Chapter 3H – COMPENSATION: U.S. NAVY RESEARCH INITIATIVES AND APPLICATIONS

Tanja F. Blackstone, Ph.D.
Navy Personnel Research, Studies, and Technology, USA

tanja.blackstone@navy.mil

3H.1 ABSTRACT

Historically, the U.S. Navy has used selective re-enlistment bonuses (SRB) to induce retention in critical skill groups. While SRBs have been shown to be effective, the effectiveness is limited and can be relatively costly. Retention bonuses are paid to all eligible individuals in critical skill groups, even if the individual's intention would be to re-enlist absent the bonus. Similarly, individuals in a given skill group all receive equivalent bonus amounts. This standardized approach to bonus awards ignores the 'true' compensation level required to induce the desired retention behavior. Further, the SRB program has had limited success in filling jobs in undesirable locations.

Alternatives to the SRB program include a flexible market based incentive system. Beginning in 2003, the U.S. Navy, in a limited test case basis, implemented Assignment Incentive Pay (AIP). AIP uses a simplified modification of a first price auction and has been shown to have success in filling positions in historically undesirable locations. Advances to AIP currently under development include the Distribution and Incentive System, (DIS). DIS incorporates an advanced optimization technique that considers the Sailor's bid and Navy costs in the job assignment process.

In this paper we discuss (1) the use and effectiveness of the U.S. Navy's SRB program, advances in modeling and estimating the impact of SRB on re-enlistment behavior, (2) experimental results of the feasibility of a multi-attribute auction in assigning military personnel to jobs, and (3) based on preliminary basic research results couched in choice theory, the net benefits of a flexible compensation plan is discussed.

3H.2 INTRODUCTION

Compensation policy and elements are updated and published by the Department of Defense (DoD) every four years.[1] Active duty personnel are entitled to three elements of "regular military compensation," basic pay, basic allowance for quarters (BAQ), and basic allowance for subsistence (BAS). The United States Congress authorizes annual adjustments to regular pays. However, BAQ and BAS reflect the variability in the cost of living by locale and therefore, vary by location. Depending on several factors, such as whether Officer or Enlisted, duty type, and skill group, military personnel can be eligible for other types of compensation.

3H.3 CURRENT COMPENSATION[2]

Compensation benefits consist of cash and non-cash benefits. Direct cash benefits consist of basic pay and allowances with indirect cash benefits being defined the tax-exempt portion of cash benefits.[3] The average

[1] See *Military Compensation Background Papers,* Fifth Edition, September 1996.

[2] Current compensation information derived from Congressional Budget Office estimates, published January 2004.

[3] For example, allowances are federally tax-exempt.

military member receives approximately 43% of his total compensation package in cash. The remaining 57% of compensation is in the form of non-cash benefits and consists of health care, deferred retirement payments, housing and subsidized goods. Of the proportion of non-cash benefits approximately 9% is deferred for retirement pay. However, only one third of officers and 15% of enlisted personnel will collect retirement pay.[4]

The Congressional Budget Office (CBO) estimates that the average active duty member receives an annual compensation package estimated at $99,000.00 for 2002. The CBO estimates does not differentiate between officer and enlisted compensation, which reasonably one would expect to observe a great degree of variability. Ignoring non-cash benefits and assuming that the average annual level of compensation of an active duty member is $43,000.00, cash bonuses designed to encourage retention can substantially increase a member's annual cash compensation. For example, bonuses for enlisted personnel range from $1800.00 to as high as $45,000.00 for an additional four years of service. The average dollar payout under the Navy's selective re-enlistment bonus program for FY2000 – 2003 is given in Table 3H-1. Similarly, Officer bonus awards for critical skill groups are authorized at $65,000.00 and higher for additional years of obligated service.

Table 3H-1: Average SRB Award by Paygrade

	Re-enlist Avg SRB Amount			
Dim Fiscal Year				
Dim_Paygrade	**2000**	**2001**	**2002**	**2003**
All Dim_Paygrade	$27,013.36	$26,487.21	$25,600.36	$22,033.40
E-2	–	–	–	$1,858.88
E-3	$16,433.75	$17,729.22	$16,481.23	$12,344.91
E-4	$24,965.71	$25,130.32	$23,898.44	$19,322.76
E-5	$29,839.60	$27,780.06	$26,370.97	$21,699.58
E-6	$26,930.01	$28,827.33	$29,430.72	$30,091.91
E-7	$16,062.12	$19,564.49	$23,001.23	$24,729.64
E-8	$18,284.16	$22,380.24	$22,027.16	$23,134.63

The use of bonuses as a retention tool is particularly effective during cycles of economic expansion. The private sector, notably the technology and aviation sectors compete directly with DOD for certain skills. Unlike DOD, the private sector is not constrained with respect to increases in basic pay and in tight labor markets the private sector can raise pay to attract highly skilled military personnel. In an effort to mitigate the flow of personnel from DOD to the private sector, DOD frequently uses bonuses to retain skilled personnel.

Several different types of incentives can be classified loosely as bonuses and can be classified into two general categories: Selective Re-enlistment Bonuses (SRB) and auctions. In the remainder of this document an overview of the SRB and auction programs are discussed. In addition, innovative advances in current compensation research efforts being undertaken in the United States Navy are discussed.

