its cdf. Then*

$$\begin{aligned}\mathbb{E}_\nu[(x^{(\text{DBMS})}-\theta^*)^2] = \Delta_0 \cdot \mathbb{E}\left[Z^2\Phi\left(\frac{|Z+\sqrt{\Delta_1/\Delta_0}|}{2}\right)\right]\\ + \Delta_1 \cdot \mathbb{E}\left[1-\Phi\left(\frac{|Z+\sqrt{\Delta_1/\Delta_0}|}{2}\right)\right].\end{aligned}$$

**Proof.** Let $Z_1$, $Z_2$ be standard normally distributed random variables with pdf $\varphi$ and cdf $\Phi$. Let $c_\nu := (\theta^* - x^{(1)})/\sqrt{\nu}$. For all $t \in \mathbb{R}$,

$$\begin{aligned}&\mathbb{E}_\nu[(x^{(\text{DBMS})}-\theta^*)^2 \mid \theta_0 = t]\\ &= \mathbb{E}_\nu[(x^{(0)}-\theta^*)^2 \mid \theta_0 = t, \theta_r \in S_0]\cdot \mathbb{P}_\nu(\theta_r \in S_0 \mid \theta_0 = t)\\ &\quad + \mathbb{E}_\nu[(x^{(1)}-\theta^*)^2 \mid \theta_0 = t, \theta_r \in S_1]\cdot \mathbb{P}_\nu(\theta_r \in S_1 \mid \theta_0 = t)\\ &= (t-\theta^*)^2\cdot(1-\mathbb{P}_\nu(\theta_r \in S_1 \mid \theta_0 = t))\\ &\quad + (x^{(1)}-\theta^*)^2\cdot(\mathbb{P}_\nu(\theta_r \in S_1 \mid \theta_0 = t))\\ &= (t-\theta^*)^2\cdot\mathbb{P}_\nu\left(Z_1 < \left|\frac{t-x^{(1)}}{2\sqrt{\nu}}\right|\right)\\ &\quad + (x^{(1)}-\theta^*)^2\cdot\mathbb{P}_\nu\left(Z_1 \geq \left|\frac{t-x^{(1)}}{2\sqrt{\nu}}\right|\right).\end{aligned}$$

By integrating,

$$\begin{aligned}&\mathbb{E}_\nu[(x^{(\text{DBMS})}-\theta^*)^2]\\ &= \int_{-\infty}^{\infty}(t-\theta^*)^2\cdot\mathbb{P}_\nu\left(Z_1 < \left|\frac{t-x^{(1)}}{2\sqrt{\nu}}\right|\right)\frac{1}{\sqrt{2\pi\nu}}\exp\left(-\frac{(t-\theta^*)^2}{2\nu}\right)dt\\ &\quad + \int_{-\infty}^{\infty}(x^{(1)}-\theta^*)^2\cdot\mathbb{P}_\nu\left(Z_1 \geq \left|\frac{t-x^{(1)}}{2\sqrt{\nu}}\right|\right)\frac{1}{\sqrt{2\pi\nu}}\exp\left(-\frac{(t-\theta^*)^2}{2\nu}\right)dt\\ &= \Delta_0\cdot\int_{-\infty}^{\infty}y^2\mathbb{P}_\nu\left(Z_1 < \left|\frac{y+c_\nu}{2}\right|\right)\varphi(y)dy\\ &\quad + \Delta_1\cdot\int_{-\infty}^{\infty}\mathbb{P}_\nu\left(Z_1 \geq \left|\frac{y+c_\nu}{2}\right|\right)\varphi(y)dy\\ &= \Delta_0\cdot\mathbb{E}_\nu\big[Z_2^2\Phi(|Z_2+c_\nu|/2)\big] + \Delta_1\cdot\mathbb{E}_\nu[1-\Phi(|Z_2+c_\nu|/2)], \qquad (7)\end{aligned}$$

using $\nu = \Delta_0$ and the variable substitution $y = (t-\theta^*)/\sqrt{\nu}$. Now, if $c_\nu \geq 0$ then we can replace $c_\nu$ by $|c_\nu|$ in Eq. (7). If $c_\nu < 0$, then $Z_2 \overset{d}{=} -Z_2$ and $|c_\nu| = -c_\nu$, and we can also replace $c_\nu$ by $|c_\nu$ in (7). Since $|c_\nu| = \sqrt{\Delta_1/\Delta_0}$, it follows that

$$\begin{aligned}\mathbb{E}_\nu[(x^{(\text{DBMS})}-\theta^*)^2]\Delta_0\cdot\mathbb{E}\left[Z_2^2\Phi\left(\frac{|Z_2+\sqrt{\Delta_1/\Delta_0}|}{2}\right)\right]\\ + \Delta_1\cdot\mathbb{E}\left[1-\Phi\left(\frac{|Z_2+\sqrt{\Delta_1/\Delta_0}|}{2}\right)\right].\end{aligned}$$

This proves the proposition. □

The results of Propositions 1–4 are illustrated in Figs. 2 and 3, for $\theta^* = 0.5$ and $x^{(1)} = 0$. In this figure, we write $p^*(\nu) := \mathbb{P}_\nu(\theta^* \in S_1)$. Fig. 2 illustrates the various monotonicity and limiting properties stated in Propositions 1–3. Fig. 3 illustrates that model $M^{(0)}$ is better than model $M^{(1)}$ when $\nu$ is small; the figure also shows that, in that case, the performance of DBMS is close to that of $M^{(0)}$. If $\nu$ is large then model $M^{(1)}$ is better than $M^{(0)}$, and, in that case, DBMS is able to reduce the loss of $M^{(0)}$.

*3.1.3. Difficulty of extending these results to more general decision problems.*

Propositions 1–4 provide detailed insights into the behavior and properties of DBMS and its relation to $\Delta_1$ and $\Delta_0$. Unfortunately, the proofs of these propositions also reveal that it is difficult to extend these insights to more general decision problems. The proofs depend on explicit expressions of the distributions of $\theta_0$, $\theta_r$, and $\theta_1$, which in many applications are not available in closed form. In some problems one might perhaps exploit asymptotic normality of estimators to obtain structural insights, but since we are primarily interested in understanding the finite-sample behavior of DBMS, such asymptotic normality results then need to be accompanied by a good understanding of the corresponding convergence rates. Other complications that arise when one tries to extend these insights to more general decision problems are that the shapes of the sets $S_0$ and $S_1$ can be highly complex, which hampers the analysis of terms like $\mathbb{P}(\theta_r \in S_1)$, and that the distribution of the estimators of different models may depend in a non-trivial way on properties of $x_1, \ldots, x_n$. Despite these difficulties to generalize the structural results from Propositions 1–4, Section 4 suggests that DBMS can successfully be applied to more complex model selection problems.

### 3.2. Alternatives to DBMS

To appreciate DBMS it is useful to consider a few alternatives. A first option is

$$\arg\max_{k\in\{0,1,\ldots,K\}} r(x^{(k)}, \tau^{(k)}(d_r)), \qquad (8)$$

which is defined if $\Theta^{(k)} \subset \Theta^{(0)}$ for all $k$. This method evaluates decision $x^{(k)}$ using the resampled estimate $\tau^{(k)}(d_r)$ instead of $\tau^{(0)}(d_r)$. A disadvantage of this approach is the potential lack of consistency: in general, $\tau^{(k)}(d_r)$ does not have to converge a.s. to $\theta^*$ as $n\to\infty$, even when $\tau^{(0)}(d_r)$ and $\tau^{(0)}(d_0)$ do converge a.s. to $\theta^*$ as $n\to\infty$. As a result, this method may structurally overestimate the performance of one of the simplified models $M^{(k)}$, and thus may select a model whose corresponding decision has a very poor performance when evaluated under the true reward function.

This drawback might perhaps be mitigated by considering

$$\arg\max_{k\in\{0,1,\ldots,K\}} r(x^{(k)}, \tau^{(i_n)}(d_r)), \qquad (9)$$

for some data-dependent $(i_n)_{n\in\mathbb{N}}$ that satisfies $\mathbb{P}(i_n = 0) \to 1$ as $n\to\infty$. A drawback of this method is that it is not clear how this sequence $(i_n)_{n\in\mathbb{N}}$ should be chosen. The 'optimal' way to do this probably depends on the unknown parameters, thus creating an additional source of error in the model selection procedure.

Our model selection method is probabilistic: it is based on a single resampled data set $d_r$, which is a realization from $D(\theta_0)$. Alternatively, one could consider

$$\arg\max_{k\in\{0,1,\ldots,K\}} \mathbb{E}[r(x^{(k)}, \tau^{(0)}(D(\theta_0))) \mid d_0]. \qquad (10)$$

This method was considered in an earlier version of this paper. Although this method yields a non-random model selection method, which perhaps might be preferable for some practitioners, a disadvantage is that it does not work for a large class of problems (including many types of linear programs with parameter

