# Tight Upper Bounds for the Expected Loss of Lexicographic Heuristics in Binary Multi-attribute Choice

Juan A. Carrasco,[*] Manel Baucells[†]

## Abstract

Tight upper bounds for the expected loss of the *DEBA* (Deterministic-Elimination-By-Aspects) lexicographic selection heuristic are obtained for the case of an additive separable utility function with unknown non-negative, non-increasing attribute weights for numbers of alternatives and attributes as large as 10 under two probabilistic models: one in which attributes are assumed to be independent Bernouilli random variables and another one with positive inter-attribute correlation. The upper bounds improve substantially previous bounds and extend significantly the cases in which a good performance of *DEBA* can be guaranteed under the assumed cognitive limitations.

**Keywords:** Multiattribute decision making; Binary attributes; Lexicographic heuristics; DEBA; Performance bounds

*JEL Classification:* C44, C63.

---

[*]Departament d'Enginyeria Electrònica, Universitat Politècnica de Catalunya, Diagonal 647, plta. 9, 08028 Barcelona, Spain, carrasco@eel.upc.edu

[†]IESE Business School, Universidad de Navarra, Avda. Pearson 21, 08034 Barcelona, Spain, mbaucells@iese.edu

# 1 Introduction

The goal of this paper is the formal study of the performance of lexicographic heuristics in multi-attribute decision making. Lexicographic heuristics are psychologically appealing because in complex multi-attribute choices provide a decision rule that avoids explicit trade-offs. We have in mind a standard multi-attribute decision problem (Keeney and Raiffa 1993), with an additive, separable utility function and $m$ alternatives characterized by $k$ attributes $x_{i,r}$, $1 \leq i \leq m$, $1 \leq r \leq k$. The utility of the alternative $i$, characterized by the profile $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,k})$, is defined as

$$U_i = w_1 x_{i,1} + w_2 x_{i,2} + \cdots + w_k x_{i,k},$$

where the $w_r$ are non-negative weights subject to the constraint $w_1 + w_2 + \cdots + w_k = 1$. Cognitive limitations are introduced by assuming that the decision maker can order the weights by size, but that the exact values of the weights are unknown. Therefore, without loss of generality, we assume $w_1 \geq w_2 \geq \cdots \geq w_k \geq 0$. The problem is to identify which of the $m$ alternatives is "best" (has the largest value of $U_i$) under this cognitive limitation. To resolve this problem, we consider a decision maker that follows a lexicographic heuristic that relies on the ordering, but not on the magnitude, of the weights.

The specific lexicographic heuristic that we will examine is *DEBA*, the *deterministic* version of the *elimination-by-aspects* model proposed by Tversky (1972). *DEBA* considers the attributes in decreasing weight ordering. In the first step, *DEBA* eliminates all the alternatives without the maximum values in the first attribute. If a single alternative remains, it is chosen. Otherwise, the values of the second attribute are examined. This procedure continues until only one alternative remains or all attributes have been examined. If, after examining the last attribute, two or more alternatives remain, then the choice between them is made at random. As a procedure, *DEBA* generalizes —to more than two alternatives— the lexicographic binary-choice model Take-The-Best (*TTB*) proposed by Gigerenzer and Goldstein (1996). There is a small difference, however: in *TTB*, the attributes are ordered by their validities, which are computed using a database of previous instances of alternatives, while in *DEBA* the ordering of the attributes by decreasing weights is assumed known. Abstracting in *TTB* the first step, *TTB* can be truly considered a particular instance of *DEBA* with just two alternatives.

In this paper, we make the additional assumption that the $x_{i,r}$ are binary, i.e., their value is either 0 or 1. This will be always the case for attributes features that are either present or absent, or that take two values. Besides mathematical tractability, the binary setup has another important advantage. In a binary setup, *DEBA* agrees with other lexicographic heuristics that differ on the cut-off used to encode the attribute values as high ($x_{i,r} = 1$) or low ($x_{i,r} = 0$). This binary encoding phase, together with the rule to order the attributes, precedes the use of the common lexicographic selection rule. For example, one could assign a value 1 only to those attribute values with the best level on that attribute; or use instead a low cutoff representing a minimum acceptable level. Those two choices yield, respectively, the *EBA* and the *LEX* heuristics discussed by Payne et al. (1993). The binary encoding is also a way to incorporate the bounded rationality of decision-makers that

easily distinguish between zero (absence) and non-zero (presence) values, but are quite insensitive to the actual magnitude of the attributes (Hsee and Rottenstreich 2004).

The *DEBA* heuristic is easy to use and popular (Bröder 2000; Newell and Shanks 2003; Newell et al. 2003). In many situations, for example, there is no need to look beyond the first one or two attributes to make a decision. Several studies have shown *DEBA* to be effective in relation to alternative simple decision heuristics (Gigerenzer and Goldstein 1996; Czerlinski et al. 1999; Martignon and Hoffrage 1999, 2002) as well as having desirable properties for both binary and multivariate choice (Hogarth and Karelaia 2006; Katsikopoulos and Fasolo, 2006). Even when attributes are continuous variables, the model can be quite effective under some circumstances (Gigerenzer et al. 1999; Hogarth and Karelaia 2005). Most of these studies give exact performance measures for two or three alternatives, or provide performance estimates based on simulation, or on particular data sets.

An exception is Baucells et al. (in press), henceforth BCH, where lower bounds for the probability that *DEBA* will choose a best alternative (one with the largest utility) and upper bounds for the expected loss of *DEBA* have been obtained by exploiting the concept of cumulative dominance (Kirkwood and Sarin 1985). Two binary probabilistic models were considered in BCH: one in which attributes are assumed to be independent Bernoulli random variables and another one with positive inter-attribute correlation and attributes of the same average quality. The bounds are not restricted to *DEBA*: they apply, respectively, to the so-called cumulative dominance compliant heuristics and fully cumulative dominance compliant heuristics, of which *DEBA* is an example. The bounds were obtained using exact computational approaches allowing the analysis of cases with up to 10 alternatives and 10 attributes. The lower bounds for the probability that the selection heuristic will choose a best alternative allowed the identification of cases in which the selection heuristic is guaranteed to have a good performance under the assumed cognitive limitations. The upper bounds for the expected loss widened that identification.

The *expected loss*, or expected difference between the utility of a best alternative and the utility of the alternative chosen by *DEBA* is the focus of this paper. The upper bounds for the expected loss of *DEBA* obtained in BCH were derived by noting that, in the presence of cumulative dominance, *DEBA* is ensured to make the correct choice and incur zero utility loss, and, if cumulative dominance holds up to attribute $r^*$, then the loss is non-greater than the sum of weights from $r^* + 1$ to $k$, which can be shown to be (tightly) bounded by $(k - r^*)/k$. That approach assumes a most pessimistic all-0 path from attribute $r^* + 1$ on for the alternative chosen by *DEBA*, together with a most optimistic all-1 path from attribute $r^* + 1$ on for the best alternative, ignoring the statistical properties of the probabilistic models. Our goal here is to repair this shortcoming and exploit the peculiarities of *DEBA* to improve the upper bounds for the expected loss derived in BCH. We will consider the same two probabilistic models as in BCH. The approach we will follow is to improve the upper bounds for the expected loss conditioned on the last attribute index $r^*$ for which some alternative exhibits cumulative dominance. The resulting new upper bounds are tight, improve substantially the upper bounds obtained in BCH, and widen significantly the identification of cases in which *DEBA* is guaranteed to have a good performance under the assumed cognitive limitations. Such guarantee can

be exploited to disregard the use of more sophisticated and more costly decision rules, including the obtention of more accurate estimates for the weights. To calculate the bounds we use computational approaches similar to the ones used in BCH yielding exact values for cases with up to 10 alternatives and 10 attributes.

The rest of the paper is organized as follows. Section 2 presents the probabilistic models for the attributes which will be considered (they are the same as those considered in BCH), reviews the cumulative dominance concept, and summarizes the developments in BCH regarding the upper bounds for the expected loss under those probabilistic models. Sections 3 and 4 are quite technical, and derive the improved, new upper bounds for the probabilistic models without and with correlation, respectively. Section 5 includes the numerical calculation of the improved upper bounds for the expected loss of *DEBA* under our two probabilistic models, the comparison with the upper bounds obtained in BCH, and the analysis of the tightness of the new upper bounds. We also show in that section that the new upper bounds widen significantly the identification of cases in which *DEBA* is guaranteed to have a good performance. Section 6 concludes the paper and highlights future research directions. Finally, the Appendix includes some proofs. Throughout the paper we use the conventions that $\sum_{i=n}^{n'} u_i = 0$, $n' < n$ and that 0 by an undefined quantity is equal to 0.

## 2 Preliminaries

Two probabilistic models for the values of the attributes $x_{i,r}$, $1 \leq i \leq m$, $1 \leq r \leq k$ will be considered:

*ZIAC (Zero Inter-Attribute Correlation) model:* The $x_{i,r}$ are independent Bernoulli random variables with parameter $p_r$, $0 < p_r < 1$.

*PIAC (Positive Inter-Attribute Correlation) model:* The $x_{i,r}$ are obtained as $x_{i,r} = (1 - z_i)y_{i,r}^l + z_i y_{i,r}^h$, where the $z_i$, $y_{i,r}^l$, $y_{i,r}^h$ are independent Bernoulli random variables with parameters $p$, $p^l = p - \sqrt{\rho}p$, and $p^h = p + \sqrt{\rho}(1 - p)$, respectively, for some $0 < p < 1$ and some $0 \leq \rho < 1$.

The ZIAC model is a simple model without need for justification. We note that $E[x_{i,r}] = p_r$. Thus, the parameter $p_r$ can be looked at as measuring the average quality of the attribute $r$: higher values of $p_r$ model attributes of higher average quality. The PIAC model is intuitively appealing: if there is positive correlation among the attributes of a given alternative, it is because there is some common cause shifting the average quality of the attributes of a given alternative. In the PIAC model, this is captured by the alternatives belonging to a "good" population (with expected values for the attribute values equal to $p^h = p + \sqrt{\rho}(1 - p)$) with probability $p$ and to a "bad" population (with expected values for the attribute values equal to $p^l = p - \sqrt{\rho}p$) with probability $1 - p$. In the PIAC model, $E[x_{i,r}] = p$ and the attribute values of any given alternative have positive correlation $\rho$. The ZIAC model with $p_r = p$, $1 \leq r \leq k$, can be seen as a particular case of the PIAC model
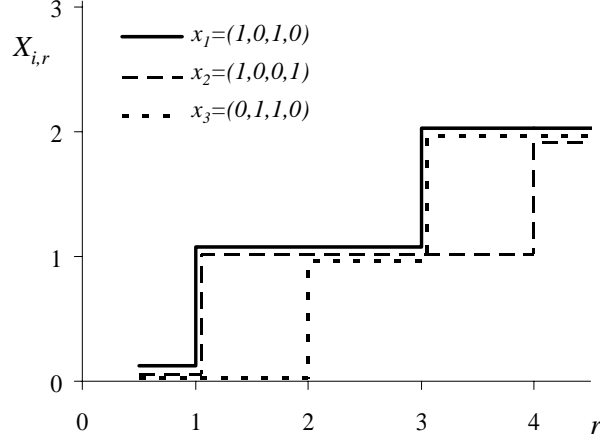
Figure 1: Alternative profiles illustrating cumulative dominance in the binary attribute case.

with $\rho = 0$. Since $\sum_{r=1}^{k} w_r = 1$, in the PIAC model the expected value of the utility of any given alternative $i$ is $E[U_i] = p$, so the parameter $p$ measures the average quality of an alternative. In the ZIAC model, $E[U_i] = \sum_{r=1}^{k} w_r p_r$, and the (unknown) weigthed sum $\sum_{r=1}^{k} w_r p_r$ is a measure of the average quality of an alternative. Of course, for $p_r = p$, $1 \leq r \leq k$, $E[U_i] = p$, and the parameter $p$ is in that case a measure of the average quality of an alternative.

In BCH, upper bounds for the expected loss of *DEBA* and similar related heuristics were obtained for the ZIAC and the PIAC models using the concept of cumulative dominance. We will start the review of those upper bounds by reviewing the concept of cumulative dominance. The cumulative profile of an alternative $i$, $1 \leq i \leq m$, is defined as $(X_{i,1}, X_{i,2}, \ldots, X_{i,k})$, where $X_{i,r} = \sum_{s=1}^{r} x_{i,s}$, $1 \leq r \leq k$. Alternative $i$ exhibits *cumulative dominance* over alternative $j$ up to attribute $r$, denoted by $c_r(i,j)$, if and only if $X_{i,s} \geq X_{j,s}$, $1 \leq s \leq r$. Alternative $i$ exhibits cumulative dominance over alternative $j$ if and only if $c_k(i,j)$, i.e. if alternative $i$ exhibits cumulative dominance over alternative $j$ up to attribute $k$. Fig. 1 illustrates cumulative dominance in the binary attribute case. Alternative 2 exhibits cumulative dominance over alternative 3 up to attribute 2, and alternative 1 exhibits cumulative dominance over alternatives 2 and 3. Further, we can use the fact that weights are non-increasing to conclude that $U_1 = w_1 + w_3 \geq U_2 = w_1 + w_4$ and $U_1 = w_1 + w_3 \geq U_3 = w_2 + w_3$. It is known that cumulative dominance characterizes optimality for non-increasing weights (Kirkwood and Sarin 1985):

**Proposition 1.** $U_i \geq U_j$ for all weights $w_1 \geq w_2 \geq \cdots \geq w_k \geq 0$, $\sum_{r=1}^{k} w_r = 1$ if and only if $c_k(i,j)$.

For $1 \leq r \leq k$, let $C_r$ denote the set of alternatives that exhibit cumulative dominance over any other alternative up to attribute $r$, i.e.,

$$C_r = \{i, 1 \leq i \leq m \, : \, c_r(i,j), 1 \leq j \leq m\}.$$

Obviously, $C_1 \supset C_2 \supset \cdots \supset C_k$. All alternatives in $C_r$ have identical cumulative attribute profile up to attribute $r$ and, therefore, they have identical attribute profile up to attribute $r$. More importantly, if $C_k$ is non-empty, then Proposition 1 guarantees that the alternatives in $C_k$ will have

4

the largest utility. In the example of Fig. 1, $C_1 = C_2 = \{1, 2\}$ and $C_3 = C_4 = \{1\}$. $C_1$ will always be non-empty. In the binary attribute case, $C_2$ will be always non-empty too. This follows by noting that $C_2$ can only be empty if there exist two alternatives $i, j$ with $x_{i,1} > x_{j,1}$ and $x_{i,1} + x_{i,2} < x_{j,1} + x_{j,2}$, which, being $x_{i,r}$ and $x_{j,r}$ binary, is impossible. For $r \geq 3$, there is no guarantee that $C_r$ will be non-empty. Consider for instance the case of two alternatives with attribute profiles $x_1 = (1, 0, 0)$ and $x_2 = (0, 1, 1)$. In that case, we have $C_3 = \emptyset$. We say that a heuristic is *cumulative dominance compliant* if, whenever $C_k \neq \emptyset$, the heuristic chooses an alternative from $C_k$. It follows from Proposition 1 that:

**Theorem 1.** *For all weights $w_1 \geq w_2 \geq \cdots \geq w_k \geq 0$, $\sum_{r=1}^{k} w_r = 1$, if $C_k$ is non-empty, then any cumulative dominance compliant heuristic will choose a best alternative.*

Theorem 1 is not restricted to the binary attribute case.