[4] Eligibility to collect retirement requires a minimum time in service of 20 years.

3H.3.1 Selective Re-enlistment Bonus

Bonuses are used by DoD to sustain targeted levels of personnel by skill group and pay grade.[5] Bonuses are typically used to obtain a targeted level of junior personnel within community (skill group) with the objective of ensuring sufficient future manning levels in the junior and senior ranks/pay grades. For the enlisted communities bonuses are directly paid to a skill group, independent of pay grade. In contrast, officer bonuses are targeted by rank and skill group.

3H.3.2 Enlistment

Typically, the Navy offers monetary incentives in order to induce re-enlistment or retention behavior in critical or undermanned skill groups. The level of monetary incentives varies across periods and Enlisted Management Communities, (EMCs).

Individual eligibility for Selective Re-enlistment Bonuses (SRB) is contingent on re-enlistment eligibility and if the individual re-enlists for a minimum of two years. Individual bonus amounts are calculated using a combination of the SRB multiplier, individual's base pay and the number of years the individual re-enlisted. The individual receives one-half of the total SRB in a lump sum payment upon re-enlistment with the remainder paid out in monthly increments over the re-enlistment contract.

The average retention rate by zone for the period for FY1996 – 2004 is given in Table 3H-2. Holding end strength requirements constant, a number of factors, such as projected retirements and quits, the U.S. Navy sets a required retention rate by Zone.[6] The retention rate for Zone A and B is approximately 36% and 50%, respectively. The Navy uses SRB to mitigate anticipated retention shortfalls by Zone and skill groups. For example, for the period from FY1996 – 2002 Zone A projected retention rates were far below the necessary critical level. In response, the Navy not only expanded their SRB program to capture a greater number of skill groups, but also increased the overall average SRB award.

Table 3H-2: Fiscal Year to Date Retention Rates by Zone U.S. Navy

Dim_Time_Period									
LOSYear	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
Zone A	24.30%	23.50%	23.80%	23.90%	27.80%	35.20%	34.50%	38.90%	37.30%
Zone B	56.00%	55.60%	53.00%	54.80%	58.30%	65.50%	69.50%	70.70%	67.80%
Zone C	78.00%	77.00%	74.80%	76.50%	78.80%	81.90%	83.60%	81.70%	84.00%
Zone D	70.80%	82.80%	85.80%	80.90%	94.40%	95.80%	96.60%	96.40%	96.00%
Zone E	18.10%	17.50%	16.40%	17.00%	18.10%	22.10%	25.90%	18.60%	25.20%

[5] The amount of monetary incentive offered in any given time period to any given EMC is largely driven by projected end strength requirements.

[6] Zones can be thought of as years of service, with individuals in Zone A have six or less years of service, individuals in Zone B having 7–11 years of service, and Zone C individuals having 12–16 years of service. Typically, SRB is targeted to Zone A and Zone B.

3H.3.3 Officer

Similar, to enlisted personnel, the Department of the Navy targets key officer communities to meet projected end-strength targets. As of FY04, Surface Warfare Officers are offered a $50K bonus given at the end of their second division tour of duty for additional obligated service at the department head level. Bonuses are offered to individuals in medical, nuclear, and aviation sub-specialties. Unlike the enlisted community, however, officer bonuses are geared towards rank and subspecialties and are relatively much more effective in shaping manning requirements.

3H.3.4 Estimating the Effect of SRB on Retention – Enlisted Personnel

The traditional approach to estimating the effect of retention to a change in compensation (bonus) is the Annualized Cost of Leaving Model (ACOL), formalized by Warner and Goldberg (1984). Simply described the ACOL model calculates the difference between military and civilian pay discounted to the present. An individual will stay in the military if the pecuniary returns to staying exceed the returns of leaving. Incremental increases to SRB, increases to the military pay side of the equation, gradually increases the cost of leaving, and hence induce the intended effect of increasing retention. The parameter estimates obtained from the ACOL models provide an indicator of the average sensitivity of retention behavior to a change in SRB, or an elasticity measure.

While the ACOL formulation is theoretically correct and represent the right approach, the underlying statistical assumptions and data limitations result in under/over estimation of the effect of change in SRB on retention. For example:

- The different ACOL calculators (see the ACOL theoretical and applied literature) use different underlying assumptions. Therefore, using the same data and same discount factors, different calculators/models yield different ACOL values.
- Most models use a discount factor of only 10 – 20 percent. This may result in an under-estimated (predicted) value of the SRB effect. Example: A one unit increase in SRB for the Weapons Control group (1995 – 2002) yields a 0.3 percentage point (ppt) increase under ACOL with a 10 percent discount factor, a 1.1 ppt increase under ACOL with a 20 percent discount factor, and a 3.4 ppt increase under ACOL with a 40 percent discount factor (or 3.6 ppt under the same model, but with three choices: leave, extend, re-enlist).
- Most of the ACOL studies employing the basic five to seven variables, as well as unemployment, have very poor statistical performance. (See Hansen and Wenger, 2002, for a detailed comparison and discussion of the different ACOL models.)
- Non-linearity issues in the data, the model, and in individuals' behavior *are not taken into account*. For example, if under an ACOL (20%) binary model, an increase in SRB by a unit (to all entitled personnel) results in a 1.1 ppt increase in re-enlistment, the current models use this estimate to forecast the personnel behavior of an increase larger than one unit in SRB. Example: If the Navy wants to increase re-enlistment (for a certain group, say Weapons Control) by 3.3 percent, then the current models suggest increasing SRB by 3 units to entitled personnel (e.g., 3 times 1.1). In doing so, these models do not take into account two major issues: non-linearity in the functional form of the econometric model, and (more important) the *decreasing marginal effects* of the SRB. Taking these two effects into consideration will yield better estimates (even under the ACOL model).
- Terminology. In the ACOL literature, the experiments are written as an increase in SRB by a unit to *all personnel*. This is misleading as the ACOL calculator "allows" an SRB (or SRB change) only for