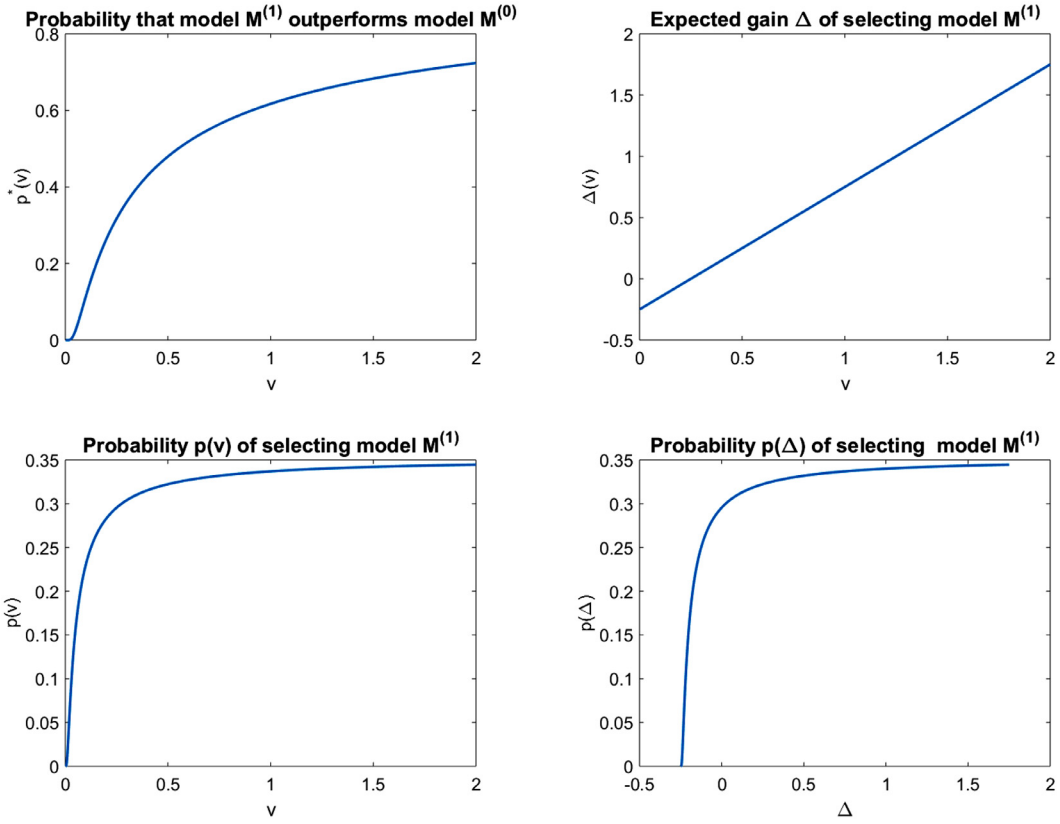


**Fig. 2.** $p^*(v)$, $\Delta(v)$, $p(v)$, and $p(\Delta)$, with $\theta^* = 0.5$ and $x^{(1)} = 0$.

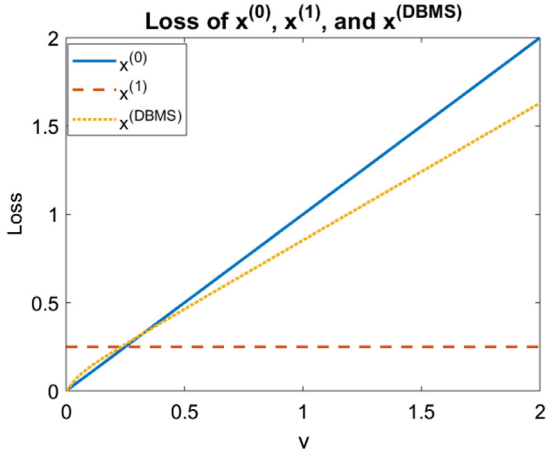


**Fig. 3.** Loss of $x^{(0)}$, $x^{(1)}$, and $x^{(\text{DBMS})}$, with $\theta^* = 0.5$ and $x^{(1)} = 0$.

uncertainty). In particular, if the reward function $r(x, \theta)$ is linear in $\theta$, $\tau^{(0)}$ is an unbiased estimator, and $\chi^{(0)}$ is exact, then $\mathbb{E}[r(x^{(k)}, \tau^{(0)}(D(\theta_0))) \mid d_0] = r(x^{(k)}, \theta_0)$, and (10) is equivalent to simply always choosing model $M^{(0)}$.

A completely different approach is to neglect the decisions suggested by the available simplified models $M^{(k)}$, $k > 0$, and to determine the optimal decision by solving

$$\max_{x \in \mathcal{X}} r(x, \tau^{(0)}(d_r)), \tag{11}$$

or

$$\max_{x \in \mathcal{X}} \mathbb{E}[r(x, \tau^{(0)}(d_r)) \mid d_0]. \tag{12}$$

At first sight Eq. (11) may seem more flexible than DBMS: why would one restrict oneself to $\{x^{(0)}, \ldots, x^{(k)}\}$ when one can optimize over the whole decision space $\mathcal{X}$? The reason is that the decisions generated by the available models are not just arbitrary numbers: the idea is that the simplified models improve upon the true model if the variance of $\theta_0$ is too high; this mitigates the poor quality of $x^{(0)}$ caused by a high variance of $\theta_0$. This property is lost with methods (11) and (12). The method (12) has the additional disadvantages that (i) it is just equivalent to the original optimization problem $\max_{x \in \mathcal{X}} r(x, \theta_0)$ if $\tau^{(0)}$ is unbiased and $r(x, \theta)$ is linear in $\theta$, and (ii) the expectation operator could make the problem more difficult to solve numerically (or even computationally intractable), leading to additional optimization errors.

In the example considered in Section 3.1.2, (11) would imply that decision $x = \theta_r$ is chosen. Since $\theta_r \sim N(\theta_0, v)$ and $\theta_0 \sim N(\theta^*, v)$, the expected loss of this decision equals

$$\begin{aligned}\mathbb{E}[\max_{x \in \mathcal{X}} r(x, \theta^*) - r(\theta_r, \theta^*)] &= \mathbb{E}[(\theta_r - \theta^*)^2] = \mathbb{E}[\mathbb{E}[(\theta_r - \theta^*)^2 \mid \theta_0]] \\ &= \mathbb{E}[\theta_0^2 + v - 2\theta_0\theta^* + (\theta^*)^2] = 2v,\end{aligned}$$

which is twice the loss of the decision $x^{(0)}$ corresponding to model $M^{(0)}$. This shows that (11) is worse than simply using model $M^{(0)}$. In the same example, decision rule (12) would imply that decision $x = x^{(0)}$ is chosen, i.e. that model $M^{(0)}$ is always followed, even if its performance is much worse than that of model $M^{(1)}$.

### 3.3. Consistency

DBMS is reward-consistent: under some conditions, the loss in reward caused by DBMS not selecting the best available model converges in probability to zero as the data size grows large. To formally state this property, we introduce some notation.

Let $(x_n)_{n \in \mathbb{N}}$ be an infinite sequence in $\mathcal{X}$, let $D_n(\theta) := (x_1, Y(x_1, \theta), \ldots, x_n, Y(x_n, \theta))$ for $n \in \mathbb{N}$ and $\theta \in \Theta^{(0)}$, and let $\|\cdot\|_\vartheta$ be a norm on $\Theta^{(0)}$. Let $\theta_k(n) = \tau^{(k)}(D_n(\theta^*))$ denote the estimate corresponding to model $M^{(k)}$ based on data $D_n(\theta^*)$, and

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

let $x^{(k)}(n) := \chi^{(k)}(\theta_k(n))$ be the optimal decision according to model $M^{(k)}$ at stage $n$, for $k = 0, 1, \ldots, K$ and $n \in \mathbb{N}$. Let $\theta_r(n) = \tau^{(0)}(D_n(\theta_0(n)))$ be the resampled estimator using $n$ data points, and let

$$k^{(\text{DBMS})}(n) := \arg\max_{k\in\{0,1,\ldots,K\}} r(x^{(k)}(n), \theta_r(n)))$$

be the model selected by DBMS at stage $n$, with ties decided in favor of the smallest maximizer. Finally, let

$$r^{(\text{DBMS})}(n) := r(x^{(k^{(\text{DBMS})}(n))}, \theta^*)$$

be the corresponding reward, and let

$$r^*(n) := \max_{k\in\{0,1,\ldots,K\}} r(x^{(k)}(n), \theta^*)$$

be the reward using the best of the available models at stage $n$.

**Proposition 5.** *Suppose that $||\theta_r(n) - \theta^*||_\vartheta$ converges in probability to zero, $r(\cdot, \cdot)$ is continuous in both variables, and $x^{(k)}(n)$ converges in probability as $n \to \infty$, for each $k = 0, 1, \ldots, K$. Then*

$$|r^{(\text{DBMS})}(n) - r^*(n)| \xrightarrow{P} 0 \text{ as } n \to \infty. \tag{13}$$

*In particular, if $x^{(k)}(n)$ converges in probability to some $x^* \in \arg\max_{x\in\mathcal{X}} r(x, \theta^*)$ as $n \to \infty$, for some $k \in \{0, 1, \ldots, K\}$, then*

$$r^{(\text{DBMS})}(n) \xrightarrow{P} r(x^*, \theta^*) \text{ as } n \to \infty. \tag{14}$$

**Proof.** Let $\epsilon > 0$, and let $k^*(n)$ be the smallest maximizer of $r(x^{(k)}, \theta^*)$ w.r.t. $k$.

$$\begin{aligned}
&\mathbb{P}\left(|r^{(\text{DBMS})}(n) - r^*(n)| > \epsilon\right)\\
&= \sum_{k=0}^{K} \mathbb{P}\left(k^{(\text{DBMS})}(n) = k \text{ and } r(x^{(k)}(n), \theta^*) < r(x^{(k^*(n))}(n), \theta^*) - \epsilon\right)\\
&\le \sum_{k=0}^{K} \mathbb{P}\left(\begin{array}{l} r(x^{(k)}(n), \theta_r(n)) \ge r(x^{(k^*(n))}(n), \theta_r(n))\\ \text{and } r(x^{(k)}(n), \theta^*) < r(x^{(k^*(n))}(n), \theta^*) - \epsilon \end{array}\right)\\
&\le \sum_{k=0}^{K} \mathbb{P}\left(\begin{array}{l} r(x^{(k)}(n), \theta_r(n)) - r(x^{(k)}(n), \theta^*) + r(x^{(k)}(n), \theta^*)\\ \ge r(x^{(k^*(n))}(n), \theta_r(n)) - r(x^{(k^*(n))}(n), \theta^*) + r(x^{(k^*(n))}(n), \theta^*)\\ \text{and } r(x^{(k)}(n), \theta^*) < r(x^{(k^*(n))}(n), \theta^*) - \epsilon \end{array}\right).
\end{aligned} \tag{15}$$

Fix $k \in \{0, 1, \ldots, K\}$, and let $x^{(k)}(\infty)$ be the limit point of $x^{(k)}(n)$ as $n \to \infty$. Since $x^{(k)}(n) \xrightarrow{P} x^{(k)}(\infty)$, $||\theta_r(n) - \theta^*||_\vartheta \xrightarrow{P} 0$, and $r(\cdot, \cdot)$ is continuous in both variables, it follows that

$$r(x^{(k)}(n), \theta_r(n)) - r(x^{(k)}(n), \theta^*) \xrightarrow{P} 0.$$

Since

$$\begin{aligned}
&|r(x^{(k^*(n))}(n), \theta_r(n)) - r(x^{(k^*(n))}(n), \theta^*)|\\
&\quad\le \sup_{l\in\{0,1,\ldots,K\}} |r(x^{(l)}(n), \theta_r(n)) - r(x^{(l)}(n), \theta^*)|,
\end{aligned}$$

this implies that also

$$|r(x^{(k^*(n))}(n), \theta_r(n)) - r(x^{(k^*(n))}(n), \theta^*)| \xrightarrow{P} 0.$$

It follows that (15) converges to

$$\sum_{k=0}^{K} \mathbb{P}\left(\begin{array}{l} r(x^{(k)}(n), \theta^*) \ge r(x^{(k^*(n))}(n), \theta^*) \text{ and}\\ r(x^{(k)}(n), \theta^*) < r(x^{(k^*(n))}(n), \theta^*) - \epsilon \end{array}\right) = 0.$$

This implies the first statement of the proposition. The second statement follows from observing that $x^{(k)}(n) \xrightarrow{P} x^*$ implies $r^*(n) \xrightarrow{P} r(x^*, \theta^*)$, since $r(\cdot, \cdot)$ is continuous. □

Observe that the statement of Proposition 5 is in terms of the rewards, and not in terms of the probability that DBMS selects the best available model. The reason is that it may happen that $r^{(\text{DBMS})}(n) < r^*(n)$ a.s. for all $n \in \mathbb{N}$, while both $r^{(\text{DBMS})}(n)$ and $r^*(n)$ converge in probability to $r(x^*, \theta^*)$. This can occur e.g. if both the true model $M^{(0)}$ and one of the simplified models, say $M^{(1)}$, are correctly specified, and the reward $r(x^{(1)}(n), \theta^*)$ converges faster to $r(x^*, \theta^*)$ than $r(x^{(0)}(n), \theta^*)$. It may then happen that the probability that DBMS selects the best available model (i.e. model $M^{(1)}$) does not converge to one, but that the reward using DBMS still converges to the optimal reward. Of course, if there is only a single model $M^{(k)}$ with $r(x^{(k)}(n), \theta^*) \xrightarrow{P} r(x^*, \theta^*)$, then Eq. (14) implies that the probability that DBMS selects this model does converge to one as $n$ grows large.