Let $r^*$ denote the highest attribute index for which some alternative exhibits cumulative dominance over all other alternatives. Formally,

$$r^* = \max_{1 \leq r \leq k} \{r \ : \ C_r \neq \emptyset\}.$$

By definition, $C_r = \emptyset$, $r^* < r \leq k$. Of course, $C_k$ is non-empty if and only if $r^* = k$. In the binary attribute case, $r^* \geq 2$. For non-binary attributes, $r^*$ could be equal to 1. A heuristic is said to be *fully cumulative dominance compliant* if it always chooses an alternative from $C_{r^*}$. Fully cumulative dominance compliance implies cumulative dominance compliance. The following result has been shown in BCH:

**Theorem 2.** DEBA *fully complies with cumulative dominance.*

Let $c$ the alternative chosen by *DEBA*. Then, the loss of the heuristic is

$$L = \max_{1 \leq i \leq m} U_i - U_c.$$

The approach taken in BCH to upper bound $E[L]$ for any fully cumulative compliant heuristic was to compute the probability mass function of $r^*$, $P(r) = P[r^* = r]$, $2 \leq r \leq k - 1$, and to use the upper bound for $E[L \mid r^* = r]$, $2 \leq r \leq k - 1$, given by the following theorem, which holds in the more general case of arbitrary attributes taking values in the interval $[0, 1]$:

**Theorem 3.** *For any fully cumulative compliant heuristic $E[L \mid r^* = r] \leq (k - r)/k$, $1 \leq r \leq k$.*

Since $L$ is 0 when $r^* = k$ and, in the binary attribute case, $r^* \geq 2$, Theorem 3 allow us to write:

$$E[L] \leq E[L]_{\text{ub}} = \sum_{r=2}^{k-1} P(r) \frac{k - r}{k}. \tag{1}$$

The procedure proposed in BCH to compute the probabilities $P(r)$ is as follows. Let $Q(r) = P[r^* \geq r]$, $2 \leq r \leq k$. Clearly:

$$P(r) = Q(r) - Q(r + 1), \quad 2 \leq r \leq k - 1,$$

5

with $Q(2) = 1$, reducing the computation of $P(r)$, $2 \leq r \leq k - 1$ to that of $Q(r)$, $3 \leq r \leq k$. The probabilities $Q(r)$, $3 \leq r \leq k$ were computed using ROBDDs (Reduced Ordered Binary Decision Diagrams) with complement 0-edges. ROBDDs (Bryant 1986) are canonical representations of Boolean functions which only depend on the ordering of the binary variables. Given a Boolean function $F(x_1, x_2, \ldots, x_n)$ of $n$ independent Bernoulli random variables, we can compute $P[F(x_1, x_2, \ldots, x_n) = 1]$ by building the ROBDD of $F()$ as a function of $x_1, x_2, \ldots, x_n$ and, then, making a depth-first traversal of the ROBDD starting at the root node. When returning from the visit to each non-terminal node $n$ with binary variable associated with it $x$, we obtain the probability that the Boolean function represented by $n$ is equal to 1 by multiplying the probability that the Boolean function represented by the 0-edge node is equal to 1 by the probability that $x$ has value 0, multiplying the probability that the Boolean function represented by the 1-edge node by the probability that $x$ has value 1, and adding up those partial results. The ROBDD of a Boolean function can be built from a description of the function using Boolean operators (for instance, NOT ($\neg$), AND($\wedge$), OR($\vee$)), which can be looked at as a combinational circuit (Ercegovac and Lang 1985) having inputs $x_1, x_2, \ldots, x_n$ and a single output, by traversing depth-first the description and using well-known procedures (Bryant 1986) to obtain the ROBDD of the output of a NOT gate from the ROBDD of the input, and the ROBDD of the output of a two-input gate (for instance, AND, OR) from the ROBDDs of the two inputs. ROBDDs with complement 0-edges (Brace et al. 1990) are a variant of ROBDDs in which the root may represent either the given Boolean function $F(x_1, x_2, \ldots, x_n)$ or its complement and non-terminal nodes may have either a 0-edge node and a 1-edge node or a complement 0-edge node an a 1-edge node, the complement 0-edge node of a node $n$ representing the complement of the Boolean function obtained from the Boolean function represented by $n$ by setting the variable $x$ associated with $n$ to 0. Adapting the previously described procedure to compute $P[F(x_1, x_2, \ldots, x_n) = 1]$ to that variant of ROBDDs is trivial. To build ROBDDs with complement 0-edges the well-known CU Decision Diagram package (Somenzi 2005) was used. The binary variables were sorted taking into account the structure of the combinational circuit representing $F(x_1, x_2, \ldots, x_n)$ using the topology heuristic (Nikolskaia et al. 1998).

The probability $Q(r)$ can be computed as the probability that the indicator function of the event $\{r^* \geq r\} = \{C_r \neq \emptyset\}$ is equal to 1. For the ZIAC model, the independent Bernoulli random variables to be considered are $x_{i,s}$, $1 \leq i \leq m$, $1 \leq s \leq r$, and the combinational circuit was built based on

$$\mathbf{1}_{\{C_r \neq \emptyset\}} = \bigvee_{i=1}^{m} \bigwedge_{\substack{j=1 \\ j \neq i}}^{m} \bigwedge_{s=1}^{r} \mathbf{1}_{\{X_{i,s} \geq X_{j,s}\}} \,.$$

For $s = 1$, the logic used to generate $\mathbf{1}_{\{X_{i,1} \geq X_{j,1}\}}$, $1 \leq i \leq m$, $1 \leq j \leq m$, $j \neq i$ was

$$\mathbf{1}_{\{X_{i,1} \geq X_{j,1}\}} = \mathbf{1}_{\{x_{i,1} \geq x_{j,1}\}} = x_{i,1} \vee \neg x_{j,1} \,.$$

For generating $\mathbf{1}_{\{X_{i,s} \geq X_{j,s}\}}$, $1 \leq i \leq m$, $1 \leq j \leq m$, $j \neq i$, $2 \leq s \leq r$ from $x_{i,s}$, $1 \leq i \leq m$, $1 \leq j \leq m$, $j \neq i$, $2 \leq s \leq r$ we used $m(r - 1)$ specialized binary carry propagate adders and $m(m - 1)(r - 1)$ binary comparators (see BCH, for details). For the PIAC model, the independent Bernoulli random variables to be considered are $z_i$, $1 \leq i \leq m$ and $y_{i,s}^l$, $y_{i,s}^h$, $1 \leq i \leq m$, $1 \leq s \leq r$ and the combinational circuit yielding $\mathbf{1}_{\{C_r \neq \emptyset\}}$ was built as for the ZIAC model, adding the logic to

obtain $x_{i,s}$, $1 \le i \le m$, $1 \le s \le r$

$$x_{i,s} = \neg z_i \wedge y_{i,s}^l \vee z_i \wedge y_{i,s}^h \,.$$

Intuitively, it is clear that the upper bounds for $E[L \mid r^* = r]$ given by Theorem 3 on which the upper bound $E[L]_{\mathrm{ub}}$ for $E[L]$ given by (1) are based can be quite rough for *DEBA*. Being *DEBA* fully cumulative dominance compliant, the alternative $c$ chosen by *DEBA* will belong to $C_{r^*}$. Then, for $r^* = r$, $c$ can only loose from attribute $r + 1$ on and the upper bounds given by Theorem 3 are obtained by assuming that $c$ will have an all-0 pattern from attribute $r + 1$ on and some other alternative will have an all-1 pattern from attribute $r+1$ on, giving a loss from attribute $r+1$ on equal to $\sum_{s=r+1}^{k} w_s$, which is (tightly) bounded by $(k - r)/k$. The improved upper bounds for $E[L]$ for *DEBA* derived in this paper are simply obtained by deriving better upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ using the properties of the probabilistic models under consideration.

## 3 Improved upper bounds for the ZIAC model

The goal of this section is to derive improved upper bounds $E[L \mid r^* = r]_{\mathrm{iub}}$ for $E[L \mid r^* = r]$, $2 \le r \le k-1$ for the ZIAC probabilistic model taking advantage of the properties of *DEBA* and the ZIAC probabilistic model. Use of those improved upper bounds in conjunction with the ROBDD-based approach used in BCH and reviewed in the previous section to compute the probability mass function of $r^*$, $P(r) = P[r^* = r]$, $2 \le r \le k - 1$, will yield the improved upper bound for $E[L]$

$$E[L]_{\mathrm{iub}} = \sum_{r=2}^{k-1} P(r)E[L \mid r^* = r]_{\mathrm{iub}} \,. \tag{2}$$

In the derivation of the improved upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ we will use the following theorem.

**Theorem 4.** *Let* $r$, $1 \le r \le k$ *and let* $\mathcal{E}_r$ *be any event depending only on* $x_{j,s}$, $1 \le j \le m$, $1 \le s \le r - 1$ *such that* $P[i$ *is chosen by DEBA* $\wedge \mathcal{E}_r] > 0$. *Let* $\mathbf{x}_{i,r}$ *be the* $(m - 1)(k - r + 1)$-*vector with components* $x_{j,s} - x_{i,s}$, $1 \le j \le m$, $j \ne i$, $r \le s \le k$. *Let* $F(u_1, \ldots, u_{(m-1)(k-r+1)})$ *be any function which is non-decreasing on each* $u_l$. *Then, for the ZIAC model,*

$$E[F(\mathbf{x}_{i,r}) \mid i \text{ is chosen by DEBA } \wedge \mathcal{E}_r] \le E[F(\mathbf{x}_{i,r})] \,.$$

*Proof.* See the Appendix. $\qquad\square$

Essentially, what Theorem 4 says is that, conditioned on any event $\mathcal{E}_r$ do not depending on the alternatives from attribute $r$ on and which does not preclude any given alternative $i$ to be chosen by *DEBA*, and in terms of a function which is non-decreasing on the differences between the values of the attributes from attribute $r$ on of the other alternatives and the corresponding attributes of

the alternative chosen by *DEBA*, for the ZIAC probabilistic model, *DEBA* cannot perform worse in average than a random selection ($E[F(\mathbf{x}_{i,r})]$ is, by symmetry, equal for all $i$, $1 \leq i \leq m$).

The improved upper bounds for $E[L \mid r^* = r]$, $2 \leq r \leq k - 1$ are given by the following theorem, where $I_+(x) = \max\{0, x\}$. The theorem also asserts that the bounds potentially improve those used in BCH. The sketch of the derivation of the bounds is as follows. First, it is relatively easy to show using linear programming results that

$$E[L \mid r^* = r] \leq \max_{r+1 \leq t \leq k} \frac{V(r,t)}{t},$$

where $V(r, t)$ is any upper bound for $E[S(r, t) \mid r^* = r]$ with

$$S(r,t) = I_+ \left( \max_{\substack{1 \leq i \leq m \\ i \neq c}} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right\} \right).$$

To derive the upper bound, we exploit the symmetries of the ZIAC model and use Theorem 4 to conclude that

$$E[S(r,t) \mid r^* = r] \leq E[Z(r,t)].$$

**Theorem 5.** *For DEBA and the ZIAC model,*

$$E[L \mid r^* = r] \leq E[L \mid r^* = r]_{\mathrm{iub}} = \max_{r+1 \leq t \leq k} \frac{V(r,t)}{t}, \quad 2 \leq r \leq k - 1,$$

*where*

$$V(r,t) = E[Z(r,t)],$$

*with*

$$Z(r,t) = I_+ \left( \max_{1 \leq i \leq m-1} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{m,s}) \right\} \right).$$

*Furthermore, $E[L \mid r^* = r]_{\mathrm{iub}} \leq (k - r)/k$, $2 \leq r \leq k$.*

*Proof.* Assume $r^* = r$, $2 \leq r \leq k - 1$, and let $c$ be the alternative chosen by *DEBA*. For any alternative $i$, $1 \leq i \leq m$, we can write

$$U_i = \sum_{s=1}^{r} (w_s - w_{s+1}) X_{i,s} + w_{r+1} X_{i,r+1} + \sum_{s=r+2}^{k} w_s x_{i,s}$$

and, since $c$ has cumulative dominance over all other alternatives up to attribute $r$ and weights are non-negative and non-increasing, for all $i$, $1 \leq i \leq m$, $i \neq c$,

$$U_i - U_c \leq w_{r+1}(X_{i,r+1} - X_{c,r+1}) + \sum_{s=r+2}^{k} w_s(x_{i,s} - x_{c,s}),$$

yielding

$$L \leq I_+ \left( \max_{\substack{1 \leq i \leq m \\ i \neq c}} \left\{ w_{r+1}(X_{i,r+1} - X_{c,r+1}) + \sum_{s=r+2}^{k} w_s(x_{i,s} - x_{c,s}) \right\} \right). \tag{3}$$

8

We continue by finding the maximum of $w_{r+1}(X_{i,r+1} - X_{c,r+1}) + \sum_{s=r+2}^{k} w_s(x_{i,s} - x_{c,s})$ subject to the constraints defining the domain of possible values for $w_{r+1}, w_{r+2}, \ldots, w_k$:

$$w_k \geq 0\,,$$

$$w_s \geq w_{s+1}, \quad r+1 \leq s \leq k-1\,,$$

$$(r+1)w_{r+1} + \sum_{s=r+2}^{k} w_s \leq 1\,,$$

where the last one comes from $w_1 \geq w_2 \geq \cdots \geq w_r \geq w_{r+1}$ and $\sum_{s=1}^{k} w_s = 1$. This is a linear programming problem (Luenberger, 2003). Using the variables $x_s = w_s - w_{s+1}, r+1 \leq s \leq k-1$ and a slack variable $y$, the problem can be put into standard form with the restrictions:

$$\sum_{s=r+1}^{k-1} sx_s + kw_k + y = 1\,,$$

$$w_k \geq 0\,,$$

$$x_s \geq 0, \quad r+1 \leq s \leq k-1\,,$$

$$y \geq 0\,,$$

which has basic feasible solutions

$$(x_{r+1}, x_{r+2}, \ldots, x_{k-1}, w_k, y) = \left(\frac{1}{r+1}, 0, \ldots, 0, 0, 0\right),$$

$$(x_{r+1}, x_{r+2}, \ldots, x_{k-1}, w_k, y) = \left(0, \frac{1}{r+2}, \ldots, 0, 0, 0\right),$$

$$\vdots$$

$$(x_{r+1}, x_{r+2}, \ldots, x_{k-1}, w_k, y) = \left(0, 0, \ldots, \frac{1}{k-1}, 0, 0\right),$$

$$(x_{r+1}, x_{r+2}, \ldots, x_{k-1}, w_k, y) = \left(0, 0, \ldots, 0, \frac{1}{k}, 0\right),$$

$$(x_{r+1}, x_{r+2}, \ldots, x_{k-1}, w_k, y) = (0, 0, \ldots, 0, 0, 1)\,.$$

Since the convex domain defined by the restrictions is bounded, the maximum is achieved at some basic feasible solution. Those basic feasible solutions correspond to the points $(w_{r+1}, w_{r+2}, \ldots, w_k)$

$$(w_{r+1}, w_{r+2}, \ldots, w_k) = \left(\frac{1}{r+1}, 0, \ldots, 0\right),$$

$$(w_{r+1}, w_{r+2}, \ldots, w_k) = \left(\frac{1}{r+2}, \frac{1}{r+2}, \ldots, 0\right),$$

$$\vdots$$

$$(w_{r+1}, w_{r+2}, \ldots, w_k) = \left(\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}\right),$$

$$(w_{r+1}, w_{r+2}, \ldots, w_k) = (0, 0, \ldots, 0)\,,$$

9

and, therefore, the maximum of $w_{r+1}(X_{i,r+1} - X_{c,r+1}) + \sum_{s=r+2}^{k} w_s(x_{i,s} - x_{c,s})$ is

$$\max_{r+1 \le t \le k} \frac{1}{t} \left( X_{i,r+1} - X_{c,r+1} + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right).$$

Using that maximum in (3):

$$L \le I_+ \left( \max_{\substack{1 \le i \le m \\ i \ne c}} \max_{r+1 \le t \le k} \frac{1}{t} \left( X_{i,r+1} - X_{c,r+1} + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right) \right)$$

$$= \max_{r+1 \le t \le k} \frac{1}{t} I_+ \left( \max_{\substack{1 \le i \le m \\ i \ne c}} \left\{ X_{i,r+1} - X_{c,r+1} + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right\} \right).$$

But, because $c$ cumulative dominates all other alternatives up to attribute $r$ and attributes are binary, necessarily $X_{i,r+1} \le X_{c,r+1} + 1$, $i \ne c$, and

$$L \le \max_{r+1 \le t \le k} \frac{S(r,t)}{t}$$

with

$$S(r,t) = I_+ \left( \max_{\substack{1 \le i \le m \\ i \ne c}} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right\} \right),$$

yielding

$$E[L \mid r^* = r] \le \max_{r+1 \le t \le k} \frac{V(r,t)}{t},$$

where $V(r,t)$ is any upper bound for $E[S(r,t) \mid r^* = r]$.

Exploiting the symmetries of the ZIAC model,

$$E[S(r,t) \mid r^* = r] = E[Z(r,t) \mid m \text{ is chosen by } \textit{DEBA} \, \wedge \, r^* = r],$$

with

$$Z(r,t) = I_+ \left( \max_{1 \le i \le m-1} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{m,s}) \right\} \right).$$

But, the event $r^* = r$ only depends on $x_{j,s}$, $1 \le j \le m$, $1 \le s \le r+1$, $P[m \text{ is chosen by } \textit{DEBA} \wedge r^* = r] > 0$, and $Z(r,t)$ is non-decreasing on each $x_{i,s} - x_{m,s}$, $1 \le i \le m-1$, $r+2 \le s \le k$, and, then, using Theorem 4,

$$E[S(r,t) \mid r^* = r] \le E[Z(r,t)].$$

It remains to show that $E[L \mid r^* = r]_{\text{iub}} \le (k-r)/k$, $2 \le r \le k-1$. Since $Z(r,t) \le t-r$, we have $E[Z(r,t)] \le t-r$. Then, $E[L \mid r^* = r]_{\text{iub}} \le \max_{r+1 \le t \le k}(t-r)/r = (k-r)/k$. $\qquad \square$

The $E[Z(r,t)]$'s involved in the improved upper bounds for $E[L \mid r^* = r]$ given by Theorem 4 can be obtained by using the recurrence-based computational scheme given by the following theorem. The recurrences can be easily obtained from the definition of $E[Z(r,t)]$.