the entitled personnel by Navy criteria (e.g., LOS 2 – 14 years, Zones A – C, certain grades, only if an individual re-enlist for at least 3 years, etc.).

- In some cases the observed number of individuals receiving SRB payments is so small (approximately 1% – 1-.5%), so the experiment is based on practically irrelevant estimates, resulting in biased estimates.

In addition, since the seminal work of Warner and Goldberg (1984) who introduced the notion of Annualized Cost of Living (ACOL) all models build on that approach under the assumption of normality, thereby using the Maximum Likelihood (ML) probit (Normally distributed errors) model. Violations of the ML assumptions can produce biased and inefficient parameter estimates. If SRB levels set by Navy policy makers are based on ill behaved data and or biased parameter estimates, underpayment or overpayment of SRB is the likely end result. Underpayment of SRB leads to missed retention goals, whereas overpayment leads to increased cost over and above the minimal threshold necessary to meet desired retention targets.

3H.3.5 Model Improvements

Navy Personnel Research, Studies, and Technology is currently engaged in an effort to address the data issues noted above and take advantage of recent advances in statistical estimation. The basic set of objectives of the research is to examine the different models with recent data and to develop the necessary tools that will help the policy makers in their decision process involving Navy personnel. Specifically, in this study the objectives are:

1) Study the impact of SRB and other pay increases on retention.
2) Study the impact of unemployment on retention.
3) Advances in statistical estimation, in particular in Information Theoretics are used to compare and contrast the Generalized Maximum Entropy (GME) non-ACOL model with the ML ACOL model. GME falls under the class of Information Theoretic models.

The models developed include ACOL-based models and new non-ACOL based models. These models are all members of the Information-Theoretic class of estimation methods and are generalization of the traditional Maximum Likelihood Logit. The basic model is built such that it accommodates for the unbalanced (different time periods for different enlisted personnel) nature of the existing data. The main advantages of these new Information-Theoretic models are that they are more flexible, use fewer assumptions, include the traditional maximum likelihood method as a special case, and allow the users to incorporate prior information and any other type of soft data.[7]

To achieve this set of goals, a detailed study of four Navy skill groups is done. These skill groups are Weapons Control, Sensor Operations, Administration, and General Seamanship. These groups were chosen because they represent the broad spectrum of Navy personnel. For each of these groups, the ACOL and non-ACOL models were used to investigate the above three objectives for the group as a whole, for subgroups (such as Zone A, B, and C) as well as over time. Further, all estimations were repeated for different choices (binary: leave, re-enlist; three-choice: leave, extend, re-enlist; four-choice: leave, short-extension, long extension, and re-enlist). For the ACOL models used here, all analyses were repeated for different discount factors (5% through 40%) and for different sets of variables. This detailed set of estimations was needed in order to compare and contrast our estimates with all other recent (and less recent) studies (i.e., Hansen and

[7] For a more detailed description of Information Theoretic Models see Golan and Perloff.

Wegner, 2002; Asch and Warner, 2001; Goldberg, 2001). For brevity, only the results form the Weapons Control group are discussed.[8]

3H.3.5.1 Weapons Control

We demonstrate our approach with the main results obtained for the Weapons Control group (1995 – 2002) consisting of 66,509 individuals. The independent variables used (in addition to the intercept) are Gender, Race, Number of Children, AFQT Score, Base Pay, Total Allowance, Education Dummies (No High School and Above High School), Sea Duty, Dollar Amount SRB, Zone Dummies, Expected Civilian Wage (Golan, 2003), Lag Real Interest Rate, Lag Value of NASDAQ index, and Unemployment Rate.

The Tables and Figures below present the results of the SRB experiments studied and developed here. Table 3H-3 presents the basic results for the full sample. *The normalized values (lower and upper bounds as reflected by zero and infinitely high discount factors) are the values recommended to use.* Table 3H-4, presents a comparison across different model scenarios: the binary and the three-choice multinomial, as well as Zone A, B, and C for the infinitely high discount factor. The main results observed here are that:

1) An analysis of small increments of SRB is more accurate since the ppt are not linear;
2) There is a major difference in ppt between a multinomial and binary models (that use the same right hand side variables); and
3) Individuals in different zones behave differently.