Ideally we would like to be able to give a finite-sample performance guarantee for DBMS. However, as already alluded to in Section 3.1, it is very difficult to state such a result in a general setting, without making further assumptions on the models, estimators, and optimization algorithms. In Section 4 we provide a numerical study of the finite-sample performance of DBMS and two alternative methods, for two well-known business optimization problems.

### 3.4. Further remarks

It is worth emphasizing that DBMS (and in fact any model selection method) can only be effective if a simplified model may outperform the true model with some positive probability. If this is not the case, then always using the true model is better than any model selection method that deviates from the true model with positive probability. Decision-based model selection is not a magic bullet: its effectiveness depends not only on the quality of the selection method, but also on the quality of the simplified models under consideration, in particular their ability to generate better decisions than the true model for some initial data sets. For example, the fact that a simple multinomial logit (MNL) model outperforms a sophisticated machine learning model in Feldman et al. (2019) strongly suggests that the MNL model captures at least some of the essential structure of the problem.

Finally, our analysis of DBMS assumes that model $M^{(0)}$ is correctly specified: there is some 'true' parameter $\theta^* \in \Theta^{(0)}$. The assumption that a posited model is correctly specified for a given data sequence is standard in the statistics literature and perhaps unavoidable for purposes of analysis; for example, practically all consistency and convergence-rate results on estimators are only meaningful in practice if one assumes that the data is generated by the postulated model, or perhaps by a model that, in some sense, is 'close' to the postulated model. In real-life applications, however, it is reasonable to expect that even $M^{(0)}$ is not correctly specified. Nothing hinders a decision maker to still apply DBMS in this case; even our consistency result in Section 3.3 remains valid (the proof of Proposition 5 uses nowhere the fact that $\theta^*$ is the 'true' parameter). Of course, 'reward-consistency' should then not be interpreted as convergence to the 'optimal reward' per se, but as convergence to $r(\chi^{(0)}(\theta^*), \theta^*)$, the optimal reward with respect to model $M^{(0)}$ and parameter $\theta^*$. The quality of the decision $\chi^{(0)}(\theta^*)$ (compared to the optimal decision with respect to the 'true' data-generating mechanism) depends of course on how accurate the (now misspecified) model $M^{(0)}$ describes the true data-generating mechanism.

Unfortunately, it is in general not possible to know with certainty whether ones model is correctly specified: it is possible to construct examples where an adversarial Nature tries to make a decision maker believe in her model and corresponding optimal decision, whilst at the same time a different, a better decision is available. This, and related questions about detecting the validity of the most general model that one has at ones disposal, is outside the scope of this paper.

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

## 4. Numerical illustrations

We illustrate the performance of DBMS by applying it to two well-known business optimization problems: assortment optimization and the newsvendor problem. These are two different types of problems. The first is a discrete optimization problem where the unknown parameter is finite-dimensional, and where the distributions of the random observations $Y(x)$ depend on the decisions $x$. The second is a continuous optimization problem where the unknown parameter is infinite-dimensional, and where the distributions of $Y(x)$ are independent of $x$.

### 4.1. Assortment optimization

Assortment optimization consists of determining which set ('assortment') of products a firm should offer to potential customers in order to maximize expected revenue.

*Setting.* A seller offers a subset (called an 'assortment') of $m \in \mathbb{N}$ products $\{1, \ldots, m\}$ for sale to its potentials customers. Selling a single item of product $j$ gives revenue $r_j$ to the firm, for some $r_1, \ldots, r_m > 0$. Upon being offered an assortment, a customer either buys nothing, in which case the firm earns nothing, or buys exactly one of the products, say product $j$, from the assortment, in which case the firm earns $r_j$.

A decision corresponds to a nonempty subset $x \subset \{1, \ldots, m\}$, and the set of feasible decisions $\mathcal{X}$ is the collection of all such subsets. Let $Y(x)$ denote the product that a customer buys when being offered assortment $x \in \mathcal{X}$. For each assortment $x$, $Y(x)$ is multinomially distributed on $x \cup \{0\}$; here $Y(x) = 0$ corresponds to buying nothing. The probability distribution of $Y(x)$ is given by $\mathbb{P}(Y(x) = j) = \theta^*_{j,x}$, for all $j \in x$ and $x \in \mathcal{X}$, and $\mathbb{P}(Y(x) = 0) = 1 - \sum_{j \in x} \theta^*_{j,x}$ for all $x \in \mathcal{X}$, for some unknown parameter $\theta^* = (\theta^*_{j,x} \mid j \in x, x \in \mathcal{X})$ in the parameter space

$$\Theta^{(0)} = \{(\theta_{j,x} \mid j \in x, x \in \mathcal{X}) \mid 0 \le \theta_{j,x} \le \sum_{i \in x} \theta_{i,x} \le 1 \text{ for all } j \in x \text{ and } x \in \mathcal{X}\}.$$

We deliberately keep $\Theta^{(0)}$ very general, without imposing assumptions such as $\theta_{j,x} \le \theta_{j,x'}$ when $j \in x' \subset x$. The expected reward function $r : \mathcal{X} \times \Theta^{(0)} \to \mathbb{R}$ is given by

$$r(x, \theta) = \sum_{i \in x} r_i \theta_{i,x}.$$

The estimator $\tau^{(0)}$ maps data $d = (x_1, y_1, \ldots, x_n, y_n)$ to

$$\tau^{(0)}(d) = (\tau^{(0)}_{j,x}(d) \mid j \in x, x \in \mathcal{X}),$$

where

$$\tau^{(0)}_{j,x}(d) = \frac{\left|\{i \in \{1, \ldots, n\} : (x_i, y_i) = (x, j)\}\right| + 1}{\left|\{i \in \{1, \ldots, n\} : x_i = x\}\right| + 1}.$$

Here $|A|$ denotes the cardinality of a set $A$. This is a small modification to the ordinary relative-frequency estimator $|\{i \in \{1, \ldots, n\} : (x_i, y_i) = (x, j)\}| \big/ |\{i \in \{1, \ldots, n\} : x_i = x\}|$; because this latter expression is undefined if the denominator equals zero, we add one to the frequency of each alternative (including the no-purchase option).

The simplified model $M^{(1)}$ assumes that customers choose according to the so-called multinomial logit model. This is a widely used discrete-choice model that exhibits certain pleasant properties (such as a concave likelihood function) but is known to be misspecified in several cases (illustrated, for example, by the infamous 'red bus / blue bus paradox'). According to the multinomial logit model, $Y(x)$ is multinomially distributed on $x \cup \{0\}$, for all $x \in \mathcal{X}$, with choice probabilities $\mathbb{P}_\theta(Y(x) = j) = 0$ for all $j \notin x$, and

$$\mathbb{P}_\theta(Y(x) = j) = \frac{\exp(\theta_j)}{1 + \sum_{k \in x} \exp(\theta_k)},$$
$$\mathbb{P}_\theta(Y(x) = 0) = \frac{1}{1 + \sum_{k \in x} \exp(\theta_k)}, \tag{16}$$

for all $j \in x$. The unknown parameter $\theta = (\theta_1, \ldots, \theta_m)$ is assumed to lie in $\Theta^{(1)} = \mathbb{R}^m$, and is estimated with maximum likelihood estimation.

For both the true model $M^{(0)}$ and the simplified model $M^{(1)}$, optimization is exact and is done by comparing the revenues of all possible assortments. (Note that this is only computationally tractable if $m$ is not too large).

*Numerical experiments.* For each number of products $m \in \{3, 5, 10\}$ and each size of the initial data set $n \in \{10, 20, 50, 100, 200, 500, 1000, 2000, 5000, 10,000, 20,000, 50,000\}$ we run 10,000 simulations. In each simulation we run three experiments:

- in experiment A, the choice probabilities $\theta$ are drawn uniformly at random, as follows: for each $l \in \{1, \ldots, m\}$ and for each assortment $x$ consisting of exactly $l$ products, the choice probabilities $\{\theta_{j,x} | j \in x \cup \{0\}\}$ are drawn uniformly at random from the $(l+1)$-dimensional simplex $\Delta_{l+1} := \{(z_1, \ldots, z_{l+1}) \in \mathbb{R}^{l+1} \mid \sum_{j=1}^{l+1} z_j = 1, z_1, \ldots, z_{l+1} \ge 0\}$.
- in experiment B, the choice probabilities $\theta$ follow a parsimonious Generalized Attraction Model (Gallego, Ratliff, & Shebalov, 2015): we draw random $\eta, v_1, \ldots, v_m$ from the uniform distribution on (0,1), and set
$$\theta_{j,x} = \frac{v_j}{1 + \sum_{i \in x} v_i + \eta \sum_{i \notin x} v_i}, \quad (j \in x, x \in \mathcal{X}).$$
- in experiment C, the choice probabilities $\theta$ follow a multinomial logit model: we draw random $v_1, \ldots, v_m$ from the uniform distribution on (0,1), and set
$$\theta_{j,x} = \frac{v_j}{1 + \sum_{i \in x} v_i}, \quad (j \in x, x \in \mathcal{X}).$$

In experiment A, the multinomial logit model $M^{(1)}$ is almost surely misspecified, whereas in experiment C it is always correctly specified. Experiment B is somewhat in between: the multinomial logit model is misspecified, but, especially if $\eta$ is small, the choice probabilities are almost of the form (16). This suggests that, for sufficiently small $n$, model $M^{(1)}$ may produce good decisions in experiment B, despite the fact that the model is misspecified.

The revenues corresponding to the individual products are set to $r_i = 100 \cdot i/m$, $i = 1, \ldots, m$. The assortments $x_1, \ldots, x_n$ in the initial data $d_0$ are chosen uniformly at random from $\mathcal{X}$.