**Theorem 6.** *For the ZIAC model, $E[Z(r,t)]$, $2 \le r \le k-1$, $r+1 \le t \le k$ can be computed using*

$$E[Z(r,t)] = \sum_{a=1}^{t-r} a\,\psi(r,t,a)$$

*and the recurrences:*

$$\psi(r,t,a) = \sum_{a'=a}^{t-r} \phi(m-1,r,t,a')\,\pi(r,t,a'-a), \quad 2 \le r \le k-1, r+1 \le t \le k, 1 \le a \le t-r,$$

$$\phi(1,r,t,a) = \pi(r,t,a-1), \quad 2 \le r \le k-1, r+1 \le t \le k, 1 \le a \le t-r,$$

$$\phi(b,r,t,a) = \phi(b-1,r,t,a)\sum_{a'=0}^{a-1}\pi(r,t,a') + \pi(r,t,a-1)\sum_{a'=1}^{a-1}\phi(b-1,r,t,a'),$$

$$2 \le b \le m-1,\, 2 \le r \le k-1,\, r+1 \le t \le k,\, 1 \le a \le t-r,$$

$$\pi(r,r+1,0) = 1, \quad 2 \le r \le k-1,$$

$$\pi(r,r+2,0) = 1 - p_{r+2}, \quad 2 \le r \le k-2,$$

$$\pi(r,r+2,1) = p_{r+2}, \quad 2 \le r \le k-2,$$

$$\pi(r,t,t-r-1) = p_{r+2}\,\pi(r+1,t,t-r-2), \quad 2 \le r \le k-3,\, r+3 \le t \le k,$$

$$\pi(r,t,a) = (1-p_{r+2})\pi(r+1,t,a) + p_{r+2}\,\pi(r+1,t,a-1),$$

$$2 \le r \le k-3,\, r+3 \le t \le k,\, 1 \le a \le t-r-2,$$

$$\pi(r,t,0) = (1-p_{r+2})\pi(r+1,t,0), \quad 2 \le r \le k-3,\, r+3 \le t \le k.$$

*Proof.* Let, for $2 \le r \le k-1$ and $r+1 \le t \le k$, be the random variables

$$H_i(r,t) = \sum_{s=r+2}^{t} x_{i,s}, \quad 1 \le i \le m,$$

$$I(b,r,t) = \max_{1 \le i \le b}\left\{1 + \sum_{s=r+2}^{t} x_{i,s}\right\}, \quad 1 \le b \le m-1,$$

and let (by symmetry, all $H_i(r,t)$, $1 \le i \le m$ have the same probability mass function)

$$\pi(r,t,a) = P[H_i(r,t) = a],$$

$$\phi(b,r,t,a) = P[I(b,r,t) = a],$$

$$\psi(r,t,a) = P[Z(r,t) = a].$$

Then, the result follows using elementary probability theory by noting that

$$I(1,r,t) = 1 + H_1(r,t),$$

$$I(b,r,t) = \max\{I(b-1,r,t), 1+H_b(r,t)\}, \quad 2 \le b \le m-1,$$

$$Z(r,t) = I_+(I(m-1,r,t) - H_m(r,t)).$$

$\square$

## 4 Improved upper bounds for the PIAC model

In this section, we derive improved upper bounds $E[L|r^* = r]_{\text{iub}}$ for $E[L|r^* = r]$, $2 \le r \le k-1$ for the PIAC probabilistic model taking advantage of the properties of *DEBA* and the PIAC probabilistic model. The bounds will be obtained in terms of the conditional probabilities $P[\mathcal{A}_r^l \mid |G| = g]$, $2 \le r \le k$, $0 \le g \le m - 1$, $P[\mathcal{A}_r^h \mid |G| = g]$, $2 \le r \le k$, $1 \le g \le m$, and $P[\mathcal{A}_r^{l,h} \mid |G| = g]$, $2 \le r \le k$, $1 \le g \le m-1$, where $G$ is the random variable "number of good alternatives", $\mathcal{A}_r^l$ is the event "some bad alternative but no good alternative cumulative dominates all other alternatives up to attribute $r$", $\mathcal{A}_r^h$ is the event "some good alternative but no bad alternative cumulative dominates all other alternatives up to attribute $r$", and $\mathcal{A}_r^{l,h}$ is the event "some good alternative and some bad alternative cumulative dominates all other alternatives up to attribute $r$". The main reason why the upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ are obtained in terms of those conditional probabilities is that the conditional probabilities can be computed using ROBDDs which turned out to have moderate sizes for values of $m$ and $k$ as large as 10. The upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ will also depend on the probability mass function of $r^*$, $P(r) = P[r^* = r]$, $2 \le r \le k - 1$. That probability mass function can be easily obtained from the conditional probabilities and can be used to obtain the improved upper bound $E[L]_{\text{iub}}$ for $E[L]$ from the upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ using (2)

$$E[L]_{\text{iub}} = \sum_{r=2}^{k-1} P(r) E[L \mid r^* = r]_{\text{iub}}.$$

We start by discussing how the conditional probabilities can be computed using ROBDDs. Then, we explain how the probability mass function of $r^*$, $P(r)$, $2 \le r \le k - 1$ can be computed from the conditional probabilities. Finally, we obtain the improved upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k - 1$ in terms of the conditional probabilities and the probability mass function of $r^*$, $P(r)$, $2 \le r \le k - 1$.

To compute the conditional probabilities using ROBDDs, we introduce two Bernouilli random variables, $a_0$ and $a_1$, independent of $z_i$, $1 \le i \le m$ and $y_{i,s}^l$, $y_{i,s}^h$, $1 \le i \le m$, $1 \le s \le k$ and define the event $\mathcal{B}_r = \mathcal{B}_r^l \cap \mathcal{B}_r^h$, where $\mathcal{B}_r^l$ is the event "either $a_0 = 0$ or some bad alternative cumulative dominates all other alternatives up to attribute $r$" and $\mathcal{B}_r^h$ is the event "either $a_1 = 0$ or some good alternative cumulative dominates all other alternatives up to attribute $r$". By the definition of the events $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, and $\mathcal{A}_r^{l,h}$, it is clear that

$$P[\mathcal{A}_r^l \cup \mathcal{A}_r^{l,h} \mid |G| = g] = P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \,\wedge\, a_0 = 1 \,\wedge\, a_1 = 0],$$

$$P[\mathcal{A}_r^h \cup \mathcal{A}_r^{l,h} \mid |G| = g] = P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \,\wedge\, a_0 = 0 \,\wedge\, a_1 = 1],$$

$$P[\mathcal{A}_r^{l,h} \mid |G| = g] = P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \,\wedge\, a_0 = 1 \,\wedge\, a_1 = 1]. \tag{4}$$

Further, since $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, and $\mathcal{A}_r^{l,h}$ are disjoint,

$$
\begin{aligned}
P[\mathcal{A}_r^l \mid |G| = g] &= P[\mathcal{A}_r^l \cup \mathcal{A}_r^{l,h} \mid |G| = g] - P[\mathcal{A}_r^{l,h} \mid |G| = g] \\
&= P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \,\wedge\, a_0 = 1 \,\wedge\, a_1 = 0] \\
&\quad - P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \,\wedge\, a_0 = 1 \,\wedge\, a_1 = 1]
\end{aligned}
\tag{5}
$$

and

$$
\begin{aligned}
P[\mathcal{A}_r^h \mid |G| = g] &= P[\mathcal{A}_r^h \cup \mathcal{A}_r^{l,h} \mid |G| = g] - P[\mathcal{A}_r^{l,h} \mid |G| = g] \\
&= P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 0 \wedge a_1 = 1] \\
&\quad - P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 1 \wedge a_1 = 1] \qquad (6)
\end{aligned}
$$

Using (4), (5), and (6), the conditional probabilities can be computed from $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 1 \wedge a_1 = 0]$, $2 \le r \le k$, $0 \le g \le m - 1$, $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 0 \wedge a_1 = 1]$, $2 \le r \le k$, $1 \le g \le m$, and $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 1 \wedge a_1 = 1]$, $2 \le r \le k$, $0 \le g \le m$. Using the symmetries of the PIAC model, the definition of conditional probability, and the fact that the underlying Bernouilli random variables $a_0$, $a_1$, $z_i$, $1 \le i \le m$, and $y_{i,s}^l$, $y_{i,s}^h$, $1 \le i \le m$, $1 \le s \le r$ are independent, $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 1 \wedge a_1 = 0]$ can be computed as $P[\mathbf{1}_{\mathcal{B}_r} = 1]$, with the "success" probability of the Bernouilli random variables $z_i$, $1 \le i \le g$ set to 1, the "success" probability of the Bernouilli random variables $z_i$, $g + 1 \le i \le m$ set to 0, the "success" probability of the Bernouilli random variable $a_0$ set to 1, and the "success" probability of the Bernouilli random variable $a_1$ set to 0. Similarly, $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 0 \wedge a_1 = 1]$ can be computed as $P[\mathbf{1}_{\mathcal{B}_r} = 1]$, with the "success" probability of the Bernouilli random variables $z_i$, $1 \le i \le g$ set to 1, the "success" probability of the Bernouilli random variables $z_i$, $g + 1 \le i \le m$ set to 0, the "success" probability of the Bernouilli random variable $a_0$ set to 0, and the "success" probability of the Bernouilli random variable $a_1$ set to 1; and $P[\mathbf{1}_{\mathcal{B}_r} = 1 \mid |G| = g \wedge a_0 = 1 \wedge a_1 = 1]$ can be computed as $P[\mathbf{1}_{\mathcal{B}_r} = 1]$, with the "success" probability of the Bernouilli random variables $z_i, 1 \le i \le g$ set to 1, the "success" probability of the Bernouilli random variables $z_i$, $g + 1 \le i \le m$ set to 0, the "success" probability of the Bernouilli random variable $a_0$ set to 1, and the "success" probability of the Bernouilli random variable $a_1$ set to 1. All those expectations can be computed from a ROBDD representation of $\mathbf{1}_{\mathcal{B}_r}$ as a function of $a_0$, $a_1$, $z_i$, $1 \le i \le m$, $y_{i,s}^l$, $y_{i,s}^h$, $1 \le i \le m$, $1 \le s \le r$. As in BCH, ROBDDs with complement 0-edges can be built with the help of the CU Decision Diagram Package. A combinational circuit yielding $\mathbf{1}_{\mathcal{B}_r}$ as a function of $a_0$, $a_1$, $z_i$, $1 \le i \le m$, $y_{i,s}^l$, $y_{i,s}^h$, $1 \le i \le m$, $1 \le s \le r$ can be built based on

$$
\mathbf{1}_{\mathcal{B}_r} = \left[ \neg a_0 + \bigvee_{i=1}^m \neg z_i \bigwedge_{\substack{j=1 \\ j \ne i}}^m \bigwedge_{s=1}^r \mathbf{1}_{\{X_{i,s} \ge X_{j,s}\}} \right] \wedge \left[ \neg a_1 + \bigvee_{i=1}^m z_i \bigwedge_{\substack{j=1 \\ j \ne i}}^m \bigwedge_{s=1}^r \mathbf{1}_{\{X_{i,s} \ge X_{j,s}\}} \right],
$$

where the indicator functions $\mathbf{1}_{\{X_{i,s} \ge X_{j,s}\}}$, $1 \le i \le m$, $1 \le j \le m$, $j \ne i$, $1 \le s \le r$ can be expressed in terms of the variables $z_i$, $1 \le i \le m$, $y_{i,s}^l$, $y_{i,s}^h$, $1 \le i \le m$, $1 \le s \le r$ as explained in Section 2.

The following theorem gives expressions for $P(r)$, $2 \le r \le k - 1$ in terms of $P[\mathcal{A}_r^l \mid |G| = g]$, $0 \le g \le m - 1$, $2 \le r \le k$, $P[\mathcal{A}_r^h \mid |G| = g]$, $1 \le g \le m$, $2 \le r \le k$, and $P[\mathcal{A}_r^{l,h} \mid |G| = g]$, $1 \le g \le m - 1$, $2 \le r \le k$. Its proof is immediate from the definitions of those conditional probabilities.

**Theorem 7.** *For the PIAC model and* $2 \leq r \leq k - 1$,

$$
\begin{aligned}
P(r) \quad = \quad & (1-p)^m \left( P[\mathcal{A}_r^l \,|\, |G| = 0] - P[\mathcal{A}_{r+1}^l \,|\, |G| = 0] \right) \\
& + \sum_{g=1}^{m-1} \binom{m}{g} p^g (1-p)^{m-g} \\
& \qquad \left( P[\mathcal{A}_r^l \,|\, |G| = g] + P[\mathcal{A}_r^h \,|\, |G| = g] + P[\mathcal{A}_r^{l,h} \,|\, |G| = g] \right. \\
& \qquad \quad \left. - P[\mathcal{A}_{r+1}^l \,|\, |G| = g] - P[\mathcal{A}_{r+1}^h \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g] \right) \\
& + p^m \left( P[\mathcal{A}_r^h \,|\, |G| = m] - P[\mathcal{A}_{r+1}^h \,|\, |G| = m] \right).
\end{aligned}
$$

*Proof.* Conditioning on $|G|$ and using $P[|G| = g] = \binom{m}{g} p^g (1-p)^{m-g}$, $0 \leq g \leq m$,

$$
\begin{aligned}
P(r) \quad = \quad & P[r^* = r] = \sum_{g=0}^{m} P[|G| = g] P[r^* = r \,|\, |G| = g] \\
= \quad & \sum_{g=0}^{m} \binom{m}{g} p^g (1-p)^{m-g} P[r^* = r \,|\, |G| = g]. \quad\quad (7)
\end{aligned}
$$

Using the definitions of $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, and $\mathcal{A}_r^{l,h}$, for $2 \leq r \leq k - 1$,

$$
\{r^* = r\} = \mathcal{A}_r^l \cup \mathcal{A}_r^h \cup \mathcal{A}_r^{l,h} - \left( \mathcal{A}_{r+1}^l \cup \mathcal{A}_{r+1}^h \cup \mathcal{A}_{r+1}^{l,h} \right),
$$

and, since $\mathcal{A}_{r+1}^l \cup \mathcal{A}_{r+1}^h \cup \mathcal{A}_{r+1}^{l,h} \subset \mathcal{A}_r^l \cup \mathcal{A}_r^h \cup \mathcal{A}_r^{l,h}$, $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, and $\mathcal{A}_r^{l,h}$ are disjoint, and $\mathcal{A}_{r+1}^l$, $\mathcal{A}_{r+1}^h$, and $\mathcal{A}_{r+1}^{l,h}$ are disjoint,

$$
\begin{aligned}
P[r^* = r \,|\, |G| = g] \quad = \quad & P[\mathcal{A}_r^l \,|\, |G| = g] + P[\mathcal{A}_r^h \,|\, |G| = g] + P[\mathcal{A}_r^{l,h} \,|\, |G| = g] \\
& - P[\mathcal{A}_{r+1}^l \,|\, |G| = g] - P[\mathcal{A}_{r+1}^h \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g],
\end{aligned}
$$

which used in (7), taking into account $P[\mathcal{A}_r^l \,|\, |G| = m] = P[\mathcal{A}_{r+1}^l \,|\, |G| = m] = P[\mathcal{A}_r^h \,|\, |G| = 0] = P[\mathcal{A}_{r+1}^h \,|\, |G| = 0] = P[\mathcal{A}_r^{l,h} \,|\, |G| = 0] = P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = 0] = P[\mathcal{A}_r^{l,h} \,|\, |G| = m] = P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = m] = 0$, $2 \leq r \leq k - 1$ gives the expression for $P(r)$, $2 \leq r \leq k - 1$. $\qquad\square$

The path to the derivation of the improved upper bounds for $E[L \,|\, r^* = r]$, $2 \leq r \leq k - 1$ starts with the following theorem, which can be looked at as a natural extension to the PIAC model of Theorem 4 in the previous section.

**Theorem 8.** *Let* $r$, $1 \leq r \leq k$, *let* $G' \subset \{1, 2, \ldots, m\}$, *and let* $\mathcal{E}_r$ *be any event depending only on* $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r - 1$ *such that* $P[i \text{ is chosen by DEBA } \wedge \mathcal{E}_r \,|\, G = G'] > 0$. *Let* $\mathbf{x}_{i,r}$ *be the* $(m-1)(k-r+1)$-*vector with components* $x_{j,s} - x_{i,s}$, $1 \leq j \leq m$, $j \neq i$, $r \leq s \leq k$. *Let* $F(u_1, \ldots, u_{(m-1)(k-r+1)})$ *be any function which is non-decreasing on each* $u_l$. *Then, for the PIAC model,*

$$
E[F(\mathbf{x}_{i,r}) \,|\, G = G' \wedge i \text{ is chosen by DEBA } \wedge \mathcal{E}_r] \leq E[F(\mathbf{x}_{i,r}) \,|\, G = G'].
$$

*Proof.* See the Appendix. □

Using Theorem 8, the derivation of the improved upper bounds for $E[L \mid r^* = r]$, $2 \leq r \leq k - 1$ is, roughly, as follows. First, retaking a result obtained in the previous section,

$$E[L \mid r^* = r] \leq \max_{r+1 \leq t \leq k} \frac{V(r, t)}{t},$$

where, $c$ being the alternative chosen by *DEBA*, $V(r, t)$ is any upper bound for $E[S(r, t) \mid r^* = r]$ with

$$S(r, t) = I_+ \left( \max_{\substack{1 \leq i \leq m \\ i \neq c}} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{c,s}) \right\} \right).$$

That upper bound is obtained in terms of upper bounds for $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^l]$, $0 \leq g \leq m - 1$, $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^h]$, $1 \leq g \leq m$, and $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^{l,h}]$, $1 \leq g \leq m - 1$, $P(r)$, and the conditional probabilities $P[\mathcal{A}_r^l \mid |G| = g]$, $P[\mathcal{A}_{r+1}^l \mid |G| = g]$, $0 \leq g \leq m - 1$, $P[\mathcal{A}_r^h \mid |G| = g]$, $P[\mathcal{A}_{r+1}^h \mid |G| = g]$, $1 \leq g \leq m$, and $P[\mathcal{A}_r^{l,h} \mid |G| = g]$, $P[\mathcal{A}_{r+1}^{l,h} \mid |G| = g]$, $1 \leq g \leq m - 1$ using elementary probability theory and the relationships between the events $\{r^* = r\}$, $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, $\mathcal{A}_r^{l,h}$, $\mathcal{A}_{r+1}^l$, $\mathcal{A}_{r+1}^h$, and $\mathcal{A}_{r+1}^{l,h}$. The upper bounds for $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^l]$, $0 \leq g \leq m - 1$, $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^h]$, $1 \leq g \leq m$, and $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^{l,h}]$, $1 \leq g \leq m - 1$ are obtained by exploiting the symmetries of the PIAC model and using Theorem 8.

The following proposition gives the upper bounds for $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^l]$, $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^h]$, and $E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^{l,h}]$.