Table 3H-3: Weapons Control – SRB Experiments for GME 3 Categories Model (Full Sample 66,509 Observations)

	Base Case	SRB+0.5	SRB+1	SRB+1.5	SRB+2	SRB+2.5	SRB+3
Implied Elasticities (0)		0.45	0.40	0.36	0.33	0.30	0.28
Prob Leave	0.4615	0.4008	0.3554	0.3224	0.2982	0.2802	0.2664
Prob Extend	0.2032	0.2118	0.2135	0.2115	0.2075	0.2025	0.1971
Prob Re-enlist (infinity)		0.3580	0.3842	0.4078	0.4289	0.4478	0.4647
Prob Re-enlist (0)	0.3353	0.3875	0.4311	0.4661	0.4943	0.5173	0.5364
% Change Re-enlist (infinity)		6.8%	14.6%	21.6%	27.9%	33.6%	38.6%
% Change Re-enlist (0)		15.6%	28.6%	39.0%	47.4%	54.3%	60.0%
Mean $SRB	1,244	1,677	2,129	2,582	3,034	3,486	3,938
Normalized % Change Re-enlist (infinity)		6.3%	12.2%	17.7%	22.7%	27.3%	31.5%
Additional Re-enlisted Personnel		3,458	7,715	11,506	14,964	18,090	20,817
Normalized % Change Re-enlist (0)		12.3%	23.0%	31.7%	38.8%	44.7%	49.5%

[8] A complete review of the results can be obtained from Navy Personnel Research, Studies, and Technology, 5720 Integrity Drive, Millington, TN 38055-1000.

Table3H-4: Weapons Control – SRB Experiments for GME Binary and 3 Categories and by Zones

	SRB+0.5	SRB+1	SRB+1.5	SRB+2	SRB+2.5	SRB+3
Normalized % Change Re-enlist (infinity) 3-Choice	6.3%	12.2%	17.7%	22.7%	27.3%	31.5%
% Change Re-enlist (infinity) Binary	1.8%	6.4%	10.8%	15.0%	19.0%	22.8%
% Change Re-enlist (0) 3-Choice	15.6%	28.6%	39.0%	47.4%	54.3%	60.0%
% Change Re-enlist (0) Binary	8.0%	15.9%	23.3%	30.2%	36.5%	42.2%
Normalized % Change Re-enlist (0) 3-Choice	12.3%	23.0%	31.7%	38.8%	44.7%	49.5%
Normalized % Change Re-enlist (0) Binary	6.7%	13.4%	19.8%	25.6%	NA	NA
Normalized % Change (infinity) Zone A 3-Choice	4.7%	9.3%	13.8%	18.2%	22.3%	26.3%
Normalized % Change (infinity) Zone B 3-Choice	9.8%	18.7%	27.5%	35.8%	43.5%	50.7%
Normalized % Change (infinity) Zone C 3-Choice	24.6%	44.8%	60.6%	72.8%	82.3%	89.7%

Table 3H-5 provides a detailed analysis of the traditional ACOL model (call it "Base") and the more detailed model (call it "extended") that includes more information on both the individuals and the macro state of the economy. Both types of models are presented for the binary and three-choice models and are based on the ML-Logit. (Probit analyses were done and yielded very similar results, so it is not presented here.).

Table 3H-5: Re-enlistment Percentage Point Increase for SRB Increase of One Unit for the Basic and Extended ACOL Models: Weapons Controls (1995 – 2002): 66,509 Observations

Groups	Basic											
	Binary			Three Choices			Extended Binary			Three Choices		
	10%	20%	40%	10%	20%	40%	10%	20%	40%	10%	20%	40%
Full Sample	0.4	1.2	3.9	0.9	2.3	4.9	0.3	1.1	3.4	0.3	1.2	3.6
1996	0.3	0.9	3.1	0.5	1.8	4.1	0.1	0.7	2.6	0.1	0.6	2.5
1997	0.5	1.5	4.6	1.2	3.1	6.0	0.3	1.1	3.7	0.3	1.3	4.0
1998	0.4	1.1	3.7	0.9	2.4	5.1	0.2	0.7	2.9	0.2	0.9	3.4
1999	0.4	1.3	4.2	0.9	2.2	4.9	0.4	1.2	3.8	0.4	1.5	4.2
2000	0.3	1.1	3.8	0.9	2.2	4.4	0.4	1.2	3.3	0.4	1.4	3.4
2001	0.4	1.2	4.0	1.0	2.4	4.8	0.4	1.3	3.5	0.5	1.7	3.9
2002	0.3	1.0	3.4	1.0	2.3	4.3	0.4	1.2	3.1	0.5	1.6	3.5
Zone A	6.3	10.3	11.5	5.9	9.9	11.5	4.3	7.3	8.6	4.2	7.3	8.6
Zone B	6.2	9.4	13.6									
TOS 4	6.5	13.7	22.0									
TOS 5	0.2	1.2	5.8									
TOS 6	0.3	1.0	3.5									

1) TOS is the Terms of service (4, 5, 6 years).

2) The Base Pay Elasticities are 0.2 (ACOL with 10% discount), and 0.4 for ACOL with 20% discount.

The main results here are that (1) the ppt increases as the discount factor increases, (2) the three choice model yields (on average) higher ppt than the binary models, (3) the basic model ("Base") yields higher ppt values than the more general model ("Extended"), and (4) that different subgroups (e.g., Zone A) behave differently.

Additional experiments comparing ACOL formulation with models controlling for individual pay components (non-ACOL) show that in all of the cases, the non-ACOL GME has good performance relative to the ACOL ML models. The GME non-ACOL has better in-sample performance as well as superior out-of-sample performance.