For each experiment we determine the optimal revenue under full information (Opt), under model $M^{(0)}$, model $M^{(1)}$, and under DBMS. We also test two alternative model-selection methods: Akaike Information Criterion (AIC), and 5-fold Cross-Validation (CV) on the estimated reward function. In particular, AIC chooses the model that minimizes AIC($k$), $k \in \{0, 1\}$, where ties are decided in favor of model $M^{(0)}$, and where

$$\text{AIC}(0) := -2 \sum_{i=1}^{n} \log \tau^{(0)}_{y_i, x_i}(d_0) + 2m2^{m-1}, \tag{17}$$

$$\text{AIC}(1) := -2 \sum_{i=1}^{n} \log \mathbb{P}_{\tau^{(1)}(d_0)}(Y(x_i) = y_i) + 2m; \tag{18}$$

here $\tau^{(1)}(d_0)$ is the maximum likelihood estimator of the unknown parameters in model $M^{(1)}$. Regarding (17), note that the number of free parameters in model $M^{(0)}$ is equal to $\sum_{i=1}^{m} i\binom{m}{i} = m2^{m-1}$.

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

**Table 1**
Average reward, 3 products.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 68.98 | 59.60 | 55.46 | 58.59 | 55.46 | 58.67 |
| 20 | 68.98 | 61.73 | 55.93 | 60.34 | 55.96 | 60.51 |
| 50 | 68.98 | 64.63 | 56.37 | 63.40 | 60.55 | 63.29 |
| 100 | 68.98 | 66.37 | 56.39 | 65.60 | 65.40 | 65.52 |
| 200 | 68.98 | 67.52 | 56.36 | 67.07 | 67.38 | 67.18 |
| 500 | 68.98 | 68.36 | 56.34 | 68.15 | 68.36 | 68.31 |
| 1000 | 68.98 | 68.67 | 56.26 | 68.57 | 68.67 | 68.66 |
| 2000 | 68.98 | 68.82 | 56.28 | 68.76 | 68.82 | 68.82 |
| 5000 | 68.98 | 68.92 | 56.31 | 68.90 | 68.92 | 68.92 |
| 10000 | 68.98 | 68.95 | 56.29 | 68.94 | 68.95 | 68.95 |
| 20000 | 68.98 | 68.97 | 56.31 | 68.96 | 68.97 | 68.97 |
| 50000 | 68.98 | 68.98 | 56.31 | 68.98 | 68.98 | 68.98 |

(a) Experiment A: general choice model

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 38.97 | 33.66 | 36.95 | 34.81 | 36.95 | 35.93 |
| 20 | 38.97 | 34.26 | 37.45 | 35.56 | 37.45 | 36.53 |
| 50 | 38.97 | 35.39 | 37.67 | 36.22 | 37.66 | 37.22 |
| 100 | 38.97 | 36.33 | 37.69 | 36.84 | 37.67 | 37.52 |
| 200 | 38.97 | 37.11 | 37.70 | 37.38 | 37.70 | 37.65 |
| 500 | 38.97 | 37.89 | 37.71 | 37.96 | 37.89 | 37.93 |
| 1000 | 38.97 | 38.29 | 37.70 | 38.25 | 38.18 | 38.18 |
| 2000 | 38.97 | 38.55 | 37.69 | 38.49 | 38.51 | 38.48 |
| 5000 | 38.97 | 38.76 | 37.69 | 38.71 | 38.77 | 38.76 |
| 10000 | 38.97 | 38.86 | 37.69 | 38.82 | 38.87 | 38.86 |
| 20000 | 38.97 | 38.90 | 37.69 | 38.88 | 38.91 | 38.90 |
| 50000 | 38.97 | 38.94 | 37.69 | 38.93 | 38.94 | 38.94 |

(b) Experiment B: generalized attraction model

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 40.44 | 37.11 | 39.51 | 38.19 | 39.51 | 38.29 |
| 20 | 40.44 | 37.21 | 39.96 | 38.38 | 39.96 | 38.84 |
| 50 | 40.44 | 37.73 | 40.27 | 38.66 | 40.25 | 39.57 |
| 100 | 40.44 | 38.15 | 40.37 | 38.90 | 40.31 | 39.89 |
| 200 | 40.44 | 38.63 | 40.40 | 39.19 | 40.32 | 40.06 |
| 500 | 40.44 | 39.23 | 40.43 | 39.58 | 40.35 | 40.19 |
| 1000 | 40.44 | 39.59 | 40.43 | 39.84 | 40.38 | 40.27 |
| 2000 | 40.44 | 39.87 | 40.43 | 40.04 | 40.41 | 40.33 |
| 5000 | 40.44 | 40.13 | 40.44 | 40.22 | 40.42 | 40.37 |
| 10000 | 40.44 | 40.26 | 40.44 | 40.31 | 40.43 | 40.40 |
| 20000 | 40.44 | 40.34 | 40.44 | 40.37 | 40.43 | 40.42 |
| 50000 | 40.44 | 40.40 | 40.44 | 40.41 | 40.44 | 40.43 |

(c) Experiment C: multinomial logit model

**Table 2**
Average reward, 5 products.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 77.70 | 61.41 | 58.43 | 60.78 | 58.43 | 60.28 |
| 20 | 77.70 | 62.47 | 59.75 | 61.61 | 59.75 | 61.60 |
| 50 | 77.70 | 64.96 | 60.42 | 63.53 | 60.42 | 63.71 |
| 100 | 77.70 | 67.67 | 60.50 | 65.95 | 60.95 | 66.00 |
| 200 | 77.70 | 70.67 | 60.61 | 69.12 | 60.63 | 69.08 |
| 500 | 77.70 | 74.04 | 60.70 | 73.29 | 73.88 | 73.72 |
| 1000 | 77.70 | 75.83 | 60.81 | 75.38 | 75.83 | 75.82 |
| 2000 | 77.70 | 76.75 | 60.86 | 76.53 | 76.75 | 76.75 |
| 5000 | 77.70 | 77.32 | 60.93 | 77.24 | 77.32 | 77.32 |
| 10000 | 77.70 | 77.51 | 60.90 | 77.49 | 77.51 | 77.51 |
| 20000 | 77.70 | 77.61 | 60.88 | 77.59 | 77.61 | 77.61 |
| 50000 | 77.70 | 77.66 | 60.93 | 77.66 | 77.66 | 77.66 |

(a) Experiment A: general choice model

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 43.14 | 36.91 | 40.28 | 37.63 | 40.28 | 39.17 |
| 20 | 43.14 | 35.89 | 41.02 | 37.27 | 41.02 | 39.57 |
| 50 | 43.14 | 35.41 | 41.46 | 37.63 | 41.46 | 40.61 |
| 100 | 43.14 | 36.20 | 41.57 | 38.32 | 41.57 | 41.14 |
| 200 | 43.14 | 37.52 | 41.66 | 39.04 | 41.66 | 41.49 |
| 500 | 43.14 | 39.36 | 41.71 | 40.07 | 41.71 | 41.65 |
| 1000 | 43.14 | 40.48 | 41.72 | 40.89 | 41.72 | 41.70 |
| 2000 | 43.14 | 41.30 | 41.73 | 41.46 | 41.73 | 41.78 |
| 5000 | 43.14 | 42.07 | 41.75 | 42.08 | 41.95 | 42.22 |
| 10000 | 43.14 | 42.46 | 41.76 | 42.43 | 42.5 | 42.59 |
| 20000 | 43.14 | 42.71 | 41.76 | 42.65 | 42.78 | 42.78 |
| 50000 | 43.14 | 42.92 | 41.75 | 42.88 | 42.95 | 42.95 |

(b) Experiment B: generalized attraction model

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 46.83 | 44.23 | 45.32 | 44.83 | 45.32 | 44.36 |
| 20 | 46.83 | 43.11 | 46.12 | 44.34 | 46.12 | 44.37 |
| 50 | 46.83 | 42.20 | 46.58 | 44.14 | 46.58 | 45.26 |
| 100 | 46.83 | 42.16 | 46.71 | 44.10 | 46.71 | 45.86 |
| 200 | 46.83 | 42.59 | 46.78 | 44.13 | 46.78 | 46.49 |
| 500 | 46.83 | 43.54 | 46.81 | 44.53 | 46.81 | 46.73 |
| 1000 | 46.83 | 44.28 | 46.82 | 44.98 | 46.82 | 46.80 |
| 2000 | 46.83 | 44.97 | 46.83 | 45.48 | 46.83 | 46.81 |
| 5000 | 46.83 | 45.68 | 46.83 | 45.99 | 46.83 | 46.82 |
| 10000 | 46.83 | 46.06 | 46.83 | 46.28 | 46.83 | 46.83 |
| 20000 | 46.83 | 46.35 | 46.83 | 46.49 | 46.83 | 46.83 |
| 50000 | 46.83 | 46.57 | 46.83 | 46.65 | 46.83 | 46.83 |

(c) Experiment C: multinomial logit model

CV chooses the model that minimizes $\mathrm{CV}(k)$, $k \in \{0, 1\}$, where ties are decided in favor of model $M^{(0)}$, and where

$$\mathrm{CV}(k) := \frac{1}{5}\sum_{l=1}^{5}\sqrt{\sum_{i\in d_l}\left(\hat{r}^{(k)}(x_i; d_0\backslash d_l) - r_{y_i}\right)^2}; \tag{19}$$

here

$$\hat{r}^{(0)}(x; d) := \sum_{j\in x} r_j \tau^{(0)}_{j,x}(d),$$

$$\hat{r}^{(1)}(x; d) := \sum_{j\in x} r_j \mathbb{P}_{\tau^{(1)}(d)}(Y(x) = j),$$

are the estimated reward functions under model $M^{(0)}$ and $M^{(1)}$ respectively, based on a data set $d$; $r_0 := 0$, and $d_l := \{(x_i, y_i) \mid i = 1 + (l-1)n/5, \ldots, ln/5\}$, $l = 1, \ldots, 5$, is a decomposition of the initial data set $d_0$ in five mutually disjoint sets of equal size.

The average rewards in the simulations are reported in Tables 1–3. All standard errors are smaller than 0.18.

*Outcomes.* In experiment A, model $M^{(0)}$ outperforms model $M^{(1)}$ for all tested value of $n$ and $m$. The average reward under model $M^{(0)}$ converges to the optimal average reward as $n$ grows large, but the average reward under model $M^{(1)}$ appears to converge to a strictly lower value. The average loss due to using model $M^{(1)}$ instead of model $M^{(0)}$ can be more than 20% (if $n \geq 2000$ and $m = 5$, or if $n = 50,000$ and $m = 10$). DBMS is at most 2.5% ($m = 5$, $n = 100$) away from the average performance of $M^{(0)}$, and, for sufficiently large $n$ ($n \geq 200$ in case $m = 3$, $n \geq 1000$ in case $m = 5$, and $n \geq 20000$ in case $m = 10$), DBMS is within 1.0 percent of the average reward under model $M^{(0)}$. Interestingly, DBMS may yield a larger average reward than both $M^{(0)}$ and $M^{(1)}$; this occurs for $m = 10$ and $n = 10, 100, 200$.