**Proposition 2.** *Let $2 \leq r \leq k - 1$ and $r + 1 \leq t \leq k$. Then, for DEBA and the PIAC model,*

$$E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^l] \leq E[Z_{l,g}(r, t)], \quad 0 \leq g \leq m - 1,$$

*with*

$$Z_{l,g}(r, t) = I_+ \left( \max \left\{ \mathbf{1}_{m-g>1} \max_{1 \leq i \leq m-g-1} \left\{ 1 + \sum_{s=r+2}^{t} (y_{i,s}^l - y_{m,s}^l) \right\}, \right.\right.$$
$$\left.\left. \mathbf{1}_{g>0} \max_{m-g \leq i \leq m-1} \left\{ 1 + \sum_{s=r+2}^{t} (y_{i,s}^h - y_{m,s}^l) \right\} \right\} \right),$$

$$E[S(r, t) \mid |G| = g \wedge r^* = r \wedge \mathcal{A}_r^h] \leq E[Z_{h,g}(r, t)], \quad 1 \leq g \leq m,$$

*with*

$$Z_{h,g}(r, t) = I_+ \left( \max \left\{ \mathbf{1}_{m-g>0} \max_{1 \leq i \leq m-g} \left\{ 1 + \sum_{s=r+2}^{t} (y_{i,s}^l - y_{m,s}^h) \right\}, \right.\right.$$
$$\left.\left. \mathbf{1}_{g>1} \max_{m-g+1 \leq i \leq m-1} \left\{ 1 + \sum_{s=r+2}^{t} (y_{i,s}^h - y_{m,s}^h) \right\} \right\} \right),$$

*and*

$$E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}]$$
$$\leq \ \max \{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\}, \ \ 1 \leq g \leq m-1 \,.$$

*Proof.* Let $0 \leq g \leq m-1$, let $\mathcal{G}_{l,g}$ be the event "alternatives $1, \ldots, m-g-1$ are bad, alternatives $m-g, \ldots, m-1$ are good, and alternative $m$ is bad", and let $\mathcal{C}_{l,g,r}$ be the event "some alternative in the subset $\{1, \ldots, m-g-1, m\}$ but no alternative in the subset $\{m-g, \ldots, m-1\}$ cumulative dominates all other alternatives up to attribute $r$". Exploiting the symmetries of the PIAC model,

$$E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l]$$
$$= \ E[Z(r,t) \mid \mathcal{G}_{l,g} \ \wedge \ m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{l,g,r}]$$

with

$$Z(r,t) = I_+ \left( \max_{1 \leq i \leq m-1} \left\{ 1 + \sum_{s=r+2}^{t} (x_{i,s} - x_{m,s}) \right\} \right).$$

But, the event $\{r^* = r\} \cap \mathcal{C}_{l,g,r}$ only depends on $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r+1$, $P[m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{l,g,r} \mid \mathcal{G}_{l,g}] > 0$, and $Z(r,t)$ is non-decreasing on each $x_{i,s} - x_{m,s}, 1 \leq i \leq m-1, r+2 \leq s \leq k$, and, then, using Theorem 8,

$$E[Z(r,t) \mid \mathcal{G}_{l,g} \ \wedge \ m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{l,g,r}]$$
$$\leq \ E[Z(r,t) \mid \mathcal{G}_{l,g}] = E[Z_{l,g}(r,t)] \,,$$

yielding

$$E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \leq E[Z_{l,g}(r,t)] \,.$$

Similarly, let $1 \leq g \leq m$, let $\mathcal{G}_{h,g}$ be the event "alternatives $1, \ldots, m-g$ are bad and alternatives $m-g+1, \ldots, m$ are good", and let $\mathcal{C}_{h,g,r}$ be the event "some alternative in the subset $\{m-g+1, \ldots, m\}$ but no alternative in the subset $\{1, \ldots, m-g\}$ cumulative dominates all other alternatives up to attribute $r$". Exploiting the symmetries of the PIAC model,

$$E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h]$$
$$= \ E[Z(r,t) \mid \mathcal{G}_{h,g} \ \wedge \ m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{h,g,r}] \,.$$

But, the event $\{r^* = r\} \cap \mathcal{C}_{h,g,r}$ only depends on $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r+1$, $P[m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{h,g,r} \mid \mathcal{G}_{h,g}] > 0$, and $Z(r,t)$ is non-decreasing on each $x_{i,s} - x_{m,s}, 1 \leq i \leq m-1, r+2 \leq s \leq k$, and, then, using Theorem 8,

$$E[Z(r,t) \mid \mathcal{G}_{h,g} \ \wedge \ m \text{ is chosen by } \textit{DEBA} \ \wedge \ r^* = r \ \wedge \ \mathcal{C}_{h,g,r}]$$
$$\leq \ E[Z(r,t) \mid \mathcal{G}_{h,g}] = E[Z_{h,g}(r,t)] \,,$$

yielding

$$E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] \leq E[Z_{h,g}(r,t)] \,.$$

16

Finally, let $1 \leq g \leq m-1$. Clearly (recall that $c$ denotes the alternative chosen by *DEBA*),

$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h}]$$
$$= \quad P[c \text{ is bad} \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h}]$$
$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is bad}]$$
$$+ \, P[c \text{ is good} \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h}]$$
$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is good}]\,,$$

and, since $P[c \text{ is bad} \,|\, |G| = g \wedge r^* = r \wedge \mathcal{A}_r^{l,h}] \geq 0$, $P[c \text{ is good} \,|\, |G| = g \wedge r^* = r \wedge \mathcal{A}_r^{l,h}] \geq 0$, and $P[c \text{ is bad} \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h}] + P[c \text{ is good} \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h}] = 1$, it suffices to prove that $E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is bad}] \leq E[Z_{l,g}(r,t)]$ and that $E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is good}] \leq E[Z_{h,g}(r,t)]$. To prove the former, let $\mathcal{G}_{l,g}$ be the event previously defined and let $\mathcal{C}'_{l,g,r}$ be the event "some alternative in the subset $\{1,\ldots,m-g-1,m\}$ and some alternative in the subset $\{m-g,\ldots,m-1\}$ cumulative dominates all other alternatives up to attribute $r$". Exploiting the symmetries of the PIAC model,

$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is bad}]$$
$$= \quad E[Z(r,t) \,|\, \mathcal{G}_{l,g} \,\wedge\, m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{l,g,r}]\,.$$

But, the event $\{r^* = r\} \cap \mathcal{C}'_{l,g,r}$ only depends on $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r+1$, $P[m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{l,g,r} \,|\, \mathcal{G}_{l,g}] > 0$, and $Z(r,t)$ is non-decreasing on each $x_{i,s} - x_{m,s}$, $1 \leq i \leq m-1$, $r+2 \leq s \leq k$, and, then, using Theorem 8,

$$E[Z(r,t) \,|\, \mathcal{G}_{l,g} \,\wedge\, m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{l,g,r}]$$
$$\leq \quad E[Z(r,t) \,|\, \mathcal{G}_{l,g}] = E[Z_{l,g}(r,t)]\,,$$

yielding

$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is bad}] \leq E[Z_{l,g}(r,t)]\,.$$

To prove the second, let $\mathcal{G}_{h,g}$ be the event previously defined and let $\mathcal{C}'_{h,g,r}$ be the event "some alternative in the subset $\{1,\ldots,m-g\}$ and some alternative in the subset $\{m-g+1,\ldots,m\}$ cumulative dominates all other alternatives up to attribute $r$". Exploiting the symmetries of the PIAC model,

$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is good}]$$
$$= \quad E[Z(r,t) \,|\, \mathcal{G}_{h,g} \,\wedge\, m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{h,g,r}]\,.$$

But, the event $\{r^* = r\} \cap \mathcal{C}'_{h,g,r}$ only depends on $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r+1$, $P[m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{h,g,r} \,|\, \mathcal{G}_{h,g}] > 0$, and $Z(r,t)$ is non-decreasing on each $x_{i,s} - x_{m,s}$, $1 \leq i \leq m-1$, $r+2 \leq s \leq k$, and, then, using Theorem 8,

$$E[Z(r,t) \,|\, \mathcal{G}_{h,g} \,\wedge\, m \text{ is chosen by } DEBA \,\wedge\, r^* = r \,\wedge\, \mathcal{C}'_{h,g,r}]$$
$$\leq \quad E[Z(r,t) \,|\, \mathcal{G}_{h,g}] = E[Z_{h,g}(r,t)]\,,$$

yielding

$$E[S(r,t) \,|\, |G| = g \,\wedge\, r^* = r \,\wedge\, \mathcal{A}_r^{l,h} \,\wedge\, c \text{ is good}] \leq E[Z_{h,g}(r,t)]\,.$$

$\square$

The improved upper bounds for $E[L \mid r^* = r]$, $2 \le r \le k-1$ are given by the following theorem.

**Theorem 9.** *For DEBA and the PIAC model,*

$$E[L \mid r^* = r] \le E[L \mid r^* = r]_{\mathrm{iub}} = \max_{r+1 \le t \le k} \frac{V(r,t)}{t}, \ \ 2 \le r \le k-1\,,$$

*where*

$$
\begin{aligned}
V(r,t) \ = \ & \frac{(1-p)^m}{P(r)} \left( P[\mathcal{A}_r^l \mid |G|=0] - P[\mathcal{A}_{r+1}^l \mid |G|=0] \right) E[Z_{l,0}(r,t)] \\
& + \sum_{g=1}^{m-1} \frac{\binom{m}{g} p^g (1-p)^{m-g}}{P(r)} \\
& \left[ P[\mathcal{A}_r^l \mid |G|=g] E[Z_{l,g}(r,t)] + P[\mathcal{A}_r^h \mid |G|=g] E[Z_{h,g}(r,t)] \right. \\
& + \left( P[\mathcal{A}_r^{l,h} \mid |G|=g] - P[\mathcal{A}_{r+1}^{l,h} \mid |G|=g] - P[\mathcal{A}_{r+1}^l \mid |G|=g] \right. \\
& \left. - P[\mathcal{A}_{r+1}^h \mid |G|=g] \right) \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& + \min\left\{ P[\mathcal{A}_r^l \mid |G|=g], P[\mathcal{A}_{r+1}^l \mid |G|=g] \right\} \\
& \left( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{l,g}(r,t)] \right) \\
& + \min\left\{ P[\mathcal{A}_r^h \mid |G|=g], P[\mathcal{A}_{r+1}^h \mid |G|=g] \right\} \\
& \left. \left( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{h,g}(r,t)] \right) \right] \\
& + \frac{p^m}{P(r)} \left( P[\mathcal{A}_r^h \mid |G|=m] - P[\mathcal{A}_{r+1}^h \mid |G|=m] \right) E[Z_{h,m}(r,t)]\,,
\end{aligned}
$$

*with the $Z_{l,g}(r,t)$ and $Z_{h,g}(r,t)$ defined in Proposition 2. Furthermore, $E[L|r^* = r]_{\mathrm{iub}} \le (k-r)/k$, $2 \le r \le k-1$.*

*Proof.* Let $c$ be the alternative chosen by *DEBA*. From the proof of Theorem 5,

$$E[L \mid r^* = r] \le \max_{r+1 \le t \le k} \frac{V(r,t)}{t}\,,$$

where $V(r,t)$ is any upper bound for $E[S(r,t) \mid r^* = r]$.

It remains to derive $V(r,t)$. To that end, we first consider the refinement of the event $\{r^* = r\}$ into the collection of disjoint non-empty events $\{|G|=g\} \cap \{r^* = r\} \cap \mathcal{A}_r^l$, $0 \le g \le m-1$, $\{|G|=g\} \cap \{r^* = r\} \cap \mathcal{A}_r^h$, $1 \le g \le m$, and $\{|G|=g\} \cap \{r^* = r\} \cap \mathcal{A}_r^{l,h}$, $1 \le g \le m-1$. Using the definition of conditional expectation, we can express $E[S(r,t) \mid r^* = r]$ in terms of $E[S(r,t) \mid |G|=g \wedge r^* = r \wedge \mathcal{A}_r^l]$, $0 \le g \le m-1$, $E[S(r,t) \mid |G|=g \wedge r^* = r \wedge \mathcal{A}_r^h]$,

18

$1 \leq g \leq m$, and $E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}]$, $1 \leq g \leq m-1$ as

$$
\begin{aligned}
P(r)&E[S(r,t) \mid r^* = r] \\
&= \ P[|G| = 0 \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \ E[S(r,t) \mid |G| = 0 \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \\
&\quad + \sum_{g=1}^{m-1} \Bigg( P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \\
&\qquad\qquad + P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] \\
&\qquad\qquad + P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}] \Bigg) \\
&\quad + P[|G| = m \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] \ E[S(r,t) \mid |G| = m \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] .
\end{aligned}
$$

Using $P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] = P[|G| = g]P[r^* = r \ \wedge \ \mathcal{A}_r^l \mid |G| = g]$, $P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] = P[|G| = g]P[r^* = r \ \wedge \ \mathcal{A}_r^h \mid |G| = g]$, $P[|G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}] = P[|G| = g]P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \mid |G| = g]$, and $P[|G| = g] = \binom{m}{g}p^g(1-p)^{m-g}$, we get

$$
\begin{aligned}
E[S(r,t) &\mid r^* = r] \\
&= \ \frac{(1-p)^m}{P(r)} \ P[r^* = r \ \wedge \ \mathcal{A}_r^l \mid |G| = 0] \ E[S(r,t) \mid |G| = 0 \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \\
&\quad + \sum_{g=1}^{m-1} \frac{\binom{m}{g}p^g(1-p)^{m-g}}{P(r)} \\
&\qquad \Bigg( P[r^* = r \ \wedge \ \mathcal{A}_r^l \mid |G| = g] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^l] \\
&\qquad\quad + P[r^* = r \ \wedge \ \mathcal{A}_r^h \mid |G| = g] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] \\
&\qquad\quad + P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \mid |G| = g] \ E[S(r,t) \mid |G| = g \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^{l,h}] \Bigg) \\
&\quad + \frac{p^m}{P(r)} \ P[r^* = r \ \wedge \ \mathcal{A}_r^h \mid |G| = m] \ E[S(r,t) \mid |G| = m \ \wedge \ r^* = r \ \wedge \ \mathcal{A}_r^h] ,
\end{aligned}
$$

and using Proposition 2,

$$
\begin{aligned}
E[S(r,t) &\mid r^* = r] \\
&\leq \ \frac{(1-p)^m}{P(r)} \ P[r^* = r \ \wedge \ \mathcal{A}_r^l \mid |G| = 0] \ E[Z_{l,0}(r,t)] \\
&\quad + \sum_{g=1}^{m-1} \frac{\binom{m}{g}p^g(1-p)^{m-g}}{P(r)} \\
&\qquad \Bigg( P[r^* = r \ \wedge \ \mathcal{A}_r^l \mid |G| = g] \ E[Z_{l,g}(r,t)] \\
&\qquad\quad + P[r^* = r \ \wedge \ \mathcal{A}_r^h \mid |G| = g] \ E[Z_{h,g}(r,t)] \\
&\qquad\quad + P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \mid |G| = g] \ \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \Bigg) \\
&\quad + \frac{p^m}{P(r)} \ P[r^* = r \ \wedge \ \mathcal{A}_r^h \mid |G| = m] \ E[Z_{h,m}(r,t)] , \quad\quad (8)
\end{aligned}
$$

For $|G| = 0$, the event $\{r^* = r\} \cap \mathcal{A}_r^l$ is identical to the event $\mathcal{A}_r^l - \mathcal{A}_{r+1}^l$ and $\mathcal{A}_{r+1}^l \subset \mathcal{A}_r^l$, and we get

$$P[r^* = r \ \wedge \ \mathcal{A}_r^l \,|\, |G| = 0] = P[\mathcal{A}_r^l \,|\, |G| = 0] - P[\mathcal{A}_{r+1}^l \,|\, |G| = 0] \,. \tag{9}$$

Similarly, for $|G| = m$, the event $\{r^* = r\} \cap \mathcal{A}_r^h$ is identical to the event $\mathcal{A}_r^h - \mathcal{A}_{r+1}^h$ and $\mathcal{A}_{r+1}^h \subset \mathcal{A}_{r+1}^h$, and we get

$$P[r^* = r \ \wedge \ \mathcal{A}_r^h \,|\, |G| = m] = P[\mathcal{A}_r^h \,|\, |G| = m] - P[\mathcal{A}_{r+1}^h \,|\, |G| = m] \,. \tag{10}$$

No such simple relationships exist for $P[r^* = r \ \wedge \ \mathcal{A}_r^l \,|\, |G| = g]$, $P[r^* = r \ \wedge \ \mathcal{A}_r^h \,|\, |G| = g]$, $P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \,|\, |G| = g]$, $1 \leq g \leq m - 1$, and we will content ourselves by deriving an upper bound for

$$\begin{aligned}
T(g, r, t) \ = \ & P[r^* = r \ \wedge \ \mathcal{A}_r^l \,|\, |G| = g] \, E[Z_{l,g}(r, t)] \\
& + P[r^* = r \ \wedge \ \mathcal{A}_r^h \,|\, |G| = g] \, E[Z_{h,g}(r, t)] \\
& + P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \,|\, |G| = g] \, \max\{E[Z_{l,g}(r, t)], E[Z_{h,g}(r, t)]\} \,, \\
& \qquad\qquad 1 \leq g \leq m - 1 \,. \tag{11}
\end{aligned}$$