3H.3.6 Unemployment Effects

As noted earlier, economic factors are a key influence on retention behavior. To see this effect the models controlled for unemployment and examined the effect of retention to changes in unemployment. In general, over the sample period (1995 – 2002) enlisted personnel became less sensitive to the unemployment rate. This result is consistent throughout subsets of the data, across models (ACOL, non-ACOL, ML, GME, etc.), and across the different unemployment measures (general, by education, and others). Table 3H-6 shows these results for the Weapons Control group.

Table 3H-6: Weapons Control – Unemployment Analysis

Evaluated at			1995 – 1996	1997 – 1999	2000 – 2002
Mean Charac.	Unemp.	GME – 3 categories	-2.17*, -.61, 3.27*	–	-.43*, -.12, .14*
	Unemp. Education	GME – 3 categories	-.54*, 1.42*, -.24	-.19, .63*, -.33*	-.41*, -.08, .12*
	Unemp.	Binary GME-1	3.2*	–	.29*
	Unemp.	Binary GME-2	1.9*	–	.07*
	Unemp.	Binary ML ACOL-20	3.53*	–	0.41*
	Unemp.	Binary ML ACOL-40	3.55*	–	0.40*
Mean over individuals	Unemp.	GME – 3 categories	-, -.31, 3.57*	–	-.3*, .01, .28*
	Unemp. Education	GME – 3 categories	-.45*, 1.51*, -.15	-.14, .69*, -.28*	-.28*, .05, .25*
Mean AFQT			72.4	74.7	76

Notes: * significant

Binary GME-1: 0 – leave and extend; 1 – re-enlist

Binary GME-2: 0 – leave; 1 – re-enlist and extend

There are probably a number of explanations for this result and more is needed from those who are familiar with the changes in rules and policies such as downsizing certain professions, however one observation is notable here: the mean quality (AFQT) of enlisted personnel (only for Weapons Control) increased during that period, which may explain part of the reduction in sensitivity toward unemployment (higher quality personnel are, on average, less sensitive to unemployment). Other explanations, observed in the other data sets include a dramatic increase in SRB (from approximately 0.3 in mid 90s on average for the Sensor Operation group) to approximately 0.9 in 2001–2. This reflects a substantial average increase of SRB payments.

To summarize, the individuals may be very sensitive to the unemployment rate (or the lag unemployment), but the macro-economic and political conditions as well as the wage/bonus increases/changes over time, make it hard to prove. Below are the main observations.

1) First, an analysis of the whole period (1995 – 2002) by groups reveals that except for the Administration group, all unemployment elasticities (with respect to Re-enlistment) are positive.
2) For all groups the sensitivity to unemployment declines over time (or becomes statistically insignificant).

Both general unemployment and unemployment by education yield similar results.

3H.3.7 Conclusions on SRB Model Improvements

The main objectives of the research were to investigate the sensitivity of enlisted personnel to SRB and other pay components, as well as to investigate the sensitivity of the personnel to the basic economic conditions. Using data from 1995 through 2002, these effects were studied for four basic skill groups: Weapons Control, Sensor Operations, General Seamanship, and Administration. A comparison of different scenarios of the Maximum Likelihood (ML) ACOL models with the Information Theoretic non-ACOL GME method was examined. The main results of that research are:

1) The sensitivity of personnel to SRB payments is skill specific and in all cases the enlisted personnel became less sensitive to the SRB payment over the period analyzed.
2) In all cases, enlisted personnel became less sensitive to unemployment levels and to other macroeconomic indicators over time.
3) In cases (skill groups) where we don't observe many individuals that received SRB, the SRB experiments should be done differently than in those cases where a significant portion of the enlisted personnel received SRBs. We provide such an analysis for the General Seamanship skill group.
4) The non-ACOL GME method yielded superior estimates and forecasts for all cases.
5) The ML-ACOL models are very sensitive to the discount factor used and to other specifications.

Spin offs of the SRB program include Targeted SRB. Where SRB has been successful in shaping the force by skill group, the tool does not ensure that the right individual is in the right job. Targeted SRB is tied to skill and location. Bonuses are paid to eligible individuals who agree to hard to fill locality assignments or locales that have skill group shortages. Unfortunately, data and/or metrics of effectiveness on this program are not available.

3H.3.8 Auctions

Relatively new incentive programs in use by the U.S. Navy are auctions. The main advantage of auctions over SRB is the individual elicitation of the actual dollar value needed to induce retention. With SRBs it is possible that given the fixed nature of the bonus that the bonus offer is insufficient to elicit the desired retention behavior. Similarly, the SRB bonus offer may be set such that individuals who re-enlisted in response to the bonus would have stayed for some lesser amount. In the first, case the direct impact is manning shortages and in the second case the Navy incurs a pecuniary cost.

Auctions elicit the valuation that the individual requires to re-enlist. And unlike a fixed SRB policy, auction rules can be designed to allow payouts to be reflective of the average private sector valuation of a given job, hence, mitigating the ability of the private sector's ability to use pay to compete for skilled labor.