All three model selection criteria DBMS, AIC, and CV outperform the other two of these criteria in some instances: DBMS for $m = 3$, $n = 50, 100$ and $m = 5$, $n = 10, 20, 200$ and $m = 10$, $n \leq 2000$; AIC for $m = 3$, $n = 200, 500, 1000, 2000$ and $m = 5$, $n = 500, 1000$; and CV for $m = 3$, $n = 10, 20$ and $m = 5$, $n = 50, 100$ and $m = 10$, $n = 5000, 10000, 20000$. For all other combinations of $m$ and $n$ there is no single best model selection method among these three.

AIC can lose up to 13.2% of the reward of DBMS ($m = 10$, $n = 10,000$), but DBMS loses never more than 1.0% of the reward obtained by AIC ($m = 10$, $n = 20,000$). The performance of CV and DBMS are closer to each other: CV is at most 1.7% worse than DBMS ($m = 10$, $n = 10$), and DBMS is at most 1.8 percent worse than CV ($m = 10$, $n = 10,000$).

In experiment B, model $M^{(1)}$ may outperform model $M^{(0)}$. This occurs if $m = 3$ and $n \leq 200$, $m = 5$ and $n \leq 2000$, or $m = 10$ and $n$ is any of the tested values. For all other pairs $m$, $n$, model $M^{(0)}$ is better than $M^{(1)}$. The loss of using model $M^{(0)}$ instead of model $M^{(1)}$ can be more than 30 percent ($m = 10$, $n = 1000, 2000$), and conversely, the loss due to using model $M^{(1)}$ instead of model $M^{(0)}$ can be up to 3.2 percent ($m = 3$, $n = 50,000$).

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

**Table 3**
Average reward, 10 products.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
|---|---|---|---|---|---|---|
| 10 | 86.64 | 68.05 | 64.66 | 68.05 | 64.66 | 66.87 |
| 20 | 86.64 | 68.05 | 66.41 | 68.00 | 66.41 | 67.51 |
| 50 | 86.64 | 68.01 | 67.47 | 67.99 | 67.47 | 67.80 |
| 100 | 86.64 | 67.97 | 67.68 | 68.03 | 67.68 | 67.84 |
| 200 | 86.64 | 67.90 | 67.77 | 68.05 | 67.77 | 67.76 |
| 500 | 86.64 | 68.91 | 67.95 | 68.77 | 67.95 | 68.50 |
| 1000 | 86.64 | 70.09 | 68.01 | 69.37 | 68.01 | 69.20 |
| 2000 | 86.64 | 72.21 | 68.08 | 70.87 | 68.08 | 70.65 |
| 5000 | 86.64 | 76.23 | 68.06 | 74.55 | 68.06 | 75.12 |
| 10000 | 86.64 | 79.93 | 68.07 | 78.45 | 68.07 | 79.91 |
| 20000 | 86.64 | 82.96 | 68.07 | 82.17 | 82.96 | 82.96 |
| 50000 | 86.64 | 85.12 | 68.07 | 84.94 | 85.12 | 85.12 |
| (a) Experiment A: general choice model | | | | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
| 10 | 48.08 | 38.25 | 44.01 | 38.30 | 44.01 | 42.65 |
| 20 | 48.08 | 38.15 | 45.17 | 38.29 | 45.17 | 44.19 |
| 50 | 48.08 | 37.84 | 45.72 | 38.21 | 45.72 | 45.41 |
| 100 | 48.08 | 37.36 | 45.85 | 38.14 | 45.85 | 45.73 |
| 200 | 48.08 | 36.62 | 45.94 | 37.91 | 45.94 | 45.90 |
| 500 | 48.08 | 35.59 | 45.98 | 37.87 | 45.98 | 45.97 |
| 1000 | 48.08 | 35.14 | 46.00 | 38.02 | 46.00 | 46.00 |
| 2000 | 48.08 | 35.37 | 46.00 | 38.51 | 46.00 | 46.00 |
| 5000 | 48.08 | 37.49 | 46.00 | 40.12 | 46.00 | 46.00 |
| 10000 | 48.08 | 40.01 | 45.99 | 41.75 | 45.99 | 45.99 |
| 20000 | 48.08 | 42.44 | 45.99 | 43.27 | 45.99 | 45.99 |
| 50000 | 48.08 | 44.73 | 45.99 | 45.02 | 45.99 | 45.99 |
| (b) Experiment B: generalized attraction model | | | | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | AIC | CV |
| 10 | 56.42 | 55.26 | 53.99 | 55.30 | 53.99 | 54.70 |
| 20 | 56.42 | 55.08 | 55.43 | 55.22 | 55.43 | 55.25 |
| 50 | 56.42 | 54.59 | 56.08 | 54.97 | 56.08 | 55.56 |
| 100 | 56.42 | 53.93 | 56.25 | 54.60 | 56.25 | 55.75 |
| 200 | 56.42 | 52.90 | 56.34 | 54.09 | 56.34 | 55.92 |
| 500 | 56.42 | 51.52 | 56.38 | 53.32 | 56.38 | 56.23 |
| 1000 | 56.42 | 50.82 | 56.40 | 53.07 | 56.40 | 56.35 |
| 2000 | 56.42 | 50.40 | 56.41 | 52.81 | 56.41 | 56.41 |
| 5000 | 56.42 | 50.60 | 56.41 | 52.75 | 56.41 | 56.41 |
| 10000 | 56.42 | 51.30 | 56.42 | 52.90 | 56.42 | 56.42 |
| 20000 | 56.42 | 52.20 | 56.42 | 53.27 | 56.42 | 56.42 |
| 50000 | 56.42 | 53.46 | 56.42 | 54.07 | 56.42 | 56.42 |
| (c) Experiment C: multinomial logit model | | | | | | |

In the cases that $M^{(1)}$ is better than $M^{(0)}$, DBMS can improve upon $M^{(0)}$ by more than 8 percent ($m = 10$, $n = 1000, 2000$). In the cases that model $M^{(0)}$ is better than $M^{(1)}$, DBMS loses never more than 0.2% of the average reward of $M^{(0)}$. Again we see that DBMS may yield a larger average reward than both $M^{(0)}$ and $M^{(1)}$; this occurs for $m = 3$, $n = 500$ and $m = 5$, $n = 5000$.

Again all three model selection criteria DBMS, AIC, and CV outperform the other two of these criteria in some instances: DBMS for $m = 3$, $n = 500, 1000$; AIC for $m = 3$, $n \notin \{500, 1000\}$ and $m = 5$, $n \leq 1000$ or $n = 50,000$, and $m = 10$, $n \leq 1000$; and CV for $m = 5$, $n = 2000, 5000, 10,000$. For all other combinations of $m$ and $n$ there is no single best model selection method.

AIC seems to be the winner in this experiment: DBMS can lose up to 17.7% ($m = 10$, $n = 500$) and CV up to 3.5% ($m = 5$, $n = 20$), compared to AIC, whereas AIC loses up to 0.3% compared to DBMS ($m = 5$, $n = 5000$) and 0.6% compared to CV ($m = 5$, $n = 5000$).

In experiment C, the multinomial logit model $M^{(1)}$ is correctly specified, and outperforms model $M^{(0)}$ in all instances. Both the average reward of $M^{(0)}$ and $M^{(1)}$ converge to the optimum as $n$ grows large, but the reward of $M^{(1)}$ appears to converge faster than that of $M^{(0)}$. The average loss due to using model $M^{(0)}$ instead of model $M^{(1)}$ can be more than 10 percent ($m = 10$, $n = 2000, 5000$). DBMS always outperforms $M^{(0)}$; the relative improvement can be up to 4.8% ($m = 10$, $n = 2000$).

AIC is the clear winner in this experiment: for all pairs $m$, $n$ except $m = 10$, $n = 10$, the average reward under AIC is larger than or equal to the average reward under AIC or CV. The reason is that AIC almost always selects model $M^{(1)}$, which outperforms $M^{(0)}$ in all instances: for $m = 3$, AIC has average reward within 0.2% of that of $M^{(1)}$, and for $m = 5, 10$, AIC has exactly the same average reward as $M^{(1)}$.

An overall conclusion from these three experiments is that there is no clear winner between DBMS, AIC, and CV. In the most general case considered in experiment A, both DBMS and CV have a somewhat similar performance, with sometimes one outperforming the other and sometimes the other way around. Both outperform AIC by a (sometimes) large margin. In experiments B and C, where the multinomial logit model is correctly specified or 'almost' correctly specified, AIC performs better than DBMS and CV.

Regarding DBMS, we observe that the largest relative amount that DBMS loses compared to $M^{(0)}$ (i.e. 2.6 percent in experiment A and 0.2 percent in experiment B) is a magnitude smaller than the largest relative amount that DBMS can improve upon $M^{(0)}$ (i.e. more than 8 percent in experiment B, and 4.8 percent in experiment C).

Regarding AIC, we observe a 'sudden' jump in the average reward in experiment A if $m = 5$ and $n$ is between 200 and 500, and even more pronounced if $m = 10$ and $n$ is between 10,000 and 20,000. AIC almost behaves like an indicator function: if $n$ is smaller than a certain critical value it selects model $M^{(1)}$ with high probability, and if $n$ is larger than this value then it selects model $M^{(0)}$ with high probability. This behavior is illustrated in Fig. 4, where we repeat experiment A for $m = 10$ products and $n = 17,000, 17,100, 17,200, \ldots, 20,000$ (and with an increased number of simulations of 80,000 instead of 10,000 for each $n$, because of the finer granularity of $n$). The figure shows the relative frequency that DBMS, AIC, and CV select model $M^{(0)}$, together with the 'optimal' model selector that always selects the best of the two. It turns out that, for these values of $n$, model $M^{(0)}$ is preferable to $M^{(1)}$ in about 89–91% of the cases; DBMS is close but slightly underestimates this with approximately two percentage points, and CV structurally overestimates this fraction to 1.0. However, AIC hardly ever selects model $M^{(0)}$ if $n \leq 17,000$, and almost always selects model $M^{(0)}$ if $n \geq 20,000$. This sudden change may explain the poor performance of AIC in Experiment A, compared to DBMS or CV.

### 4.2. Newsvendor problem

Our second numerical illustration applies DBMS to the newsvendor problem, an archetypal optimization problem in inventory management. The problem consists of determining an order quantity that optimally balances between the costs of stock-outs ('backorder' costs) and overstocking ('holding' costs). The newsvendor problem has been studied in many variants; we consider the most basic version.