Clearly, from the definitions of $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, and $\mathcal{A}_r^{l,h}$,

$$\begin{aligned}
\{r^* = r\} \cap \mathcal{A}_r^l \ &= \ \mathcal{A}_r^l - \left( \mathcal{A}_{r+1}^l \cup \mathcal{A}_{r+1}^h \cup \mathcal{A}_{r+1}^{l,h} \right) , \\
\{r^* = r\} \cap \mathcal{A}_r^h \ &= \ \mathcal{A}_r^h - \left( \mathcal{A}_{r+1}^l \cup \mathcal{A}_{r+1}^h \cup \mathcal{A}_{r+1}^{l,h} \right) , \tag{12} \\
\{r^* = r\} \cap \mathcal{A}_r^{l,h} \ &= \ \mathcal{A}_r^{l,h} - \left( \mathcal{A}_{r+1}^l \cup \mathcal{A}_{r+1}^h \cup \mathcal{A}_{r+1}^{l,h} \right) .
\end{aligned}$$

Also, $\mathcal{A}_{r+1}^l \subset \mathcal{A}_r^l \cup \mathcal{A}_r^{l,h}$, $\mathcal{A}_{r+1}^h \subset \mathcal{A}_r^h \cup \mathcal{A}_r^{l,h}$, and $\mathcal{A}_{r+1}^{l,h} \subset \mathcal{A}_r^{l,h}$. Define $\mathcal{A}_{r+1}^{l\prime} = \mathcal{A}_{r+1}^l \cap \mathcal{A}_r^l$, $\mathcal{A}_{r+1}^{l\prime\prime} = \mathcal{A}_{r+1}^l \cap \mathcal{A}_r^{l,h}$, $\mathcal{A}_{r+1}^{h\prime} = \mathcal{A}_{r+1}^h \cap \mathcal{A}_r^h$, and $\mathcal{A}_{r+1}^{h\prime\prime} = \mathcal{A}_{r+1}^h \cap \mathcal{A}_r^{l,h}$. The Venn diagram of Fig. 2 shows the relationships between the events $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, $\mathcal{A}_r^{l,h}$, $\mathcal{A}_{r+1}^l$, $\mathcal{A}_{r+1}^{l\prime}$, $\mathcal{A}_{r+1}^{l\prime\prime}$, $\mathcal{A}_{r+1}^h$, $\mathcal{A}_{r+1}^{h\prime}$, $\mathcal{A}_{r+1}^{h\prime\prime}$, and $\mathcal{A}_{r+1}^{l,h}$. Using them in (12),

$$\begin{aligned}
\{r^* = r\} \cap \mathcal{A}_r^l \ &= \ \mathcal{A}_r^l - \mathcal{A}_{r+1}^{l\prime}, \\
\{r^* = r\} \cap \mathcal{A}_r^h \ &= \ \mathcal{A}_r^h - \mathcal{A}_{r+1}^{h\prime}, \\
\{r^* = r\} \cap \mathcal{A}_r^{l,h} \ &= \ \mathcal{A}_r^{l,h} - \left( \mathcal{A}_{r+1}^{l\prime\prime} \cup \mathcal{A}_{r+1}^{h\prime\prime} \cup \mathcal{A}_{r+1}^{l,h} \right) ,
\end{aligned}$$

and using $\mathcal{A}_{r+1}^{l\prime} \subset \mathcal{A}_r^l$, $\mathcal{A}_{r+1}^{h\prime} \subset \mathcal{A}_r^h$, $\mathcal{A}_{r+1}^{l\prime\prime} \cup \mathcal{A}_{r+1}^{h\prime\prime} \cup \mathcal{A}_{r+1}^{l,h} \subset \mathcal{A}_r^{l,h}$, and the fact that $\mathcal{A}_{r+1}^{l\prime\prime}$, $\mathcal{A}_{r+1}^{h\prime\prime}$, and $\mathcal{A}_{r+1}^{l,h}$ are disjoint,

$$P[r^* = r \ \wedge \ \mathcal{A}_r^l \,|\, |G| = g] = P[\mathcal{A}_r^l \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g] \,,$$

$$P[r^* = r \ \wedge \ \mathcal{A}_r^h \,|\, |G| = g] = P[\mathcal{A}_r^h \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g] \,,$$

$$\begin{aligned}
P[r^* = r \ \wedge \ \mathcal{A}_r^{l,h} \,|\, |G| = g] &= P[\mathcal{A}_r^{l,h} \,|\, |G| = g] \\
&\quad - P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g] \,,
\end{aligned}$$

which used in (11) give

$$
\begin{aligned}
T(g, r, t) \;=\;& P[\mathcal{A}_r^l \,|\, |G| = g]\, E[Z_{l,g}(r,t)] + P[\mathcal{A}_r^h \,|\, |G| = g] E[Z_{h,g}(r,t)] \\
& + \left( P[\mathcal{A}_r^{l,h} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g] \right) \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& - P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g]\, E[Z_{l,g}(r,t)] - P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g]\, E[Z_{h,g}(r,t)] \\
& - P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g] \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& - P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g] \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& \qquad\qquad\qquad 1 \le g \le m-1 \,.
\end{aligned}
$$

The upper bound for $T(g,r,t)$, $1 \le g \le m-1$, is obtained by finding the maximum of $T(g,r,t)$ as a function of $P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g]$, $P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g]$, $P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g]$, and $P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g]$, subject to the constraints that those variables are known to satisfy. The constraints are (see Fig. 2):

$$
P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g] + P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g] = P[\mathcal{A}_{r+1}^{l} \,|\, |G| = g]\,,
$$

$$
P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g] + P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g] = P[\mathcal{A}_{r+1}^{h} \,|\, |G| = g]\,,
$$

$$
0 \le P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r}^{l} \,|\, |G| = g]\,,
$$

$$
0 \le P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r}^{h} \,|\, |G| = g]\,,
$$

$$
P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g] \ge 0\,,
$$

$$
P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g] \ge 0\,,
$$

$$
P[\mathcal{A}_{r+1}^{l\prime\prime} \,|\, |G| = g] + P[\mathcal{A}_{r+1}^{h\prime\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r}^{l,h} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g]\,.
$$

This is equivalent to find the maximum of

$$
\begin{aligned}
T(g, r, t) \;=\;& P[\mathcal{A}_r^l \,|\, |G| = g]\, E[Z_{l,g}(r,t)] + P[\mathcal{A}_r^h \,|\, |G| = g]\, E[Z_{h,g}(r,t)] \\
& + \bigg( P[\mathcal{A}_r^{l,h} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \,|\, |G| = g] - P[\mathcal{A}_{r+1}^{l} \,|\, |G| = g] \\
& \quad - P[\mathcal{A}_{r+1}^{h} \,|\, |G| = g] \bigg) \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& + P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g]\bigg( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{l,g}(r,t)] \bigg) \\
& + P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g]\bigg( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{h,g}(r,t)] \bigg)\,, \\
& \qquad\qquad\qquad 1 \le g \le m-1
\end{aligned}
$$

as a function of $P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g]$ and $P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g]$ subject to the constraints

$$
0 \le P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r}^{l} \,|\, |G| = g]\,,
$$

$$
0 \le P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r}^{h} \,|\, |G| = g]\,,
$$

$$
P[\mathcal{A}_{r+1}^{l\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r+1}^{l} \,|\, |G| = g]\,,
$$

$$
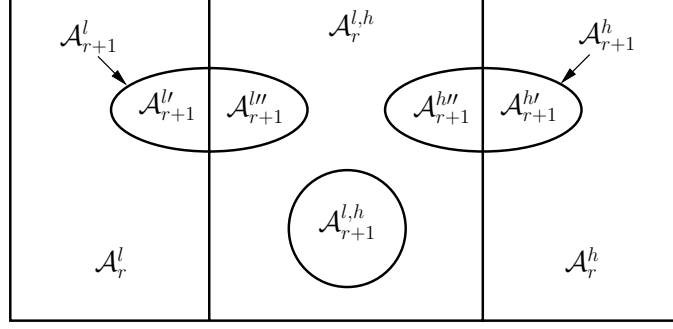P[\mathcal{A}_{r+1}^{h\prime} \,|\, |G| = g] \le P[\mathcal{A}_{r+1}^{h} \,|\, |G| = g]\,,
$$

Figure 2: Venn diagram showing the relationships between the events $\mathcal{A}_r^l$, $\mathcal{A}_r^h$, $\mathcal{A}_r^{l,h}$, $\mathcal{A}_{r+1}^l$, $\mathcal{A}_{r+1}^{l\prime}$, $\mathcal{A}_{r+1}^{l\prime\prime}$, $\mathcal{A}_{r+1}^h$, $\mathcal{A}_{r+1}^{h\prime}$, $\mathcal{A}_{r+1}^{h\prime\prime}$, and $\mathcal{A}_{r+1}^{l,h}$.

$$P[\mathcal{A}_{r+1}^{l\prime} \mid |G| = g] + P[\mathcal{A}_{r+1}^{h\prime} \mid |G| = g]$$
$$\geq P[\mathcal{A}_{r+1}^l \mid |G| = g] + P[\mathcal{A}_{r+1}^h \mid |G| = g] - P[\mathcal{A}_r^{l,h} \mid |G| = g] + P[\mathcal{A}_{r+1}^{l,h} \mid |G| = g].$$

Since the function is non-decreasing on both $P[\mathcal{A}_{r+1}^{l\prime} \mid |G| = g]$ and $P[\mathcal{A}_{r+1}^{h\prime} \mid |G| = g]$, a point $(a, b)$ at which $T(g, r, t)$ achieves its maximum is

$$a = \min\{P[\mathcal{A}_r^l \mid |G| = g], P[\mathcal{A}_{r+1}^l \mid |G| = g]\}$$

$$b = \min\{P[\mathcal{A}_r^h \mid |G| = g], P[\mathcal{A}_{r+1}^h \mid |G| = g]\}$$

yielding (11)

$$
\begin{aligned}
& P[r^* = r \ \wedge\ \mathcal{A}_r^l \mid |G| = g]\, E[Z_{l,g}(r,t)] + P[r^* = r \ \wedge\ \mathcal{A}_r^h \mid |G| = g]\, E[Z_{h,g}(r,t)] \\
& + P[r^* = r \ \wedge\ \mathcal{A}_r^{l,h} \mid |G| = g] \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& \qquad \leq\ P[\mathcal{A}_r^l \mid |G| = g]\, E[Z_{l,g}(r,t)] + P[\mathcal{A}_r^h \mid |G| = g]\, E[Z_{h,g}(r,t)] \\
& \qquad\quad + \Big( P[\mathcal{A}_r^{l,h} \mid |G| = g] - P[\mathcal{A}_{r+1}^{l,h} \mid |G| = g] - P[\mathcal{A}_{r+1}^l \mid |G| = g] \\
& \qquad\qquad - P[\mathcal{A}_{r+1}^h \mid |G| = g] \Big) \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} \\
& \qquad\quad + \min\{P[\mathcal{A}_r^l \mid |G| = g], P[\mathcal{A}_{r+1}^l \mid |G| = g]\} \\
& \qquad\qquad \Big( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{l,g}(r,t)] \Big) \\
& \qquad\quad + \min\{P[\mathcal{A}_r^h \mid |G| = g], P[\mathcal{A}_{r+1}^h \mid |G| = g]\} \\
& \qquad\qquad \Big( \max\{E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)]\} - E[Z_{h,g}(r,t)] \Big), \\
& \qquad\qquad\qquad\qquad 1 \leq g \leq m - 1.
\end{aligned}
\tag{13}
$$

The upper bound $V(r, t)$ for $E[S(r, t) \mid r^* = r]$ is obtained by combining (8), (9), (10), and (13).

It remains to show that $E[L \mid r^* = r]_{\text{iub}} \leq (k - r)/k$, $2 \leq r \leq k - 1$. Since $Z_{l,g}(r,t), Z_{h,g}(r,t) \leq t - r$, we have $E[Z_{l,g}(r,t)], E[Z_{h,g}(r,t)] \leq t - r$. Furthermore, careful analysis of the procedure followed to derive $V(r, t)$ reveals that $V(r, t)$ is given by the right-hand side of (8) with $P[r^* = r \ \wedge\ \mathcal{A}_r^l \mid |G| = g]$, $P[r^* = r \ \wedge\ \mathcal{A}_r^h \mid |G| = g]$, and $P[r^* = r \ \wedge\ \mathcal{A}_r^{l,h} \mid |G| = g]$,

$1 \leq g \leq m - 1$ replaced by quantities $R_g^l(r)$, $R_g^h(r)$, and $R_g^{l,h}(r)$, $1 \leq g \leq m - 1$ satisfying $R_g^l(r), R_g^h(r), R_g^{l,h}(r) \geq 0$ and $R_g^l(r) + R_g^h(r) + R_g^{l,h}(r) = P[r^* = r \mid |G| = g]$. Then, we have

$$
\frac{V(r,t)}{t-r} \;\leq\; \frac{1}{P(r)}\Bigg[(1-p)^m \, P[r^* = r \,\wedge\, \mathcal{A}_r^l \mid |G| = 0]
$$

$$
+ \sum_{g=1}^{m-1} \binom{m}{g} p^g \, (1-p)^{m-g} P[r^* = r \mid |G| = g]
$$

$$
+ p^m \, P[r^* = r \,\wedge\, \mathcal{A}_r^h \mid |G| = m]\Bigg],
$$

and taking into account that, for $|G| = 0$, $\{r^* = r\} \subset \mathcal{A}_r^l$ and, for $|G| = m$, $\{r^* = r\} \subset \mathcal{A}_r^h$, and using $P[|G| = g] = \binom{m}{g} p^g (1-p)^{m-g}$,

$$
\frac{V(r,t)}{t-r} \leq \frac{1}{P(r)} \sum_{g=0}^{m} P[|G| = g] P[r^* = r \mid |G| = g] = \frac{P[r^* = r]}{P(r)} = 1 \,.
$$

Then, $E[L \mid r^* = r]_{\text{iub}} = \max_{r+1 \leq t \leq k} V(r,t)/t \leq \max_{r+1 \leq t \leq k}(t-r)/t = (k-r)/k$. $\qquad\square$

It remains to discuss the computation of the $E[Z_{l,g}(r,t)]$'s and $E[Z_{h,g}(r,t)]$'s involved in the improved upper bounds for $E[L|r^* = r]$, $2 \leq r \leq k-1$ given by Theorem 9. They can be computed using a recurrence-based computational scheme similar to the one given by Theorem 6 to compute the $E[Z(r,t)]$'s involved in the improved upper bounds for the ZIAC model. The recurrence-based computational scheme is given by the following theorem.

**Theorem 10.** *For the PIAC model, $E[Z_{l,g}(r,t)]$, $0 \leq g \leq m-1$, $2 \leq r \leq k-1$, $r+1 \leq t \leq k$ and $E[Z_{h,g}(r,t)]$, $1 \leq g \leq m$, $2 \leq r \leq k-1$, $r+1 \leq t \leq k$ can be computed using*

$$
E[Z_{l,g}(r,t)] = W_{l,g}(t-r) \,,
$$

$$
E[Z_{h,g}(r,t)] = W_{h,g}(t-r) \,,
$$

$$
W_{l,g}(u) = \sum_{a=1}^{u} a\,\xi_{l,g}(u,a) \,, \quad 0 \leq g \leq m-1 \,, 1 \leq u \leq k-2 \,,
$$

$$
W_{h,g}(u) = \sum_{a=1}^{u} a\,\xi_{h,g}(u,a) \,, \quad 1 \leq g \leq m \,, 1 \leq u \leq k-2 \,,
$$

*and the recurrences:*

$$
\xi_{l,g}(u,a) = \sum_{a'=a}^{u} \psi(m-g-1,u,a')\,\pi^l(u,a'-a) \,, \quad 0 \leq g \leq m-1 \,, 1 \leq u \leq k-2 \,, 1 \leq a \leq u \,,
$$

$$
\xi_{h,g}(u,a) = \sum_{a'=a}^{u} \psi(m-g,u,a')\,\pi^h(u,a'-a) \,, \quad 1 \leq g \leq m \,, 1 \leq u \leq k-2 \,, 1 \leq a \leq u \,,
$$

$$
\psi(0,u,a) = \phi^h(m-1,u,a) \,, \quad 1 \leq u \leq k-2 \,, 1 \leq a \leq u \,,
$$

$$
\psi(m-1,u,a) = \phi^l(m-1,u,a) \,, \quad 1 \leq u \leq k-2 \,, 1 \leq a \leq u \,,
$$

$$\psi(b,u,a) = \phi^l(b,u,a) \sum_{a'=1}^{a} \phi^h(m-b-1,u,a') + \phi^h(m-b-1,u,a) \sum_{a'=1}^{a-1} \phi^l(b,u,a'),$$

$$1 \le b \le m-2\,,\ 1 \le u \le k-2\,,\ 1 \le a \le u\,,$$

$$\phi^l(1,u,a) = \pi^l(u,a-1)\,,\quad 1 \le u \le k-2\,,\ 1 \le a \le u\,,$$

$$\phi^l(b,u,a) = \phi^l(b-1,u,a) \sum_{a'=0}^{a-1} \pi^l(u,a') + \pi^l(u,a-1) \sum_{a'=1}^{a-1} \phi^l(b-1,u,a')\,,$$

$$2 \le b \le m-1\,,\ 1 \le u \le k-2\,,\ 1 \le a \le u\,,$$

$$\phi^h(1,u,a) = \pi^h(u,a-1)\,,\quad 1 \le u \le k-2\,,\ 1 \le a \le u\,,$$

$$\phi^h(b,u,a) = \phi^h(b-1,u,a) \sum_{a'=0}^{a-1} \pi^h(u,a') + \pi^h(u,a-1) \sum_{a'=1}^{a-1} \phi^h(b-1,u,a')\,,$$

$$2 \le b \le m-1\,,\ 1 \le u \le k-2\,,\ 1 \le a \le u\,,$$

$$\pi^l(1,0) = 1\,,$$

$$\pi^l(2,0) = 1 - p^l\,,$$

$$\pi^l(2,1) = p^l\,,$$

$$\pi^l(u,u-1) = p^l\,\pi^l(u-1,u-2)\,,\quad 3 \le u \le k-2\,,$$

$$\pi^l(u,a) = (1-p^l)\pi^l(u-1,a) + p^l\,\pi^l(u-1,a-1)\,,\quad 3 \le u \le k-2\,,\ 1 \le a \le u-2\,,$$

$$\pi^l(u,0) = (1-p^l)\pi^l(u-1,0)\,,\quad 3 \le u \le k-2\,.$$

$$\pi^h(1,0) = 1\,,$$

$$\pi^h(2,0) = 1 - p^h\,,$$

$$\pi^h(2,1) = p^h\,,$$

$$\pi^h(u,u-1) = p^h\,\pi^h(u-1,u-2)\,,\quad 3 \le u \le k-2\,,$$

$$\pi^h(u,a) = (1-p^h)\pi^h(u-1,a) + p^h\,\pi^h(u-1,a-1)\,,\quad 3 \le u \le k-2\,,\ 1 \le a \le u-2\,,$$

$$\pi^h(u,0) = (1-p^h)\pi^h(u-1,0)\,,\quad 3 \le u \le k-2\,.$$

*Proof.* See the Appendix. $\qquad\square$

# 5  Analysis

In this section we will illustrate the improved upper bounds for the expected loss of *DEBA* obtained in this paper, will compare them with those obtained in BCH, and will analyze their tightness. We will also discuss how the improved upper bounds allow us to extend the conclusions in BCH regarding the performance of *DEBA*.

