

Table 1: Comparisons of  $D$ ,  $G$ ,  $IV$  and  $A$  ( $k = 3$ )

DESIGN 1		DESIGN 2								
		BOX- DRAPER	CCD	HOKE D2	HOKE D6	NOTZ	SCD	ECHIP (N=11)	ECHIP (N=18)	OPTEX
BOX- DRAPER (N=10)	$D$	–	97.7	37.2	0.0	30.2	97.7	20.9	15.5	29.0
	$G$	–	20.9	100.0	7.0	76.7	81.4	51.2	30.2	56.2
	$IV$	–	27.9	95.3	67.4	72.1	76.7	72.1	72.1	72.1
	$A$	–	27.9	86.0	27.9	72.1	76.7	72.1	68.2	72.1
CCD (N=18)	$D$	2.3	–	9.3	2.3	7.0	90.7	7.0	3.1	11.6
	$G$	79.1	–	100.0	16.3	83.7	93.0	62.8	61.2	70.2
	$IV$	72.1	–	81.4	72.1	72.1	88.4	72.1	72.1	72.1
	$A$	72.1	–	74.4	72.1	72.1	88.4	72.1	72.1	72.1
HOKE D2 (N=10)	$D$	62.8	90.7	–	7.0	9.3	90.7	2.3	12.4	25.2
	$G$	0.0	0.0	–	0.0	7.0	67.4	14.0	4.7	15.5
	$IV$	4.7	18.6	–	0.0	9.3	62.8	34.9	26.4	46.1
	$A$	14.0	25.6	–	2.3	62.8	67.4	58.1	23.3	62.4
HOKE D6 (N=13)	$D$	100.0	97.7	93.0	–	60.5	100.0	41.9	61.2	62.4
	$G$	93.0	83.7	100.0	–	97.7	88.4	79.1	88.4	98.4
	$IV$	32.6	27.9	100.0	–	88.4	74.4	72.1	100.0	72.9
	$A$	72.1	27.9	97.7	–	81.4	76.7	72.1	86.0	72.1
NOTZ (N=10)	$D$	70.8	93.0	90.7	39.5	–	90.7	7.0	35.7	43.0
	$G$	23.3	16.3	100.0	32.6	–	83.7	34.9	17.1	31.0
	$IV$	27.9	27.9	90.7	11.6	–	60.5	48.8	51.2	55.8
	$A$	27.9	27.9	37.2	18.6	–	60.5	44.2	38.0	56.2
SCD (N=11)	$D$	2.3	9.3	9.3	0.0	9.3	–	9.3	7.8	10.1
	$G$	18.6	7.0	32.6	11.6	16.3	–	18.6	14.7	17.4
	$IV$	23.3	11.6	37.2	25.6	39.5	–	37.2	34.9	38.8
	$A$	23.3	14.0	32.6	23.3	39.5	–	37.2	28.7	37.2
ECHIP (N=11)	$D$	79.1	93.0	97.7	58.1	93.0	90.7	–	70.0	92.2
	$G$	48.8	37.2	86.0	20.9	65.1	81.4	–	38.0	68.2
	$IV$	27.9	27.9	65.1	27.9	51.2	62.8	–	51.2	81.4
	$A$	27.9	27.9	41.9	27.9	55.8	62.8	–	71.3	62.8
ECHIP (N=18)	$D$	84.5	96.9	87.6	38.8	64.3	92.2	30.0	–	52.3
	$G$	69.8	38.8	96.9	11.6	82.9	85.3	62.0	–	70.5
	$IV$	27.9	27.9	73.6	0.0	48.8	65.1	48.8	–	56.6
	$A$	31.8	27.9	76.7	14.0	62.0	71.3	53.5	–	60.1
OPTEX (N=11)	$D$	70.9	88.4	74.8	37.6	57.0	89.9	7.8	47.7	–
	$G$	43.8	29.8	84.5	1.6	69.0	82.6	31.8	29.5	–
	$IV$	27.9	27.9	53.9	27.1	44.2	61.2	18.6	43.4	–
	$A$	27.9	27.9	37.6	27.9	43.8	62.8	19.4	39.9	–

Note: Values are % of all models across all permutations of labels in which DESIGN 1 is superior to DESIGN 2.

(superior implies the optimality criterion difference is  $\geq 0$  for  $D$ ,  $A$ , and  $G$  and is  $\leq 0$  for  $IV$ )

Table 2: Comparisons of  $D$ ,  $G$ ,  $IV$  and  $A$  ( $k = 4$ )