*Setting.* The decision to take is an order quantity $x$ from the nonnegative reals $\mathcal{X} = [0, \infty)$. After selecting $x$, an observation of demand $Y(x)$ is observed. The distribution of $Y(x)$ is independent of the decision $x$, and we write $Y = Y(x)$. The unknown cumulative distribution function (cdf) of $Y$ is denoted by $\theta^*$, and lies in the collection $\Theta^{(0)}$ of cdfs of nonnegative random variables with finite expectation:

$$\Theta^{(0)} = \Big\{ \text{all cdfs } \theta : (-\infty, \infty) \to [0,1] \text{ with } \lim_{y \uparrow 0} \theta(y) = 0 \text{ and } \int_0^\infty y \, \mathrm{d}\theta(y) < \infty \Big\}.$$

The expected cost function $c : \mathcal{X} \times \Theta^{(0)} \to [0, \infty)$ is defined as

$$c(x, \theta) = h \int_{y=0}^{x} (x - y) \, \mathrm{d}\theta(y) + b \int_{y=x}^{\infty} (y - x) \, \mathrm{d}\theta(y), \tag{20}$$

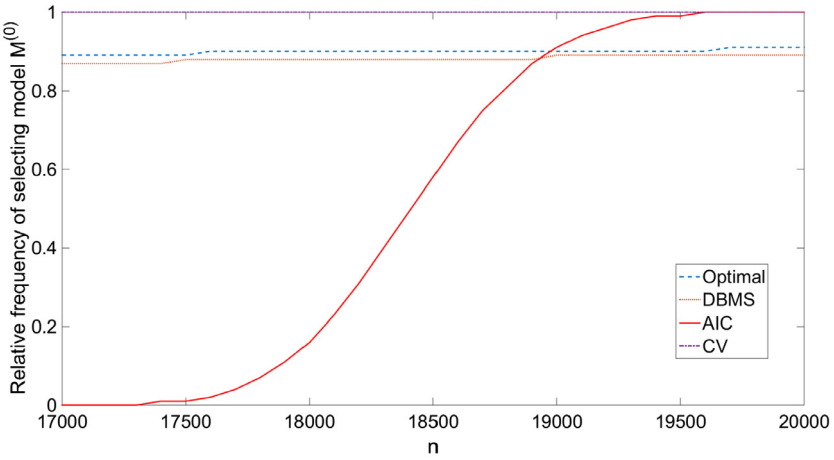


**Fig. 4.** Relative frequency of selecting model $M^{(0)}$, as function of $n$.

for some known holding costs $h > 0$ and backorder costs $b > 0$. Note that this problem is about costs minimization instead of reward maximization. We can still use DBMS by applying it to the reward function $r(x, \theta) := -c(x, \theta)$.

The demand distribution is estimated by the empirical distribution function: $\tau^{(0)}$ maps a data sequence $(x_1, y_1, \ldots, x_n, y_n)$ to the distribution function

$$y \mapsto \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{y_i \le y\}. \tag{21}$$

Optimization is exact: for each $\theta \in \Theta^{(0)}$, (20) is minimized by

$$\chi^{(0)}(\theta) := \inf\left\{z \ge 0 : \theta(z) \ge \frac{b}{b+h}\right\}. \tag{22}$$

If $\theta$ has an inverse $\theta^{-1}$ then $\chi^{(0)}(\theta) = \theta^{-1}(b/(b+h))$.

The simplified model $M^{(1)}$ assumes that demand $Y(x)$ is exponentially distributed with mean $\theta$, for all $x \in \mathcal{X}$ and some $\theta \in \Theta^{(1)} := [0, \infty)$. The estimator $\tau^{(1)}$ maps data $(x_1, y_1, \ldots, x_n, y_n)$ to the sample mean $(y_1 + \ldots + y_n)/n$, and optimization is again exact; for exponential distributions with mean $\theta$, Eq. (22) equals $\chi^{(1)}(\theta) = -\theta \log(h/(b+h))$.

*Numerical experiments.* Fix $h = 1$. For each backorder costs $b \in \{2.0, 1.5, 1.0, 0.5\}$ and each size of the initial data set $n \in \{10, 50, 100, 500, 1000\}$ we run 10,000 simulations. In each simulation we run three experiments:

- in experiment A, we let $Y$ be lognormally distributed with mean $m$ and variance $v$, where we draw $m$ uniformly at random from [0,5] and $v$ uniformly at random from [0, 25];
- in experiment B, we let $Y$ be lognormally distributed with mean $m$ and variance $m^2$, where we draw $m$ uniformly at random from [0,5];
- in experiment C, we let $Y$ be exponentially distributed with mean $m$, where we draw $m$ uniformly at random from [0,5].

In experiment A the exponential-demand model $M^{(1)}$ is almost surely misspecified, whereas in experiment C it is always correctly specified. Experiment B is somewhat in between: the exponential-demand model is misspecified, but our requirement $\mathrm{Var}(Y) = \mathbb{E}[Y]^2$ on $Y$ is satisfied by exponentially distributed demand. Thus, in experiment B, the distribution of $Y$ is, in some sense, closer to an exponentially distributed random variable than in experiment A, which might imply that model $M^{(1)}$ sometimes outperforms the true model for sufficiently small $n$.

The decisions $x_1, \ldots, x_n$ in the initial data set are drawn uniformly at random from the interval [0,5]. Note that these quantities are only needed to apply cross-validation, and are not used by DBMS.

For each experiment we determine the optimal costs under full information (Opt), under model $M^{(0)}$, model $M^{(1)}$, and under DBMS. We also test 5-fold cross-validation (CV), which chooses the model that minimizes CV($k$), $k \in \{0, 1\}$, where ties are decided in favor of model $M^{(0)}$, and where

$$\mathrm{CV}(k) := \frac{1}{5}\sum_{l=1}^{5}\sqrt{\sum_{i \in d_l}\left(\hat{c}^{(k)}(x_i, d_0 \backslash d_l) - c_i\right)^2}; \tag{23}$$

here

$$\hat{c}^{(0)}(x; \tilde{d}) := \frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}} h(x - \tilde{y}_i)^+ + b(\tilde{y}_i - x)^+,$$

$$\begin{aligned}\hat{c}^{(1)}(x; \tilde{d}) &:= h\int_{y=0}^{x}(x-y)e^{-y/\bar{y}}/\bar{y}\,\mathrm{d}y + b\int_{y=x}^{\infty}(y-x)e^{-y/\bar{y}}/\bar{y}\,\mathrm{d}y \\ &= h(x + (\exp(-x/\bar{y}) - 1)\bar{y}) + b\exp(-x/\bar{y})\bar{y},\end{aligned}$$

are the estimated cost functions under model $M^{(0)}$ and $M^{(1)}$ respectively, based on a data set $\tilde{d} = (\tilde{x}_i, \tilde{y}_i)_{1 \le i \le \tilde{n}}$; $\bar{y} = \frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}} y_i$, $c_i = h(x_i - y_i)^+ + b(y_i - x_i)^+$ are the observed costs associated to $(x_i, y_i)$, for $i = 1, \ldots, n$, and $d_l := \{(x_i, y_i) \mid i = 1 + (l-1)n/5, \ldots, ln/5\}$, $l = 1, \ldots, 5$, is a decomposition of the initial $d_0$ in five mutually disjoint sets of equal size. We omit comparing DBMS to AIC, since the true model is infinite-dimensional.

The average costs in the simulations are reported in Tables 4–6. All standard errors are smaller than 0.02.

*Outcomes.* In experiment A, model $M^{(0)}$ outperforms model $M^{(1)}$ in all instances of $b$ and $n$. The average costs under model $M^{(1)}$ can be more than 10 percent higher than that of $M^{(0)}$ ($b = 0.5$, $n \ge 50$). The average costs under DBMS are close to that of model $M^{(0)}$: never more than 0.6 percent higher ($b = 1.0$, $n = 10$), and for $n \ge 50$ the difference is never more than 0.3% ($b = 1.5$, $n = 50$). Cross-validation performs worse than DBMS in all instances of $b$ and $n$. It loses up to 3.9% compared to $M^{(0)}$ ($b = 1.0$, $n = 10$), and up to 2.0% if we only consider $n \ge 50$ ($b = 1.5$, $n = 50$).

In experiment B, model $M^{(1)}$ sometimes outperforms model $M^{(0)}$. The average costs under model $M^{(0)}$ can be more than 1.2% higher than under $M^{(1)}$ ($b = 0.5$, $n = 10$), and conversely, the average costs under model $M^{(1)}$ can be more than 1.5 percent higher than under $M^{(0)}$ ($b = 0.5$, $n = 1000$). DBMS is always within 0.7% of the best performing model ($b = 0.5$, $n = 10$), CV is always within 0.9 percent ($b = 0.5$, $n = 50$). There is no clear winner between DBMS and CV: both sometimes outperform the other, but never by a large margin. The costs under DBMS can be up to 0.2% higher

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

**Table 4**
Average costs in experiment A. $Y\sim$ lognormal$(m, v)$.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
|---|---|---|---|---|---|
| 10 | 2.42 | 2.59 | 2.66 | 2.59 | 2.67 |
| 50 | 2.42 | 2.46 | 2.52 | 2.46 | 2.51 |
| 100 | 2.42 | 2.44 | 2.51 | 2.45 | 2.48 |
| 500 | 2.42 | 2.43 | 2.49 | 2.43 | 2.44 |
| 1000 | 2.42 | 2.43 | 2.49 | 2.43 | 2.43 |
| | | (a) $b = 2.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.98 | 2.09 | 2.17 | 2.10 | 2.16 |
| 50 | 1.98 | 2.00 | 2.07 | 2.01 | 2.04 |
| 100 | 1.98 | 1.99 | 2.05 | 1.99 | 2.02 |
| 500 | 1.98 | 1.98 | 2.04 | 1.98 | 1.98 |
| 1000 | 1.98 | 1.98 | 2.04 | 1.98 | 1.98 |
| | | (b) $b = 1.5$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.46 | 1.53 | 1.62 | 1.54 | 1.59 |
| 50 | 1.46 | 1.47 | 1.56 | 1.48 | 1.50 |
| 100 | 1.46 | 1.46 | 1.55 | 1.47 | 1.48 |
| 500 | 1.46 | 1.46 | 1.54 | 1.46 | 1.46 |
| 1000 | 1.46 | 1.46 | 1.54 | 1.46 | 1.46 |
| | | (c) $b = 1.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 0.83 | 0.88 | 0.96 | 0.89 | 0.90 |
| 50 | 0.83 | 0.84 | 0.93 | 0.84 | 0.85 |
| 100 | 0.83 | 0.84 | 0.93 | 0.84 | 0.84 |
| 500 | 0.83 | 0.83 | 0.92 | 0.83 | 0.83 |
| 1000 | 0.83 | 0.83 | 0.92 | 0.83 | 0.83 |
| | | (d) $b = 0.5$ | | | |