To illustrate the improved upper bounds for the expected loss of *DEBA* and compare them with those obtained in BCH, we will set as the base case scenario the ZIAC model with $p_r = 0.5$. We will then consider the following six variations. First, to examine the effect of average attribute quality, we will change the probability level $p_r$ in the ZIAC model to the values $p_r = 0.2$ and $p_r = 0.8$. Second, to observe the effect of varying $p_r$, we will consider a *low-to-high* pattern in which $p_r$ increases linearly from $p_1 = 0.2$ to $p_k = 0.8$, i.e. $p_r = 0.2 + 0.6(r-1)/(k-1)$, and a *high-to-low* pattern in which $p_r$ decreases linearly with $r$ from $p_1 = 0.8$ to $p_k = 0.2$, i.e. $p_r = 0.8 - 0.6(r-1)/(k-1)$. The average of the $p_r$'s is 0.5 in both cases. Third, to examine the effect of positive correlation, we will set $p = 0.5$ in the PIAC model and will explore the correlation levels $\rho = 0.2$ and $\rho = 0.5$. For these seven scenarios, Fig. 3 plots the upper bounds for the expected loss of *DEBA*, $E[L]_{\mathrm{ub}}$, obtained in BCH for $2 \leq m \leq 10$ and $3 \leq k \leq 10$. Fig. 4 plots the improved upper bounds, $E[L]_{\mathrm{iub}}$, obtained in this paper. We can first note that the improved upper bounds are substantially better than those obtained in BCH: for large $k$ they are about half the previously obtained upper bounds. Another important point is that the improved upper bounds seem to increase relatively less rapidly with $k$ than the upper bounds obtained in BCH, implying that the comparison with the upper bounds obtained in BCH would probably get better for larger values of $k$. The behavior of $E[L]_{\mathrm{iub}}$ with respect to the average quality of the attributes, a varying average attribute quality, and the presence of inter-attribute positive correlation is similar to the behavior with respect to those characteristics of $E[L]_{\mathrm{ub}}$: $E[L]_{\mathrm{iub}}$ decreases when the average quality of the attributes increases beyond 0.5, when there is a high-to-low pattern in the average attribute qualities, and in the presence of increasing positive inter-attribute correlation. The behavior of $E[L]_{\mathrm{iub}}$ with respect to $k$ and $m$ is also similar to the behavior of $E[L]_{\mathrm{ub}}$. First, for fixed $m$ and increasing $k$, $E[L]_{\mathrm{iub}}$ increases. This is because in (2) the $P(r)$ are independent of $k$ and $E[L \mid r^* = r]_{\mathrm{iub}}$ is non-decreasing with $k$ (see Theorems 5 and 9). Second, for fixed $k$, $E[L]_{\mathrm{iub}} \rightarrow 0$ as $m \rightarrow \infty$. This is because, for fixed $k$, for both the ZIAC and the PIAC probabilistic models, the probability that there will be some all-1 alternative goes to 1 as $m \rightarrow \infty$ (see BCH), implying $P(k) \rightarrow 1$, which by (2) implies $E[L]_{\mathrm{iub}} \rightarrow 0$. As for $E[L]_{\mathrm{ub}}$, for fixed $k$, there seems to exist a turning point for $m$, $m^*$, before which $E[L]_{\mathrm{iub}}$ increases with $m$ and beyond which $E[L]_{\mathrm{iub}}$ decreases with $m$. The turning point $m^*$ seems to decrease as $k$ decreases, as the average quality of the attributes increases, and in the presence of increasing positive inter-attribute correlation.

Before analyzing the tightness of $E[L]_{\mathrm{iub}}$, we will compare $E[L \mid r^* = r]_{\mathrm{iub}}$, $2 \leq r \leq k-1$ with the upper bounds, $E[L \mid r^* = r]_{\mathrm{ub}} = (k-r)/k$, used in BCH and analyze the tightness of the former. This will allow us to discuss till what extent the upper bounds obtained in this paper can be further improved. The event $r^* = r$, $2 \leq r \leq k-1$ implies the existence of some alternative $i$
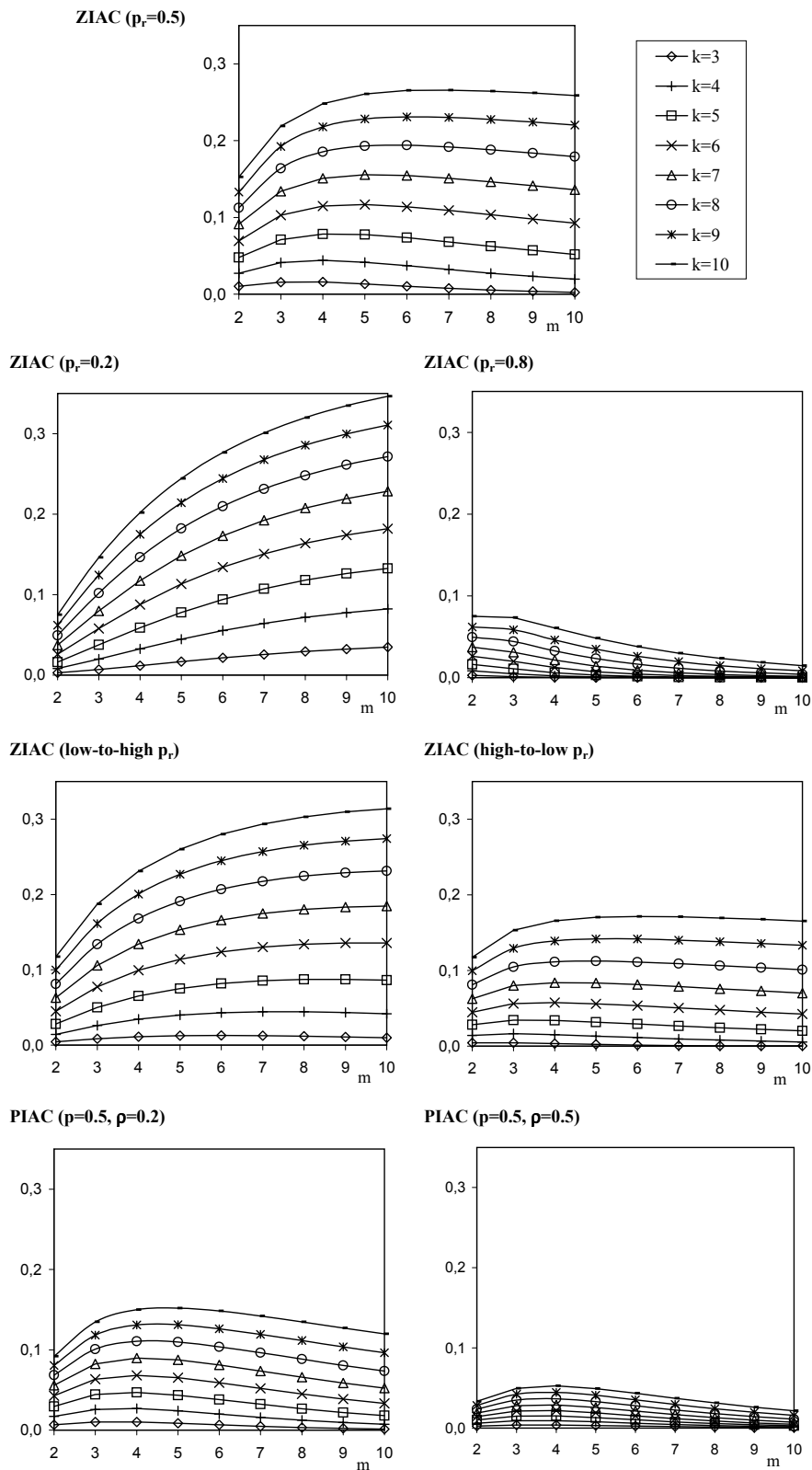
Figure 3: $E[L]_{\mathrm{ub}}$ as a function of $m$ and $k$ for seven scenarios (base case and six variations).
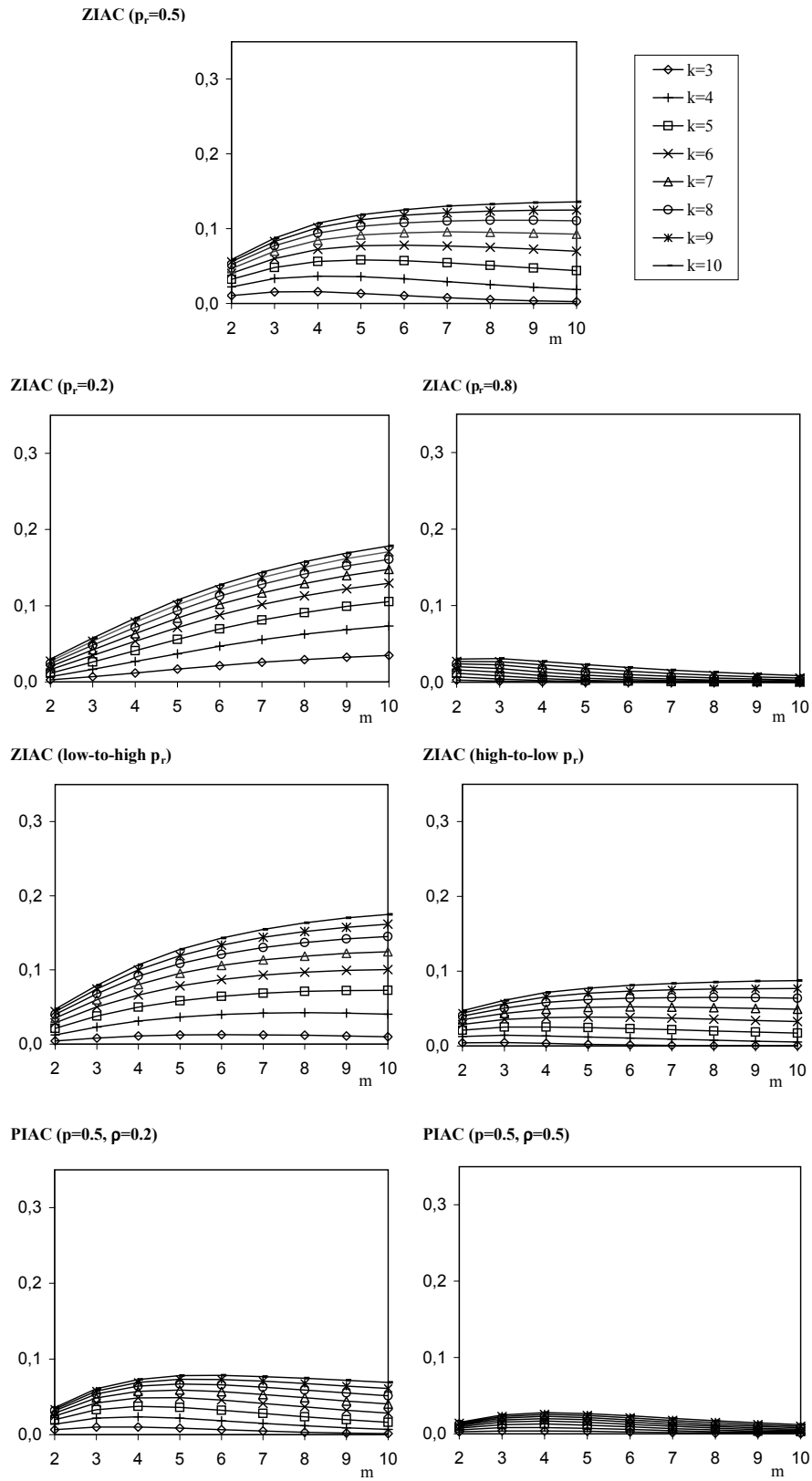
26

Figure 4: $E[L]_{\text{iub}}$ as a function of $m$ and $k$ for seven scenarios (base case and six variations).

different from the alternative $c$ chosen by *DEBA* with $X_{i,r} = X_{c,r}$ and $X_{i,r+1} = X_{c,r+1} + 1$. The feasible set of weights $w_1 = w_2 = \cdots = w_{r+1} = 1/(r+1)$, $w_{r+2} = w_{r+3} = \cdots = w_k = 0$ give to alternative $i$ a utility $U_i$ that is $1/(r+1)$ larger than the utility $U_c$ of the alternative chosen by *DEBA*. This implies that any upper bound for $E[L \mid r^* = r]$ independent on the precise values of the weights cannot be smaller than $1/(r+1)$. In summary, $1/(r+1)$ is a lower bound on any feasible upper bound for $E[L \mid r^* = r]$. For the ZIAC model, Fig. 5 compares $E[L \mid r^* = r]_{\text{iub}}$ with $1/(r+1)$, and with $E[L \mid r^* = r]_{\text{ub}} = (k-r)/k$. Fig. 6 performs the same comparison for the PIAC model. First, notice that both figures show a significant difference between $E[L | r^* = r]_{\text{ub}}$ and $1/(r+1)$, suggesting room for significant improvement of the upper bounds for $E[L \mid r^* = r]$ used in BCH. That observation was the motivation for our work. For the ZIAC model, $E[L \mid r^* = r]_{\text{iub}}$ are close to $1/(r+1)$ in all cases, indicating that they cannot be further improved significantly. For the PIAC model, however, $E[L \mid r^* = r]_{\text{iub}}$ are not close to $1/(r+1)$ for small $r$. Thus, in principle, there could exist room for significant improvement of the upper bounds $E[L \mid r^* = r]_{\text{iub}}$ for the PIAC model. However, we doubt that this is actually the case. The reason is that there exists an intuitive explanation for the departure from $1/(r+1)$: for small values of $r$ there is a non-negligible probability that, conditioned on $r^* = r$, the alternatives in $C_{r^*}$ are from the "bad" population and, in that scenario, with appropriate sets of feasible weights, the conditional expected loss of *DEBA* can be high compared to $1/(r+1)$.

We next analyze the tightness of the improved upper bounds for the expected loss of *DEBA* obtained in this paper. To that end, we will compare $E[L]_{\text{iub}}$ with the maximum value of $E[L]$ over all feasible weights, $E[L]_{\text{max}}$. Exact computation of $E[L]_{\text{max}}$ does not seem feasible: one has first to derive an exact procedure for the computation of $E[L]$ for any given set of feasible weights, and then take the maximum over all feasible sets of weights. To estimate $E[L]_{\text{max}}$ we used the following, expensive approach. First, we sampled 100,000 realizations of sets of alternatives. Second, using those samples, we obtained point estimates for $E[L]$ for 100,000 uniformly sampled sets of feasible weights, and took the maximum of those point estimates to estimate $E[L]_{\text{max}}$. Table 1 compares $E[L]_{\text{iub}}$ with the estimates for $E[L]_{\text{max}}$ for the ZIAC model. Table 2 performs the comparison for the PIAC model. For $E[L]_{\text{max}}$ we give the point estimate and its 90 % confidence interval. We also give the percent increase of $E[L]_{\text{iub}}$ with respect to the point estimate for $E[L]_{\text{max}}$. For the ZIAC model, we exclude the case $m = 10$, $k = 5$, $p_r = 0.8$ because in that case $E[L]_{\text{max}}$ was so small that the point estimate for $E[L]_{\text{max}}$ had a confidence interval of too poor a quality. In all cases, the upper bound $E[L]_{\text{iub}}$ is reasonably tight. Broadly speaking, the tightness of $E[L]_{\text{iub}}$ improves with the average quality of the attributes and in the presence of inter-attribute correlation and degrades as $k$ increases. Given those results and the discussed good tightness of $E[L \mid r^* = r]_{\text{iub}}$, it is unlikely that the upper bounds for the expected loss of *DEBA* obtained in this paper can be further improved significantly unless more sophisticated approaches than conditioning on $r^*$ are used.