Beginning in FY03 Assignment Incentive Pay (AIP) was launched by the Navy to increase "volunteerism" for hard to fill locations. Certain locations, such as Sigonella, Italy, were experiencing skill shortages by pay grades. Forced moves of individuals to historically undesirable locations, especially those at higher pay grades with the option to retire, resulted in a high quit (retirement) rate, exacerbating shortages at senior levels. In FY02 the United States Congress authorized funding of AIP. AIP is tied to locale, paygrade and skill type.

Each period the Navy sets a maximum incentive pay or price that it is willing to pay to eligible individuals to voluntarily accept a hard to fill assignment. The maximum price varies by locality, pay grade and skill. The Navy's reserve price, the maximum amount the Navy is willing to pay an individual for accepting a hard

to fill position, is known to the individual. Eligible Individuals who are in their rotation window can bid on a job up to the Navy's reserve price. Individual's bids reflect the additional monthly pay the individual would require to accept a hard to fill position, hence the Navy's reserve price and likewise bids can be considered incremental pay to monthly salaries.

AIP auction rules are similar to those of a first-price sealed bid auction. Sealed bids are accepted for a given time period, with the job being awarded to the lowest bidder. The winning bidder receives a bonus equal to his bid for each month that the individual remains in the hard to fill job. As of April 2004 the Navy's reserve price ranged from $50.00 to $1200.00. The lowest received bid was $0.00 with the average bid being $310.00.

A much more advanced auction under development is the Distribution Incentive System (DIS). DIS is being designed to consider not only the individual's preferences for a given job, as revealed by the individual's bid, but Navy constraints and or policies. DIS assigns a weighted fitness score in determining the auction winner.

Unlike AIP, which assigns winners strictly based on lowest bid, DIS is designed to optimize over individual bids, productivity evaluations, PCS costs, and readiness requirements. DIS uses an advanced algorithm to minimize the next cost of assignments subject to Navy manpower readiness constraints. As of this writing DIS is being beta tested on university subjects for parameter validation.

A precursor to DIS was the Job Market Labor Allocation Model (JMLAM).

3H.3.9 Job Market Labor Allocation Model (JMLAM)

Rather than the traditional single unit auctions, where a single unit of a good or service is brought to a market, the use of multi-object auctions are steadily becoming the preferred auction mechanism for goods and services that are either complements or substitutes, (Matsui and Watanabe, 2003). The multi-object/unit auctions have been used to auction off Federal Communications Commission (FCC) licenses, transportation services, and delivery routes (Matsui and Watanabe, 2003). A key feature of multi-unit objects is that they allow the buyer to submit a single bid on a combination of objects of the buyer's choosing. This type of auction mechanism allows for complementary (substitutes) goods to be sold as a single unit, thereby, increasing the value of the bundle to the buyer.[9] The primary difficulty with multi-unit objects is evaluating the winner of a given auction. For N buyers in the market over M goods there are [N! / r!(N-r)!] possible combinations of objects. The ability of the seller to evaluate each buyer's bid is likely to be computationally intractable.

Numerous auction algorithms have been used to determine winners in multi-unit auctions. Typically, optimization algorithms used in multi-unit auctions are designed to maximize the auctioneer's revenue, maximize allocative efficiency and/or determine location specific prices, (McCabe, Rassenti, and Smith, 1991). Examples of the use of such algorithms can be found in the FCC spectrum, airline slots, gas pipeline networks, and sales routes auctions, (Ledyard, Olson, Porter, Swanson, and Torma, 2002; McCabe et al., 1991).

Another difficulty inherent in multi-unit objects is that buyers do not necessarily bid their true valuation for each item in the bundle ((Matsui and Watanabe, 2003). In other words, the buyer must determine his or her valuation for each possible bundle combination. Determination by any buyer of the value of a given bundle of objects is likely to be cognitively prohibitive in particular as the possible bundle combinations grows in relation to the number of objects. Various methods are discussed in the literature to overcome the difficulty of

[9] Complementarity is defined as a bundle of items that are worth more than the sum of their parts. Substitutability is defined as a bundle that is worth less than the sum of its parts, (Hudson and Sandholm, 2002).

preference elicitation and elicitation of the buyer's marginal valuation of a bundle, (Hudson and Sandholm, 2002; Conen and Sandholm, 2001). Our research, however, does not focus on solving the problem of preference elicitation.

In an effort to shed some insight into alternative methods to allocate labor across jobs, we explore the feasibility of using a multi-unit auction approach. The auction environment is designed to allow multiple sellers (sellers of labor) and a single buyer (buyer of labor) to negotiate bids over multiple attributes. A seller sells a single unit of the hypothetical good or service (analogous to a unit of labor) by submitting bids on one or more attributes.[10] Unlike the general case in multi-unit auctions, sellers cannot, per se, submit bids on a subset of the available objects, but must bid on the bundle (attributes) at hand. Sellers, however, can choose the desired level for each attribute. Likewise the buyer, using a scoring rule (see Annex A) must evaluate the bids as a bundle.

In an effort to mimic some of the constraints faced in the military labor market, the experimental auction allows for (1) forced market convergence, that is a hypothetical job must be filled, (2) control of a seller's value, where seller value can be defined as a measure of the marginal productivity (marginal cost) of a seller. The results reported from this research effort report findings from base case experiments.