**Table 5**
Average costs in experiment B. $Y\sim$ lognormal$(m, m^2)$.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
|---|---|---|---|---|---|
| 10 | 2.42 | 2.60 | 2.58 | 2.59 | 2.59 |
| 50 | 2.42 | 2.46 | 2.46 | 2.46 | 2.46 |
| 100 | 2.42 | 2.44 | 2.45 | 2.44 | 2.44 |
| 500 | 2.42 | 2.42 | 2.43 | 2.42 | 2.42 |
| 1000 | 2.42 | 2.42 | 2.43 | 2.42 | 2.42 |
| | | (a) $b = 2.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.99 | 2.12 | 2.11 | 2.11 | 2.11 |
| 50 | 1.99 | 2.02 | 2.02 | 2.02 | 2.02 |
| 100 | 1.99 | 2.00 | 2.01 | 2.01 | 2.00 |
| 500 | 1.99 | 1.99 | 2.00 | 1.99 | 1.99 |
| 1000 | 1.99 | 1.99 | 2.00 | 1.99 | 1.99 |
| | | (b) $b = 1.5$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.49 | 1.57 | 1.56 | 1.57 | 1.57 |
| 50 | 1.49 | 1.50 | 1.50 | 1.50 | 1.50 |
| 100 | 1.49 | 1.50 | 1.49 | 1.50 | 1.50 |
| 500 | 1.49 | 1.49 | 1.49 | 1.49 | 1.49 |
| 1000 | 1.49 | 1.49 | 1.49 | 1.49 | 1.49 |
| | | (c) $b = 1.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 0.86 | 0.92 | 0.90 | 0.91 | 0.91 |
| 50 | 0.86 | 0.87 | 0.88 | 0.87 | 0.88 |
| 100 | 0.86 | 0.87 | 0.88 | 0.87 | 0.87 |
| 500 | 0.86 | 0.86 | 0.88 | 0.86 | 0.86 |
| 1000 | 0.86 | 0.86 | 0.88 | 0.86 | 0.86 |
| | | (d) $b = 0.5$ | | | |

than under CV ($b = 1.5$, $n = 10$), and the costs under CV can be up to 0.7% higher than under DBMS ($b = 0.5$, $n = 50$).

In experiment C, model $M^{(1)}$ outperforms model $M^{(0)}$ in all instances of $b$ and $n$; the average costs under $M^{(0)}$ can be up to 4.7% ($b = 0.5$, $n = 10$) higher than under $M^{(1)}$. DBMS improves upon $M^{(0)}$ in all instances, by up to 1.4% ($b = 0.5$, $n = 10$). CV does better than DBMS in all instances, and can improve upon $M^{(0)}$ by up to 3.6% ($b = 0.5$, $n = 10$).

An overall conclusion from these three experiments is that sticking to a single model $M^{(0)}$ or $M^{(1)}$ may induce losses up to 10 percent. The costs under DBMS stay close to that of $M^{(0)}$ in case

**Table 6**
Average costs in experiment C. $Y\sim$ exponential with mean $m$.

| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
|---|---|---|---|---|---|
| 10 | 2.75 | 2.97 | 2.89 | 2.94 | 2.92 |
| 50 | 2.75 | 2.80 | 2.78 | 2.79 | 2.78 |
| 100 | 2.75 | 2.77 | 2.76 | 2.77 | 2.76 |
| 500 | 2.75 | 2.75 | 2.75 | 2.75 | 2.75 |
| 1000 | 2.75 | 2.75 | 2.75 | 2.75 | 2.75 |
| | | (a) $b = 2.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 2.29 | 2.45 | 2.39 | 2.43 | 2.41 |
| 50 | 2.29 | 2.33 | 2.31 | 2.32 | 2.32 |
| 100 | 2.29 | 2.31 | 2.30 | 2.31 | 2.30 |
| 500 | 2.29 | 2.29 | 2.29 | 2.29 | 2.29 |
| 1000 | 2.29 | 2.29 | 2.29 | 2.29 | 2.29 |
| | | (b) $b = 1.5$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.73 | 1.84 | 1.79 | 1.83 | 1.81 |
| 50 | 1.73 | 1.76 | 1.75 | 1.75 | 1.75 |
| 100 | 1.73 | 1.75 | 1.74 | 1.74 | 1.74 |
| 500 | 1.73 | 1.74 | 1.73 | 1.74 | 1.73 |
| 1000 | 1.73 | 1.74 | 1.73 | 1.73 | 1.73 |
| | | (c) $b = 1.0$ | | | |
| $n$ | Opt | $M^{(0)}$ | $M^{(1)}$ | DBMS | CV |
| 10 | 1.01 | 1.08 | 1.03 | 1.07 | 1.04 |
| 50 | 1.01 | 1.03 | 1.02 | 1.02 | 1.02 |
| 100 | 1.01 | 1.02 | 1.02 | 1.02 | 1.02 |
| 500 | 1.01 | 1.02 | 1.01 | 1.01 | 1.01 |
| 1000 | 1.01 | 1.01 | 1.01 | 1.01 | 1.01 |
| | | (d) $b = 0.5$ | | | |

$M^{(1)}$ is misspecified, but if $M^{(1)}$ is better than $M^{(0)}$, then part of the potential gain is captured by DBMS. Cross-validation performs worse than DBMS in experiment A, comparable to DBMS in experiment B, and better than DBMS in experiment C. The largest observed gain of CV compared to $M^{(0)}$ is 3.6%, and the largest loss 3.9%. For DBMS these values are 1.4% and 0.6%. Thus, in some sense, DBMS is closer to $M^{(0)}$ and CV is closer to $M^{(1)}$: both the highest gains and the largest losses of DBMS compared to the true model are smaller than the gains and losses of CV compared to $M^{(0)}$.

## 5. Concluding remarks

Data-driven decision making revolves around mathematical models, statistical estimators, and optimization algorithms. While the properties of estimators and optimization algorithms have been studied extensively in a wide variety of contexts, the question how to select a proper mathematical model from a decision-making viewpoint has received little attention in the literature. In many situations, for example, there is a choice between a simple model and more complex model. Determining which of these models leads to the best decision is a very relevant question, but the existing literature does not describe a generic method to answer it. An extensive model-selection apparatus has been developed in the past decades, but these methods either do not take quality-of-decisions as the discriminating factor between models (and thus, in a sense, decouple model selection from optimization), or can only be applied to a subset of the class of decision-problems that we consider.

This paper aims to take a step in the direction of connecting model-selection with data-driven decision making. To this end, we propose a generic decision-based model selection method, named DBMS, that judges the quality of a model by the (estimated) quality of the decision it supports. The method is applicable to a wide class of decision-problems. It is easy to use in practice, does not require large computation times, and does not depend on hyper-parameters that are difficult to tune. Under some conditions, the method is reward-consistent (meaning that the reward using DBMS converges to the optimal reward). Our numerical illustrations show

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025

that DBMS is frequently on par and sometimes better than existing model-selection methods; this suggests that DBMS is a step in the right direction, but that there also still is room for further improvement and fine-tuning of the method.

The main practical insight of this work is that decision makers who have to select a model for a data-driven decision problem should not confine themselves to a single model; instead, they can select multiple models with different degrees of complexity, and use a decision-based model-selection method such as the one proposed in this paper to determine, for each data set at hand, which model is expected to produce the best decision.

Whilst the focus of this paper is on static problems, we expect that decision-based model selection can be a powerful tool in dynamic decision problems under uncertainty (so-called *multi-armed bandit problems*). In the majority of these problems, the model is fixed throughout the whole time horizon. As an alternative, we suggest to incorporate decision-based model selection into the multi-armed bandit framework, such that the complexity of the model upon which decisions are based grows with the size and richness of the data that is available. In other words, the complexity of the used model should be 'justified' by what the data can support, and when the data set is growing, the complexity of the model should be growing as well. Integrating decision-based model selection method in such dynamic decision making problems may lead to significant improved performance in a wide variety of contexts. In several of such applications, dynamic model selection can only be implemented if its computation times are sufficiently small. Because the method proposed in this paper scores well on this aspect - compared, for example, to cross-validation - it may lend itself very well for such dynamic decision-making applications.

**Acknowledgements**

Fruitful discussions and conversations with Thomas Vossen, Bill Cooper, Peter Grünwald, and Wouter Koolen have been very beneficial for the development of ideas presented in this paper. We thank Peter Grünwald for his valuable comments and literature suggestions, which have greatly improved the paper. We thank Gah-Yi Ban for reading and commenting on an earlier version the manuscript. We also gratefully thank the referee team for their constructive comments. The first author was partially supported by an NWO VENI grant, and part of his research was done while affiliated with University of Twente and Centrum Wiskunde & Informatica, Amsterdam. The second author was with Centrum Wiskunde & Informatica, Amsterdam and VU University, Amsterdam, during the research leading to this paper.

## Appendix A. loss incurred by suboptimal model selection criteria

In this supplementary section we show by an example that model selection based on assessing the quality of the estimated parameters or of the estimated reward function, instead of the quality of decisions, may induce unbounded losses. To this end, consider the linear program

$$\max r(x,\theta) := \max (\theta-\alpha)x \quad \text{s.t. } 0 \le x \le 1,$$

for $\alpha \ge 0$, $\theta \in \mathbb{R}$, which is maximized by $x^* := \mathbf{1}\{\theta-\alpha>0\}$. The value of $\alpha$ is known, the value of $\theta$ is unknown, but it can be estimated from data of the form $(x_i, y_i)_{1\le i\le n}$, where $n$ is an integer, $x_1,\dots,x_n \in [0,1]$ are nonrandom and not all zero, $y_i = \theta x_i + \epsilon_i$ for $i=1,\dots,n$, and $\epsilon_1,\dots,\epsilon_n$ are i.i.d. normally distributed random variables with mean zero and (unknown) variance $\sigma^2>0$. The ordinary least squares estimates of $\theta$ and $\sigma^2$ are $\hat\theta_0 := (\sum_{i=1}^n x_i^2)^{-1}\sum_{i=1}^n x_i y_i$ and $\hat\sigma_0^2 := n^{-1}\sum_{i=1}^n (y_i-\hat\theta_0 x_i)^2$, the corresponding estimated objective function is $\hat r_0 : x \mapsto (\hat\theta_0-\alpha)x$, and the corresponding estimated optimal decision is $\hat x_0^* := \mathbf{1}\{\hat\theta_0-\alpha>0\}$. These quantities correspond to what we call the 'true' model. We also consider a simplified model, where the decision maker assumes $\theta=0$. In this case, she estimates $\theta$ and $\sigma^2$ by $\hat\theta_1 := 0$ and $\hat\sigma_1^2 := n^{-1}\sum_{i=1}^n y_i^2$, the objective function $r$ by $\hat r_1 : x \mapsto (0-\alpha)x$, and the corresponding optimal decision by $\hat x_1^* := 0$, the maximizer of $\hat r_1$.