The relevance of the improved upper bounds stems from the fact that they widen the identification of cases in which *DEBA* is guaranteed to have a good performance under the assumed cognitive limitations regarding the values of the weights of the utility function. More specifically, $E[L]_{\text{iub}}$ is reasonably small in all seven considered scenarios, indicating a reasonably good performance of *DEBA* for numbers of alternatives and attributes as large as 10 when attribute values are uncorre-
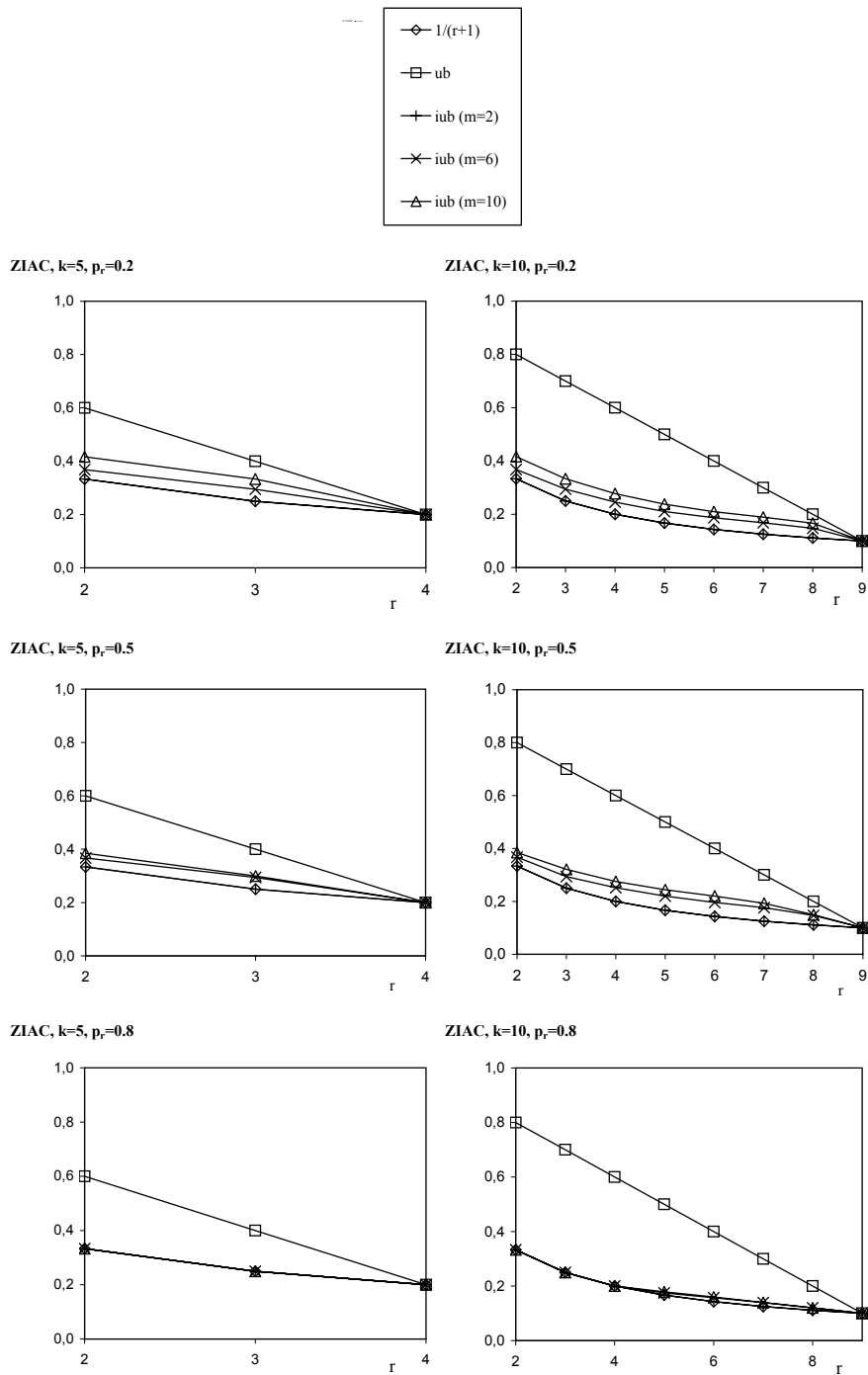
Figure 5: $1/(r + 1)$, $E[L \mid r^* = r]_{\mathrm{ub}}$ (ub), and $E[L \mid r^* = r]_{\mathrm{iub}}$ (iub) for the ZIAC probabilistic model.
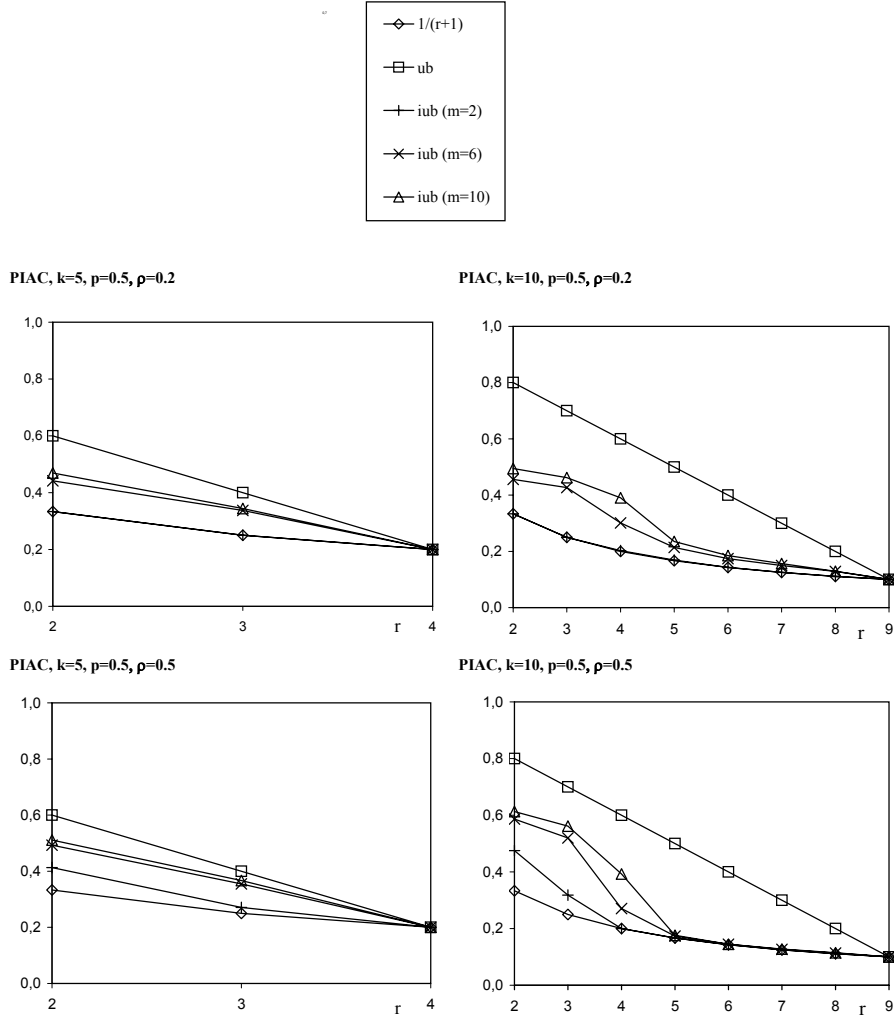
Figure 6: $1/(r + 1)$, $E[L \mid r^* = r]_{\mathrm{ub}}$ (ub), and $E[L \mid r^* = r]_{\mathrm{iub}}$ (iub) for the PIAC probabilistic model.

Table 1: Tightness of $E[L]_{\mathrm{iub}}$ for the ZIAC model.

| $m$ | $k$ | $p_r$ | $E[L]_{\max}$ | $E[L]_{\mathrm{iub}}$ | increase |
|---|---|---|---|---|---|
| 5 | 5 | 0.2 | $0.04181 \pm 0.00049$ | 0.05576 | 33 % |
| 5 | 5 | 0.5 | $0.04521 \pm 0.00052$ | 0.05836 | 29 % |
| 5 | 5 | 0.8 | $0.002420 \pm 0.000119$ | 0.002520 | 4 % |
| 5 | 10 | 0.2 | $0.06026 \pm 0.00046$ | 0.10741 | 78 % |
| 5 | 10 | 0.5 | $0.07381 \pm 0.00055$ | 0.11878 | 61 % |
| 5 | 10 | 0.8 | $0.01976 \pm 0.00027$ | 0.02303 | 17 % |
| 10 | 5 | 0.2 | $0.06495 \pm 0.00059$ | 0.10536 | 62 % |
| 10 | 5 | 0.5 | $0.03694 \pm 0.00045$ | 0.04394 | 19 % |
| 10 | 10 | 0.2 | $0.08479 \pm 0.00052$ | 0.17827 | 110 % |
| 10 | 10 | 0.5 | $0.08487 \pm 0.00056$ | 0.13581 | 60 % |
| 10 | 10 | 0.8 | $0.008324 \pm 0.000167$ | 0.008728 | 5 % |

Table 2: Tightness of $E[L]_{\mathrm{iub}}$ for the PIAC model.

| $m$ | $k$ | $p$ | $\rho$ | $E[L]_{\max}$ | $E[L]_{\mathrm{iub}}$ | increase |
|---|---|---|---|---|---|---|
| 5 | 5 | 0.5 | 0.2 | $0.02835 \pm 0.00046$ | 0.03612 | 27 % |
| 5 | 5 | 0.5 | 0.5 | $0.009502 \pm 0.000299$ | 0.011422 | 20 % |
| 5 | 10 | 0.5 | 0.2 | $0.05618 \pm 0.00060$ | 0.07788 | 39 % |
| 5 | 10 | 0.5 | 0.5 | $0.02103 \pm 0.00042$ | 0.02651 | 26 % |
| 10 | 5 | 0.5 | 0.2 | $0.01363 \pm 0.00030$ | 0.01635 | 20 % |
| 10 | 5 | 0.5 | 0.5 | $0.002130 \pm 0.000133$ | 0.002552 | 20 % |
| 10 | 10 | 0.5 | 0.2 | $0.04757 \pm 0.00049$ | 0.06899 | 45 % |
| 10 | 10 | 0.5 | 0.5 | $0.009768 \pm 0.000232$ | 0.012141 | 24 % |

lated or when there exists positive inter-attribute correlation irrespectively of the average quality of the attributes. Using the upper bounds obtained in BCH, a good performance of *DEBA* for up to 10 alternatives and 10 attributes is only predicted when the average quality of the attributes is high, when there exists positive inter-attribute correlation, and when, not being the average qualities of the attributes low, there is a high-to-low pattern in the average qualities of the attributes.

## 6 Conclusions

The *DEBA* heuristic is a popular and intuitively appealing heuristic for multi-attribute choice when weights are non-increasing which has been extensively studied in the literature (Bröeder 2000). For any heuristic, the expected loss is a reasonable measure for the performance, arguably better than the probability of making the best choice. In this paper, we have derived improved upper bounds for the expected loss of *DEBA* independent of the precise values of the weights for the binary attribute case and for two probabilistic models: one in which the attribute values are assumed to be independent Bernoulli random variables, and another one with positive inter-attribute correlation and attributes of the same average quality. It has been shown that the new upper bounds improve substantially the upper bounds derived in BCH. Furthermore, they have been shown to be tight. The new upper bounds extend significantly the identification of cases in which a good performance of *DEBA* can be guaranteed under the assumed cognitive limitations. In particular, a good performance of the heuristic is guaranteed for numbers of alternatives and numbers of attributes as large as 10 when attribute values are uncorrelated or when there is positive inter-attribute correlation. A low value in the upper bound for the expected loss of *DEBA* can be used to recommend use of the simple *DEBA* decision heuristic and disregard the use of more sophisticated and costly decision rules, including the obtention of more accurate estimates for the attribute weights.

An obvious direction for future research is the extension of the work presented in this paper to a probabilistic model for the attribute values allowing arbitrary average qualities for all attributes

and allowing any feasible inter-attribute correlation matrix. First attempts in that direction seem to indicate that the extension is a challenging one.

## Acknowledgements

## Appendix A

The proof of Theorem 4 is preceeded by a lemma and a proposition. The lemma establishes a result regarding conditional expected values of a non-decreasing function of the the differences of independent Bernouilli random variables and a common independent Bernouilli random variable.

**Lemma 1.** *Let $n \geq 2$. Let $F(u_1, \ldots, u_{n-1})$ be a function which is non-decreasing on each $u_i$. Let $v_i$, $1 \leq i \leq n$, be independent Bernoulli random variables with parameters $p_i$, $0 < p_i < 1$. Let $\mathcal{E}_\nu$ be the event $\{v_1 = 1 \vee v_1 = \cdots = v_\nu = 0\}$, $2 \leq \nu \leq n$. Then,*

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}_\nu] \leq E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \overline{\mathcal{E}_\nu}].$$

*Proof.* Let the events $\mathcal{E}^1 = \{v_1 = 1\}$, $\mathcal{E}_\nu^2 = \{v_1 = \cdots = v_\nu = 0\}$. Clearly, $\mathcal{E}^1 \cup \mathcal{E}_\nu^2 = \mathcal{E}_\nu$ and $\mathcal{E}^1 \cap \mathcal{E}_\nu^2 = \emptyset$. Because

$$
\begin{aligned}
&E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}_\nu] \\
&\quad = \frac{P[\mathcal{E}^1]\, E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}^1] + P[\mathcal{E}_\nu^2]\, E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}_\nu^2]}{P[\mathcal{E}^1] + P[\mathcal{E}_\nu^2]},
\end{aligned}
$$

it is enough to prove

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}^1] \leq E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \overline{\mathcal{E}_\nu}]$$

and

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \mathcal{E}_\nu^2] \leq E[F(v_2 - v_1, \ldots, v_n - v_1) \,|\, \overline{\mathcal{E}_\nu}].$$

Let $\mathcal{E}_\nu^3$ be the event $\{v_2 = \cdots = v_\nu = 0\}$. Since $\overline{\mathcal{E}_\nu} = \{v_1 = 0\} \cap \overline{\mathcal{E}_\nu^3}$ and, for $v_1 = 0$, $F(v_2 - v_1, \ldots, v_n - v_1) = F(v_2, \ldots, v_n)$,

$$E[F(v_2 - v_1 \ldots, v_n - v_1) \,|\, \overline{\mathcal{E}_\nu}] = E[F(v_2, \ldots, v_n) \,|\, \overline{\mathcal{E}_\nu^3}]. \tag{14}$$

We will start by proving

$$E[F(v_2, \ldots, v_n) \mid \overline{\mathcal{E}_\nu^3}] \geq E[F(v_2, \ldots, v_n)]. \tag{15}$$

Being $F(u_1, \ldots, u_{n-1})$ non-decreasing on each $u_i$,

$$E[F(v_2, \ldots, v_n) \mid \mathcal{E}_\nu^3] = E[F(0, \ldots, 0, v_{\nu+1}, \ldots, v_n)] \leq E[F(v_2, \ldots, v_n)]$$

and, then, (15) follows from

$$\begin{aligned}
E[F(v_2, \ldots, v_n)] &= \prod_{2 \leq i \leq \nu} (1 - p_i) \; E[F(v_2, \ldots, v_n) \mid \mathcal{E}_\nu^3] \\
&\quad + \left[ 1 - \prod_{2 \leq i \leq \nu} (1 - p_i) \right] E[F(v_2, \ldots, v_n) \mid \overline{\mathcal{E}_\nu^3}].
\end{aligned}$$

Using (14) and (15), it is enough to prove

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \mid \mathcal{E}^1] \leq E[F(v_2, \ldots, v_n)]$$

and

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \mid \mathcal{E}_\nu^2] \leq E[F(v_2, \ldots, v_n)].$$

When the event $\mathcal{E}^1$ is realized $F(v_2 - v_1, \ldots, v_n - v_1) = F(v_2 - 1, \ldots, v_n - 1)$ and, being $F(u_1, \ldots, u_{n-1})$ non-decreasing on each $u_i$,

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \mid \mathcal{E}^1] = E[F(v_2 - 1, \ldots, v_n - 1)] \leq E[F(v_2, \ldots, v_n)],$$

proving the first result. When the event $\mathcal{E}^2(\nu)$ is realized $F(v_2 - v_1, \ldots, v_n - v_1) = F(0, \ldots, 0, v_{\nu+1}, \ldots, v_n)$ and, being $F(u_1, \ldots, u_{n-1})$ non-decreasing on each $u_i$,

$$E[F(v_2 - v_1, \ldots, v_n - v_1) \mid \mathcal{E}_\nu^2] = E[F(0, \ldots, 0, v_{\nu+1}, \ldots, v_n)] \leq E[F(v_2, \ldots, v_n)],$$

proving the second result.