3H.4 RESEARCH OBJECTIVE

Unlike the private sector, the Navy cannot offer individually tailored compensation packages. Institutional constraints mandate that the Navy pay Sailors of equivalent paygrade and length of service equivalent wages. The Navy can, however, negotiate, that is offer incentives, on various job attributes, such as bonuses, reduced sea-shore rotation, and promotion points, for example.

An efficient means of negotiation is to couch the Sailor (seller) – Navy (buyer) negotiation process in an auction environment. The Navy labor market, however, has unique constraints that require modifications to any proposed auction approach. These constraints include the requirement that jobs cannot go unfilled and auction accessibility by participants.[11]

Using a first price sealed bid auction approach; this paper explores the bidding behavior of subjects in a multi-attribute market setting. The proposed auction is designed to examine bidding behavior and market outcomes in a controlled experimental environment that mimics Navy constraints and the effects of possible policy proposals.

The Job Market Labor Allocation Model (JMLAM) is designed to provide a bidding environment where a buyer and multiple sellers can submit offers and or counteroffers on the attributes of a hypothetical good/service. The model is designed to allow flexibility for experimental testing of Navy constraints and possible policy impacts. Experimental parameters can be set by the experimenter to examine the effect of bidding behaviors under varying constraints and or rules.

[10] An attribute is the terminology used to refer to a form of compensation with each attribute being a component of a bid.

[11] Given the worldwide dispersion of military personnel and the relatively short duration of a member's rotation window an online auction is necessitated. Bandwidth and security issues restrict web access, thereby preventing sailors from participating in a real time online auction.

It is important to note that for purposes of experimental testing, observing of bidding behavior, and market outcomes, jobs are assigned generic names and the incentive attributes are referred to as "attributes." This is done in order to prevent individuals from bringing in egalitarian, altruistic, or other preferences from everyday social life into the experiment.

The JMLAM user interface is simple in design, (see Figure 3H-1 below) and allows the auction participants to view bidding history and earnings information, and participate in one or more auctions in a given time period. The history screen provides information to the seller that may help the seller in formulating future bids. The option of participation in multiple auctions is analogous to a Sailor bidding on multiple assignments.[12]

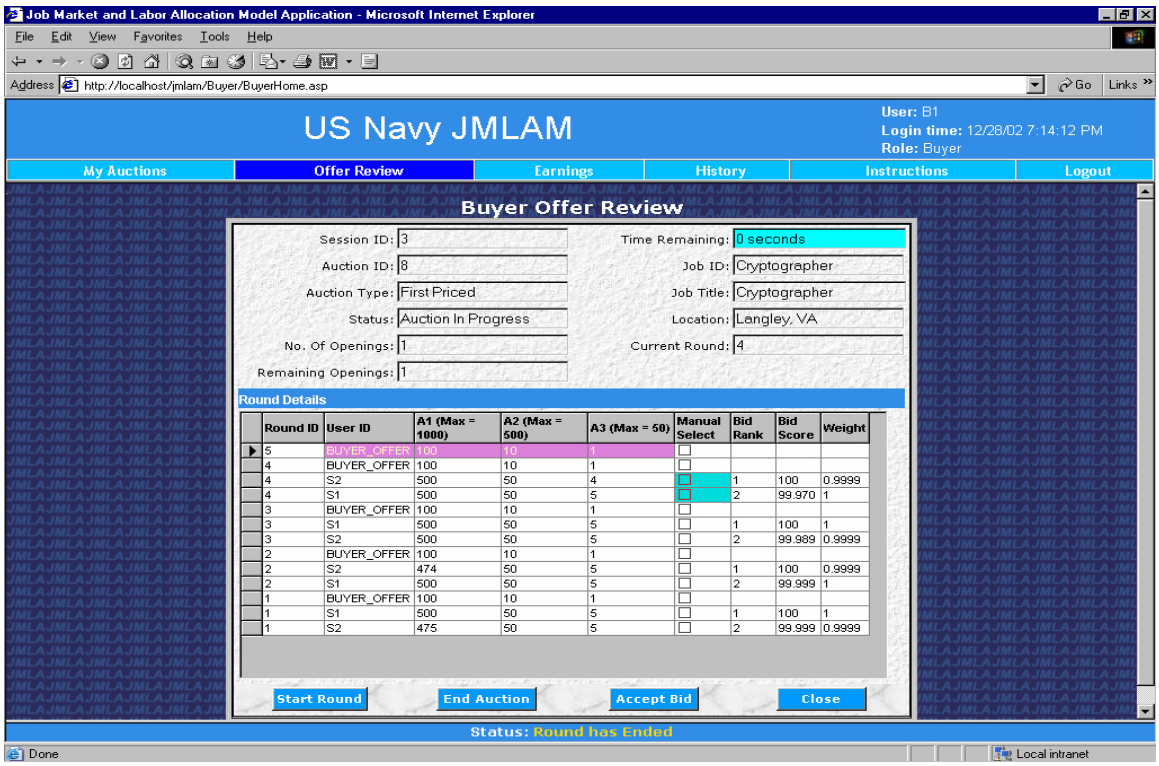

Figure 3H-1: JMLAM User Interface.