Now, consider the following three model-selection criteria:

(i) the quality of the estimated optimal decision $\hat x_k^*$, measured by

$$\text{Regret}(\hat x_k^*) := r(x^*) - r(\hat x_k^*), \qquad k \in \{0,1\};$$

(ii) the quality of the estimated reward function $\hat r_k$, measured by its $L_2$ distance to the true reward function:

$$\big|\big|\hat r_k - r\big|\big|_2 := \left(\int_0^1 (\hat r_k(x)-r(x))^2 \mathrm{d}x\right)^{1/2}, \qquad k \in \{0,1\};$$

(iii) the quality of the estimated parameters $\hat\theta_k$, $\hat\sigma_k^2$, measured by the expected KL-divergence between the true and estimated distribution of $y$ at a randomly selected $x$:

$$\mathrm{KL}(\hat\theta_k,\hat\sigma_k^2) := \int_0^1\int_{-\infty}^{\infty} \phi(y \mid \theta x,\sigma^2)\log\left(\frac{\phi(y\mid\theta x,\sigma^2)}{\phi(y\mid\hat\theta_k x,\hat\sigma_k^2)}\right)\mathrm{d}y\,\mathrm{d}x, \quad k\in\{0,1\};$$

here $\phi(y|\mu,\varsigma^2)$ is the pdf of a $N(\mu,\varsigma^2)$ distributed random variable, evaluated at $y$.

Let

$$k_{\text{Regret}} := \arg\min_{k\in\{0,1\}} \text{Regret}(\hat x_k^*),$$
$$k_{L_2} := \arg\min_{k\in\{0,1\}} \big|\big|\hat r_k - r\big|\big|_2,$$
$$k_{\text{KL}} := \arg\min_{k\in\{0,1\}} \mathrm{KL}(\hat\theta_k,\hat\sigma_k^2),$$

be the best models according to these three criteria, with ties decided in favor of model $k=0$. The expected regret under these criteria is given by

$$\mathbb{E}[\text{Regret}(x_{k_{\text{Regret}}})] = \begin{cases} (\theta-\alpha)\mathbb{P}\Big(\hat\theta_0 \le \alpha\Big) & \text{if } \theta>\alpha \\ 0 & \text{if } \theta\le\alpha \end{cases} \tag{24}$$

$$\mathbb{E}[\text{Regret}(x_{k_{L_2}})] = \begin{cases} (\theta-\alpha)\Big(\mathbb{P}\Big(\hat\theta_0\le\alpha\Big)+\mathbb{P}\Big(\hat\theta_0>\alpha \text{ and } (\hat\theta_0-\theta)^2>\theta^2\Big)\Big) & \text{if } \theta>\alpha \\ (\alpha-\theta)\mathbb{P}\Big(\hat\theta_0>\alpha \text{ and } (\hat\theta_0-\theta)^2\le\theta^2\Big) & \text{if } \theta\le\alpha \end{cases} \tag{25}$$

and

$$\mathbb{E}[\text{Regret}(x_{k_{\text{KL}}})] = \begin{cases} (\theta-\alpha)\Big(\mathbb{P}\Big(\hat\theta_0\le\alpha\Big)+\mathbb{P}\Big(\hat\theta_0>\alpha \text{ and } \mathrm{KL}(\hat\theta_0,\hat\sigma_0^2)>\mathrm{KL}(\hat\theta_1,\hat\sigma_1^2)\Big)\Big) & \text{if } \theta>\alpha \\ (\alpha-\theta)\mathbb{P}\Big(\hat\theta_0>\alpha \text{ and } \mathrm{KL}(\hat\theta_0,\hat\sigma_0^2)\le\mathrm{KL}(\hat\theta_1,\hat\sigma_1^2)\Big) & \text{if } \theta\le\alpha \end{cases} \tag{26}$$

If $\sigma^2/\sum_{i=1}^n x_i^2 = \theta^2$ and $\alpha = \theta + 0.75\sqrt{\theta}$, for some $\theta>0$, then $\mathbb{E}[\text{Regret}(x_{k_{L_2}})]$ can be made arbitrary large by choosing $\theta$ large, while $\mathbb{E}[\text{Regret}(x_{k_{\text{Regret}}})]$ remains zero.

If $\sum_{i=1}^n x_i^2 = c_1 n$, $\theta = \sigma^2 = c_2^2 c_1 n$, and $\alpha = \theta + 0.75c_2$, for some $0<c_1<1<c_2$, then, as $n$ grows large, $\mathbb{P}\Big(\mathrm{KL}(\hat\theta_0,\hat\sigma_0^2)\le\mathrm{KL}(\hat\theta_1,\hat\sigma_1^2)\Big)$ converges to one, $\mathbb{E}[\text{Regret}(x_{k_{\text{KL}}})]$ converges to $0.75c_2\mathbb{P}(N(0,1)>0.75)\approx 0.17c_2$, whereas

$\mathbb{E}[\text{Regret}(x_{k_{\text{Regret}}})] = 0$ for all $n$. The difference in expected regret can be arbitrarily large by choosing $c_2$ large.

Note that the model-selection criteria in this example depend on the unknown parameters. They still need to be estimated from data before they can be applied. Cross-validation and AIC are often used to estimate $k_{L_2}$ and $k_{\text{KL}}$, and DBMS is an estimator of $k_{\text{Regret}}$. The purpose of this example is not to compare the performance of DBMS with AIC or CV, but to argue that, in a decision-making context, it makes sense to design model-selection procedures that estimate $k_{\text{Regret}}$, i.e. the quality of decisions, instead of $k_{L_2}$, $k_{\text{KL}}$, or other criteria not connected to the decision problem at hand.

## References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest: Academiai Kiado.
Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4*, 40–79.
Bastani, H., & Bayati, M. (2019). Online Decision Making with High-Dimensional Covariates. *Operations Research, 68*(1). https://doi.org/10.1287/opre.2019.1902.
Besbes, O., Phillips, R., & Zeevi, A. (2010). Testing the validity of a demand model: an operations perspective. *Manufacturing & Service Operations Management, 12*(1), 162–183.
Besbes, O., & Zeevi, A. (2015). On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science, 61*(4), 723–739.
Bousquet, O., Boucheron, S., & Lugosi, G. (2004). Introduction to statistical learning theory. In O. Bousquet, U. von Luxburg, & G. Rätsch (Eds.), *Advanced lectures on machine learning*. In *Lecture Notes in Computer Science: 3176* (pp. 169–207). Springer Berlin Heidelberg.
Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer-Verlag.
Cachon, G. P., & Kök, A. G. (2007). Implementation of the newsvendor model with clearance pricing: How to (and how not to) estimate a salvage value. *Manufacturing & Service Operations Management, 9*(3), 276–290.
Chu, L. Y., Shanthikumar, J. G., & Shen, Z.-J. M. (2008). Solving operational statistics via a Bayesian analysis. *Operations Research Letters, 36*(1), 110–116.
Claeskens, G., & Hjort, N. L. (2003). The focused information criterion. *Journal of the American Statistical Association, 98*(464), 900–916.
Claeskens, G., & Hjort, N. L. (2008). *Model selection and model averaging*. Cambridge: Cambridge University Press.
Cooper, W. L., & Li, L. (2012). On the use of buy up as a model of customer choice in revenue management. *Production and Operations Management, 21*(5), 833–850.
Cooper, W. L., Homem-de Mello, T., & Kleywegt, A. J. (2006). Models of the spiral–down effect in revenue management. *Operations Research, 54*(5), 968–987.
Cooper, W. L., Homem-de Mello, T., & Kleywegt, A. J. (2015). Learning and pricing with models that do not explicitly incorporate competition. *Operations Research, 63*(1), 86–103.
Feldman, J., Zhang, D., Liu, X., & Zhang, L. (2019). Customer choice models versus machine learning: finding optimal product displays on Alibaba. Available at SSRN. https://ssrn.com/abstract=3232059.
Gallego, G., Ratliff, R., & Shebalov, S. (2015). A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research, 63*(1), 212–232.
Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association, 70*(350), 320–328.
Grünwald, P. D. (2007). *The minimum description length principle*. Cambridge, Massachusetts: MIT Press.
Grünwald, P. D., & van Ommen, T. (2017). Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis, 12*(4), 1069–1103.
Guyon, I., Saffari, A., Dror, G., & Cawley, G. (2010). Model selection: beyond the Bayesian/frequentist divide. *Journal of Machine Learning Research, 11*, 61–87.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. New York: Springer.
Jeffreys, H. (1935). Some tests of significance, treated by the theory of probability. *Mathematical Proceedings of the Cambridge Philosophical Society, 31*(02), 203–222.
Jeffreys, H. (1961). *Theory of probability*. New York: Oxford University Press.
Kao, Y. H., & Van Roy, B. (2013). Learning a factor model via regularized PCA. *Machine Learning, 91*(3), 279–303.
Kao, Y. H., & Van Roy, B. (2014). Directed principal component analysis. *Operations Research, 62*(4), 957–972.
Kao, Y. H., Van Roy, B., & Yan, X. (2009). Directed regression. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems 22* (pp. 889–897). Curran Associates, Inc.
Karoui N.El., Lim A.E.B., Vahn G.-Y. (2014). Performance-based regularization in mean-CVaR portfolio optimization. Working paper, arXiv1111.2091.
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.
(2001). Model selection. In P. Lahiri (Ed.). *Lecture Notes - Monograph Series, Volume 38*. Beachwood, Ohio: Institute of Mathematical Statistics.
Lee, S., Homem-de Mello, T., & Kleywegt, A. J. (2012). Newsvendor-type models with decision-dependent uncertainty. *Mathematical Methods of Operations Research, 76*(2), 189–221.
Lim, A. E. B., Shanthikumar, J. G., & Shen, Z. J. M. (2006). Model uncertainty, robust optimization, and learning. In M. P. Johnson, B. Norman, N. Secomandi, P. Gray, & H. J. Greenberg (Eds.), *Models, methods, and applications for innovative decision making* (pp. 66–94). INFORMS.
Liyanage, L. H., & Shanthikumar, J. G. (2005). A practical inventory control policy using operational statistics. *Operations Research Letters, 33*(4), 341–348.
Mallows, C. L. (1973). Some comments on Cp. *Technometrics, 15*(4), 661–675.
Nickerson, R. C., & Boyd, D. W. (1980). The use and value of models in decision analysis. *Operations Research, 28*(1), 139–155.
Ramamurthy, V., Shanthikumar, J. G., & Shen, Z.-J. M. (2012). Inventory policy with parametric demand: operational statistics, linear correction, and regression. *Production and Operations Management, 21*(2), 291–308.
Rissanen, J. (1978). Modeling by shortest data description. *Automatica, 14*(5), 465–471.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583–639.
Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological), 36*(2), 111–147.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological), 58*(1), 267–288.
Vapnik, V. N. (1998). *Statistical learning theory*. New York: Wiley-Interscience.
Vapnik, V. N. (2000). *The nature of statistical learning theory*. New York: Springer.
Wasserman, L. (2000). Bayesian model selection and model averaging.. *Journal of Mathematical Psychology, 44*(1), 92–107.
Watson, J., & Holmes, C. (2016). Approximate models and robust decisions. *Statistical Science*. Forthcoming
Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology, 44*(1), 41–61.

Please cite this article as: A.V. den Boer and D.D. Sierag, Decision-based model selection, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.08.025