DESIGN 1		DESIGN 2								
		BOX- DRAPER	CCD	HOKE D2	HOKE D5	NOTZ	SCD	PBCD	ECHIP (N=15)	ECHIP (N=28)
BOX- DRAPER (N=15)	$D$	–	19.6	1.3	1.3	1.3	78.6	38.4	41.7	0.1
	$G$	–	0.9	73.2	35.7	78.6	57.6	65.2	72.9	4.2
	$IV$	–	12.9	3.6	2.7	8.9	43.3	19.2	75.1	8.1
	$A$	–	12.1	3.1	4.0	7.6	44.2	19.2	75.7	0.1
CCD (N=28)	$D$	80.4	–	18.8	0.4	19.2	93.3	72.8	61.5	1.7
	$G$	99.1	–	96.4	81.7	100.0	86.6	82.1	87.3	83.7
	$IV$	87.1	–	84.4	83.5	87.1	79.5	76.8	73.3	80.8
	$A$	87.9	–	84.4	64.7	84.8	81.2	57.1	73.9	80.8
HOKE D2 (N=15)	$D$	98.7	81.2	–	44.2	99.5	93.8	83.9	72.8	15.4
	$G$	26.8	3.6	–	21.9	61.6	66.5	62.9	72.7	8.4
	$IV$	96.4	15.6	–	12.9	100.0	62.9	19.2	79.0	71.9
	$A$	96.9	15.6	–	13.8	21.4	57.6	19.2	78.8	33.8
HOKE D5 (N=19)	$D$	98.7	99.6	55.8	–	63.4	99.1	99.1	76.3	21.9
	$G$	64.3	18.3	78.1	–	76.8	65.6	84.4	73.3	26.1
	$IV$	97.3	16.5	87.1	–	100.0	65.6	19.2	77.9	73.2
	$A$	96.0	35.3	86.2	–	92.0	65.6	22.8	77.5	72.8
NOTZ (N=15)	$D$	98.7	80.2	5.8	36.6	–	93.3	83.5	71.9	12.8
	$G$	21.4	0.0	39.7	23.2	–	63.4	57.6	71.8	6.5
	$IV$	91.1	12.9	0.0	0.0	–	62.1	19.2	75.4	71.1
	$A$	92.4	15.2	78.6	8.0	–	57.6	19.2	77.5	35.2
SCD (N=17)	$D$	21.4	6.7	6.2	0.9	6.7	–	2.2	21.2	1.9
	$G$	42.4	13.4	33.5	34.4	36.6	–	38.4	42.5	27.8
	$IV$	56.7	20.5	37.1	34.4	37.9	–	36.2	72.5	49.4
	$A$	55.8	18.8	42.4	34.4	42.4	–	35.3	66.4	37.1
PBCD (N=21)	$D$	61.6	27.2	16.1	0.9	16.5	97.8	–	47.8	3.3
	$G$	34.8	17.9	37.1	15.6	42.4	61.6	–	66.6	19.6
	$IV$	80.8	23.2	80.8	80.8	80.8	63.8	–	73.7	77.6
	$A$	80.8	42.9	80.8	77.2	80.8	64.7	–	73.4	76.0
ECHIP (N=15)	$D$	58.3	38.5	27.2	23.7	28.1	78.8	52.2	–	11.5
	$G$	27.1	12.7	27.3	26.7	28.2	57.5	33.4	–	19.9
	$IV$	24.9	26.7	21.0	22.1	24.6	27.5	26.3	–	12.9
	$A$	24.3	26.1	21.2	22.5	22.5	33.6	26.6	–	15.1
ECHIP (N=28)	$D$	99.9	98.3	84.6	78.1	87.2	98.1	96.7	88.5	–
	$G$	95.8	16.3	91.6	73.9	93.5	72.2	80.4	80.1	–
	$IV$	91.9	19.2	28.1	26.8	28.9	50.6	22.4	87.1	–
	$A$	99.9	19.2	66.2	27.2	64.8	62.9	24.0	84.9	–

Note: Values are % of all models across all permutations of labels in which DESIGN 1 is superior to DESIGN 2.

(superior implies the optimality criterion difference is  $\geq 0$  for  $D$ ,  $A$ , and  $G$  and is  $\leq 0$  for  $IV$ )

Table 3: Comparisons of  $D$ ,  $G$ ,  $IV$  and  $A$  ( $k = 5$ )

DESIGN 1		DESIGN 2					
		BOX-DRAPER	CCD	HOKE D2	HOKE D7	PBCD	NOTZ
BOX-DRAPER (N=21)	$D$	–	15.3	0.0	2.3	68.2	0.0
	$G$	–	1.7	0.0	0.0	74.0	0.0
	$IV$	–	4.8	0.0	0.0	19.9	0.0
	$A$	–	7.5	0.0	0.1	20.3	0.0
CCD (N=30)	$D$	84.7	–	1.5	0.1	91.9	1.3
	$G$	98.3	–	66.3	19.4	94.9	83.9
	$IV$	95.2	–	87.7	87.7	91.3	87.7
	$A$	92.5	–	83.1	24.7	81.4	87.7
HOKE D2 (N=21)	$D$	100.0	98.5	–	67.8	96.8	31.5
	$G$	100.0	33.7	–	29.8	90.5	88.3
	$IV$	100.0	12.3	–	12.3	58.5	87.7
	$A$	100.0	16.9	–	15.0	58.2	84.6
HOKE D7 (N=26)	$D$	97.7	99.9	32.2	–	99.8	27.7
	$G$	100.0	80.6	70.2	–	96.8	87.0
	$IV$	100.0	12.3	87.7	–	78.9	87.7
	$A$	99.9	75.3	85.0	–	82.1	87.2
PBCD (N=23)	$D$	31.8	8.1	3.2	0.2	–	2.8
	$G$	26.0	5.1	9.5	3.2	–	13.2
	$IV$	80.1	8.7	41.5	21.1	–	64.5
	$A$	79.7	18.6	41.8	17.9	–	56.1
NOTZ (N=21)	$D$	100.0	98.7	68.5	72.3	97.2	–
	$G$	100.0	16.1	11.8	13.0	86.8	–
	$IV$	100.0	12.3	12.3	12.3	35.5	–
	$A$	100.0	12.3	15.4	12.3	43.9	–

Note: Values are % of all models across all permutations of labels in which DESIGN 1 is superior to DESIGN 2.

(superior implies the optimality criterion difference is  $\geq 0$  for  $D$ ,  $A$ , and  $G$  and is  $\leq 0$  for  $IV$ )