$\square$

Let $A_r$, $1 \leq r \leq k$ be the set of alternatives which remain after the $r$th *DEBA* elimination step, and let conventionally $A_0 = \{1, 2, \ldots, m\}$. More specifically, $A_1$ includes the alternatives with maximum first attribute value, $A_2$ includes the alternatives in $A_1$ with maximum second attribute value, etc. We clearly have $A_0 \supset A_1 \supset A_2 \supset \cdots \supset A_k \neq \emptyset$. *DEBA* selects at random any alternative from $A_k$. We continue the proof of Theorem 4 by the following key proposition, from which Theorem 4 will follow almost immediately. The proposition establishes that, conditioned on any event $\mathcal{E}_r$ do not depending on the alternatives from attribute $r$ on and which does not preclude any given alternative $i$ to belong to the subset $A_{r-1}$ (and to any subset $A_n$, $0 \leq n \leq k$), and in terms of a function which is non-decreasing on the differences between the values of the attributes from attribute $r$ on of the other alternatives and the corresponding attributes of alternative $i$, the expected value of the function conditioned on $i \in A_n$ is non-increasing on $n$. The sketch of the proof of the

proposition is as follows. Since in the ZIAC probabilistic model, the $x_{j,s}$'s are independent, $\mathcal{E}_r$ only depends on $x_{j,s}$, $1 \le j \le m$, $1 \le s \le r-1$, and $\{i \in A_n\}$, $\{i \in A_{n+1}\}$ only depend on $x_{j,s}$, $1 \le j \le m$, $1 \le s \le n+1$, trivially $E[F(\mathbf{x}_{i,r}) | i \in A_{n+1} \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r}) | i \in A_n \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r})]$ for $0 \le n \le r-2$. For the case $r-1 \le n \le k-1$, either $P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r] = 0$ and we also have $E[F(\mathbf{x}_{i,r}) | i \in A_{n+1} \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r}) | i \in A_n \wedge \mathcal{E}_r]$, or, otherwise, being $A_{n+1} \subset A_n$, the result follows if $E[F(\mathbf{x}_{i,r}) | i \in A_{n+1} \wedge \mathcal{E}_r] \le E[F(\mathbf{x}_{i,r}) | i \in A_n - A_{n+1} \wedge \mathcal{E}_r]$. The proof of the latter is done by considering that $\{i \in A_{n+1}\} = \{i \in A_n\} \cap \mathcal{F}_{i,n+1}$ and $\{i \in A_n - A_{n+1}\} = \{i \in A_n\} \cap \overline{\mathcal{F}_{i,n+1}}$, where $\mathcal{F}_{i,n+1} = \{x_{i,n+1} = 1 \vee x_{j,n+1} = 0, j \in A_n\}$, and invoking Lemma 1, using the fact that $x_{j,n+1}$, $1 \le j \le m$ are independent of $x_{j,s}$, $1 \le j \le m$, $r \le s \le k$, $s \ne n+1$.

**Proposition 3.** *Let $r$, $1 \le r \le k$ and let $\mathcal{E}_r$ be any event depending only on $x_{j,s}$, $1 \le j \le m$, $1 \le s \le r-1$ such that $P[i \in A_{r-1} \wedge \mathcal{E}_r] > 0$ (which implies $P[i \in A_n \wedge \mathcal{E}_r] > 0$, $0 \le n \le k$). Let $\mathbf{x}_{i,r}$ be the $(m-1)(k-r+1)$-vector with components $x_{j,s} - x_{i,s}$, $1 \le j \le m$, $j \ne i$, $r \le s \le k$. Let $F(u_1, \ldots, u_{(m-1)(k-r+1)})$ be any function which is non-decreasing on each $u_l$. Then, for the ZIAC model and $0 \le n \le n' \le k$,*

$$E[F(\mathbf{x}_{i,r}) \mid i \in A_{n'} \wedge \mathcal{E}_r] \le E[F(\mathbf{x}_{i,r}) \mid i \in A_n \wedge \mathcal{E}_r].$$

*Proof.* It is enough to show

$$E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] \le E[F(\mathbf{x}_{i,r}) \mid i \in A_n \wedge \mathcal{E}_r]$$

for $0 \le n \le k - 1$. Since the events $\{i \in A_n\}$ and $\{i \in A_{n+1}\}$ only depend on $x_{j,s}$, $1 \le j \le m$, $1 \le s \le n+1$ and the $x_{j,s}$ are independent Bernouilli random variables, for $0 \le n \le r-2$, we trivially have $E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r}) \mid i \in A_n \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r})]$. It remains to consider the case $r-1 \le n \le k-1$. For $P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r] = 0$, the events $\{i \in A_n\} \cap \mathcal{E}_r$ and $\{i \in A_{n+1}\} \cap \mathcal{E}_r$ are identical, and we trivially have $E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] = E[F(\mathbf{x}_{i,r}) \mid i \in A_n \wedge \mathcal{E}_r]$. Therefore, assume $P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r] > 0$. Since, being $A_{n+1} \subset A_n$,

$$
\begin{aligned}
E[F(\mathbf{x}_{i,r}) \mid i \in A_n \wedge \mathcal{E}_r] &= P[i \in A_{n+1} \wedge \mathcal{E}_r \mid i \in A_n \wedge \mathcal{E}_r]\, E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] \\
&\quad + P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r \mid i \in A_n \wedge \mathcal{E}_r]\, E[F(\mathbf{x}_{i,r}) \mid i \in A_n - A_{n+1} \wedge \mathcal{E}_r] \\
&= \Big(1 - P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r \mid i \in A_n \wedge \mathcal{E}_r]\Big) E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] \\
&\quad + P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r \mid i \in A_n \wedge \mathcal{E}_r]\, E[F(\mathbf{x}_{i,r}) \mid i \in A_n - A_{n+1} \wedge \mathcal{E}_r] \\
&= E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] \\
&\quad + P[i \in A_n - A_{n+1} \wedge \mathcal{E}_r \mid i \in A_n \wedge \mathcal{E}_r] \\
&\quad \Big(E[F(\mathbf{x}_{i,r}) \mid i \in A_n - A_{n+1} \wedge \mathcal{E}_r] - E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r]\Big),
\end{aligned}
$$

and it is enough to prove that

$$E[F(\mathbf{x}_{i,r}) \mid i \in A_{n+1} \wedge \mathcal{E}_r] \le E[F(\mathbf{x}_{i,r}) \mid i \in A_n - A_{n+1} \wedge \mathcal{E}_r].$$

Let $\mathcal{F}_{i,n+1}$ be the event $\{x_{i,n+1} = 1 \vee x_{j,n+1} = 0, j \in A_n\}$. We have $\{i \in A_{n+1}\} = \{i \in A_n\} \cap \mathcal{F}_{i,n+1}$ and $\{i \in A_n - A_{n+1}\} = \{i \in A_n\} \cap \overline{\mathcal{F}_{i,n+1}}$. It is enough to show that

$$E[F(\mathbf{x}_{i,r}) \,|\, \mathcal{F}_{i,n+1} \wedge i \in A_n \wedge \mathcal{E}_r] \leq E[F(\mathbf{x}_{i,r}) \,|\, \overline{\mathcal{F}_{i,n+1}} \wedge i \in A_n \wedge \mathcal{E}_r]\,.$$

Letting $\widehat{\mathbf{x}}_{i,n+1}$ the $(m-1)$-vector with components $x_{j,n+1} - x_{i,n+1}$, $1 \leq j \leq m$, $j \neq i$, letting $\widetilde{\mathbf{x}}_{r,n+1}$ the $m(k-r)$-vector with components $x_{j,s}$, $1 \leq j \leq m$, $r \leq s \leq k$, $s \neq n+1$, denoting by $\pi(\widetilde{\mathbf{u}}_{r,n+1})$ the joint probability mass function of $x_{j,s}$, $1 \leq j \leq m$, $r \leq s \leq k$, $s \neq n+1$, conditioned on $\{i \in A_n\} \wedge \mathcal{E}_r$, and denoting by $F(\widetilde{\mathbf{u}}_{r,n+1}, \widehat{\mathbf{x}}_{i,n+1})$ the function of $\widehat{\mathbf{x}}_{i,n+1}$ which is obtained from $F(\mathbf{x}_{i,r})$ by setting each $x_{j,s}$ in $\widetilde{\mathbf{x}}_{r,n+1}$ to $u_{j,s}$, since $x_{j,n+1}$, $1 \leq j \leq m$, are independent of $x_{j,s}$, $1 \leq j \leq m$, $r \leq s \leq k$, $s \neq n+1$,

$$E[F(\mathbf{x}_{i,r}) \,|\, \mathcal{F}_{i,n+1} \wedge i \in A_n \wedge \mathcal{E}_r] = \sum_{\widetilde{\mathbf{u}}_{r,n+1} \in \{0,1\}^{m(k-r)}} \pi(\widetilde{\mathbf{u}}_{r,n+1}) E[F(\widetilde{\mathbf{u}}_{r,n+1}, \widehat{\mathbf{x}}_{i,n+1}) \,|\, \mathcal{F}_{i,n+1}]$$

$$E[F(\mathbf{x}_{i,r}) \,|\, \overline{\mathcal{F}_{i,n+1}} \wedge i \in A_n \wedge \mathcal{E}_r] = \sum_{\widetilde{\mathbf{u}}_{r,n+1} \in \{0,1\}^{m(k-r)}} \pi(\widetilde{\mathbf{u}}_{r,n+1}) E[F(\widetilde{\mathbf{u}}_{r,n+1}, \widehat{\mathbf{x}}_{i,n+1}) \,|\, \overline{\mathcal{F}_{i,n+1}}]$$

and, since, by Lemma 1,

$$E[F(\widetilde{\mathbf{u}}_{r,n+1}, \widehat{\mathbf{x}}_{i,n+1}) \,|\, \mathcal{F}_{i,n+1}] \leq E[F(\widetilde{\mathbf{u}}_{r,n+1}, \widehat{\mathbf{x}}_{i,n+1}) \,|\, \overline{\mathcal{F}_{i,n+1}}]$$

the result follows. $\qquad\square$

**Proof of Theorem 4.** The event "$i$ is chosen by *DEBA*" is identical to $\{i \in A_k\}$. Further, since $A_{r-1} \subset A_k$, $P[i$ is chosen by *DEBA* $\wedge \mathcal{E}_r] > 0$ implies $P[i \in A_{r-1} \wedge \mathcal{E}_r] > 0$. Applying, then, Proposition 3 with $n = 0$ and $n' = k$, taking into account that $\{i \in A_0\}$ is the "true" event, we obtain

$$E[F(\mathbf{x}_{i,r}) \,|\, i \text{ is chosen by } DEBA \wedge \mathcal{E}_r] \leq E[F(\mathbf{x}_{i,r}) \,|\, \mathcal{E}_r]\,.$$

But, since the $x_{j,s}$'s are independent and $\mathcal{E}_r$ only depends on $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq r-1$,

$$E[F(\mathbf{x}_{i,r}) \,|\, \mathcal{E}_r] = E[F(\mathbf{x}_{i,r})]\,. \quad \square$$

**Proof of Theorem 8.** Conditioned on $G = G'$, the $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq k$ are independent Bernouilli random variables. Nothing on the proof of Theorem 4 depended on the (for the ZIAC model) independent Bernouilli random variables $x_{j,s}$, $1 \leq j \leq m$, $1 \leq s \leq k$ having a "success" parameter depending only on $s$. Then, the result follows as a trivial generalization of Theorem 4. $\quad \square$

**Proof of Theorem 10.** Let, for $2 \leq r \leq k-1$ and $r+1 \leq t \leq k$, be the random variables

$$H_i^l(r,t) = \sum_{s=r+2}^{t} y_{i,s}^l\,, \quad 1 \leq i \leq m\,,$$

$$H_i^h(r,t) = \sum_{s=r+2}^{t} y_{i,s}^h\,, \quad 1 \leq i \leq m\,,$$

35

$$I^l(b, r, t) = \max_{1 \le i \le b} \left\{ 1 + \sum_{s=r+2}^{t} y_{i,s}^l \right\}, \quad 1 \le b \le m - 1,$$

$$I^h(b, r, t) = \max_{1 \le i \le b} \left\{ 1 + \sum_{s=r+2}^{t} y_{i,s}^h \right\}, \quad 1 \le b \le m - 1,$$

$$J(b, r, t) = \max \left\{ \mathbf{1}_{b>0} \max_{1 \le i \le b} \left\{ 1 + \sum_{s=r+2}^{t} y_{i,s}^l \right\}, \mathbf{1}_{m-b>1} \max_{b+1 \le i \le m-1} \left\{ 1 + \sum_{s=r+2}^{t} y_{i,s}^h \right\} \right\},$$

$$0 \le b \le m - 1,$$

and let (by symmetry, all $H_i^l(r, t)$, $1 \le i \le m$ have the same probability mass function, all $H_i^h(r, t)$, $1 \le i \le m$ have the same probability mass function, and the probability mass functions of $H_i^l(r, t)$, $H_i^h(r, t)$, $I^l(b, r, t)$, $I^h(b, r, t)$, $J(b, r, t)$, $Z_{l,g}(r, t)$, and $Z_{h,g}(r, t)$ depend on $r$ and $t$ through $t - r$)

$$P[H_i^l(r, t) = a] = \pi^l(t - r, a),$$

$$P[H_i^h(r, t) = a] = \pi^h(t - r, a),$$

$$P[I^l(b, r, t) = a] = \phi^l(b, t - r, a),$$

$$P[I^h(b, r, t) = a] = \phi^h(b, t - r, a),$$

$$P[J(b, r, t) = a] = \psi(b, t - r, a),$$

$$P[Z_{l,g}(r, t) = a] = \xi_{l,g}(t - r, a),$$

$$P[Z_{h,g}(r, t) = a] = \xi_{h,g}(t - r, a).$$

Then, the result follows using elementary probability theory by noting that

$$I^l(1, r, t) = 1 + H_1^l(r, t),$$

$$I^l(b, r, t) = \max\{I^l(b - 1, r, t), 1 + H_b^l(r, t)\}, \quad 2 \le b \le m - 1,$$

$$I^h(1, r, t) = 1 + H_1^h(r, t),$$

$$I^h(b, r, t) = \max\{I^h(b - 1, r, t), 1 + H_b^h(r, t)\}, \quad 2 \le b \le m - 1,$$

$$J(0, r, t) = I^h(m - 1, r, t),$$

$$J(m - 1, r, t) = I^l(m - 1, r, t),$$

for $1 \le b \le m - 2$, $J(b, r, t)$ is the maximum of two independent random variables with same probability mass functions as $I^l(b, r, t)$ and $I^h(m - b - 1, r, t)$, and

$$Z_{l,g}(r, t) = I_+(J(m - g - 1, r, t) - H_m^l(r, t)), \quad 0 \le g \le m - 1,$$

$$Z_{h,g}(r, t) = I_+(J(m - g, r, t) - H_m^h(r, t)), \quad 1 \le g \le m. \quad \square$$

# References

[1] Baucells, M., Carrasco, J. A., Hogarth, R. M., (in press). Cumulative dominance and heuristic performance in binary multi-attribute choice. Operations Research.

[2] Brace, K. S., R. L. Rudell, and R. E. Bryant 1990. Efficient implementation of a BDD package. In: Proc. 27th ACM/IEEE Design Automation Conference, IEEE Computer Society Press, Los Alamitos, USA, pp. 40–57.

[3] Bröeder, A., 2000. Assessing the empirical validity of the 'Take-The-Best' heuristic as a model of human probabilistic inference. Journal of Experimental Psychology: Learning Memory, and Cognition 26, 1332–1346.

[4] Bryant, R. E., 1986. Graph-based algorithms for Boolean function manipulation. IEEE Transactions on Computers C-35, 677–691.

[5] Czerlinski, J., Gigerenzer, G., Goldstein, D. G., 1999. How good are simple heuristics? In: Gigerenzer, G., Todd, P. M., and the ABC Research Group (Eds.), Simple Heuristics that Make us Smart. Oxford University Press, New York, USA, pp. 97–118.

[6] Ercegovac, M. D., Lang, T., 1985. Digital Systems and Hardware/Firmware Algorithms. John Wiley & Sons, New York, USA.

[7] Gigerenzer, G., Goldstein, D., 1996. Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review 103, 650–669.

[8] Gigerenzer, G., Todd, P. M., and the ABC Research Group, 1999. Simple Heuristics that Make us Smart. Oxford University Press, New York, USA.

[9] Hogarth, R. M., Karelaia, N., 2005. Ignoring information in binary choice with continuous variables: When is less more? Journal of Mathematical Psychology 49, 115–124.

[10] Hogarth, R. M., Karelaia, N., 2006. Take-the-Best and other simple strategies: why and when they work in binary choice. Theory and Decision 61, 205–249.

[11] Hsee, C. K., Rottenstreich, Y., 2004. Music, pandas, and muggers: On the affective psychology of value. Journal of Experimental Psychology 133, 23–30.

[12] Katsikopoulos, K., Fasolo, B., 2006. New tools for decision analysts. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans 36, 960–967.

[13] Keeney, R. L., Raiffa, H., 1993. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press, Cambridge, UK.

[14] Kirkwood, C. W., Sarin, R. K., 1985. Ranking with partial information: A method and an application. Operations Research 33, 38–48.

[15] Luenberger, D. G., 2003. Linear and Nonlinear Programming. Kluwer Academic Publishers, Boston, USA.

[16] Martignon, L., Hoffrage, U., 1999. Why does one-reason decision making work? A case study in ecological rationality. In: Gigerenzer, G., Todd, P.M., and the ABC Research Group (Eds.), Simple Heuristics that Make us Smart. Oxford University Press, New York, USA, 119–140.

[17] Martignon, L., Hoffrage, U., 2002. Fast, frugal, and fit: Simple heuristics for paired comparison. Theory and Decision 52, 29–71.

[18] Newell, B. R., Shanks, D. R., 2003. Take the best or look at the rest? Factors influencing one-reason decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition 29, 53–65.

[19] Newell, B. R., Weston, N. J., Shanks, D. R., 2003. Empirical tests of a fast-and-frugal heuristic: Not everyone takes-the-best. Organizational Behavior and Human Decision Processes 91, 82–96.

[20] Nikolskaia, M., Rauzy, A., Sherman, D. J., 1998. Almana: A BDD minimization tool integrating heuristic and rewriting methods. In: Proc. 2nd International Conference on Formal Methods in Computer-Aided Design. Springer-Verlag, Berlin, Germany, Lecture Notes in Computer Science 1522, 100–114.

[21] Payne, J. W., Bettman, J. R., Johnson, E. J., 1993. The Adaptive Decision Maker. Cambridge University Press, New York, USA.

[22] Somenzi, F., 2005. CUDD: CU decision diagram package release 2.4.1. Available at http://vlsi.colorado.edu/~fabio/CUDD/cuddIntro.html.

[23] Tversky, A., 1972. Elimination by aspects: A theory of choice. Psychological Review 79, 281–299.