The auction is designed so that each seller can sell a single unit of a hypothetical good/service, while the buyer can purchase one or more units of the good/service. The maximum number of goods/services that the buyer is willing and able to purchase is unknown to the seller. Subjects are informed that at least one good/service is available, however, that the buyer is not obligated to accept a bid from a seller(s). The buyer's objective is to maximize the difference between the buyer's reserve and the seller offers. Earnings are only accrued to the buyer if he or she accepts a seller's offer and the offer is less than the buyer's reserve. Buyer's earnings, therefore, increase the greater the difference between the buyer's reserve and the accepted seller's offer and the number of seller's bids the buyer accepts.

[12] Experiments to date; however, have been limited to the participation of subjects in only one auction at a time.

Sellers submit bids on three hypothetical attributes of a good/service. Each attribute has a reserve price. Depending on the experimental parameters, (see Tables 3H-7 and 3H-8) the reserve prices may vary across sellers and auctions. Sellers know their reserve prices; sellers do not know the other sellers' reserve prices or the distribution of reserve prices. The implication here is that while subjects can refer to the history screen for guidelines as to how to submit future bids, knowledge of past bidding behavior may or may not be helpful.

Table 3H-7: Experiment 1 Parameters

Seller Value	1
Reserve Prices, N = 360	(10, 5, 2)
Reserve Prices, N = 408	(10, 5, 2) or (13, 7, 3) or (18, 9, 4)
Job Openings	2, unknown to the seller
Seconds for Play	90 – 105 seconds
Conversion Rate	.05, $1 Experimental Dollar = $0.05 USD
Maximum Number of Rounds	10, unknown to seller
Buyer Reserve	(100, 50, 25), unknown to a seller
Maximum Number of Sellers	9, average subjects per session was 7

Table 3H-8: Experiment 2 Parameters

Seller Value	.6, .8 or 1, randomly assigned session/auction/subject
Reserve Prices	(13, 10, 6) if seller value = .6
Reserve Prices	(16, 12, 8) if seller value = .8
Reserve Prices	(20, 15, 10) if seller value = 1
Job Openings	2, unknown to the seller
Seconds for Play	60 seconds
Conversion Rate	.10, $1 Experimental Dollar = $0.10 USD
Maximum Number of Rounds	10, unknown to seller
Buyer Reserve	(100, 50, 25), unknown to the seller
Maximum Number of Sellers	9, average subjects per session was 7

The seller's earnings are determined by the difference from his or her reserve price and the submitted bid. The greater the difference between the reserve price and the submitted bid the greater the earnings. A seller only realizes earnings if the seller wins an auction. The probability of winning an auction diminishes the greater the difference between seller's reserve price and his or her submitted bid. The complexity of setting an optimal bid is compounded by the fact the seller is required to submit bids on three attributes. Bidding strategy is further complicated by the fact that each seller is assigned a "seller value."[13] Each seller knows his or her seller value, but other seller values and/or the distribution of the seller values are unknown to individual sellers. The buyer, however, knows all seller values and the distribution of seller values.

[13] See Annex B for a description an function of the 'seller value.'

The probability of winning an auction, therefore, is a function of the seller bid and seller value. Given two sellers, s1 and s2, with seller values of 1 and .5 respectively and who submit identical bids, the seller with the higher seller value will win the auction, in this case s1.

The seller's bids are evaluated using a scoring rule, described in greater detail in Annex A. Assuming that the buyer is willing and able to accept a bid, the buyer chooses that seller with the highest rank. If the buyer can purchase multiple units of the hypothetical good or service, the buyer simply chooses the *n* highest ranked sellers until the number of available units is exhausted.

3H.5 EXPERIMENTAL DESIGN

This research is an ongoing effort, with two experimental sessions having been completed to date. The experimental parameters adopted in Experiment 1 were largely designed to provide a benchmark from which to compare future experiments. In all cases seller value remained fixed across sessions and auctions. In approximately one-half of the auctions subject reserve prices were fixed across auctions. In those cases reserve prices varied, subjects were randomly assigned to one of the following three (A1, A2, A3) groupings; (10, 5, 2), (13, 7, 3) or (18, 9, 4).

Approximately 60 subjects participated in the first experimental session. Each subject group participated in 2 sessions of 10 auctions each, with a sample size of 768 observations. In general, the two highest ranked sellers were chosen as auction winners. While the auction is designed to enable multiple rounds, winners were always assigned in the first round. In general, sellers submitted bids very close to their reserve; typically within $1 – $3 experimental above their reserve and in some cases sellers bid their reserve on at least one attribute. Submitting bids equal to seller reserves may be attributed to the fact that sellers did not know the buyers reserve. Future experimentation will examine the effect of bidding behavior when the buyer's reserve is known, (see Annex B).

In Experiment 2, subjects were asked to participate in 3 sessions, for a total of 810 observations. Experiment 2 subjects were assigned seller values and corresponding reservation prices in a random order. Higher seller values imply higher productivity and, therefore, higher associated reserve prices were assigned to sellers with higher seller values. Based on the bidding behaviors observed in Experiment 1, the seller value/reserve prices were set so that there would be an equal probability of winning an auction, independent of the seller value/ reserve price combination.

3H.6 PRELIMINARY AUCTION RESULTS

At this stage of the research it is uncertain as to whether subjects focus on the summed value of their bids, a composite bid, or consider each attribute bid independently. To this extent we present analysis on composite bidding behavior and individual attributes. Median and minimum bids are examined as a cursory insight into bidding behaviors in Experiment 1 and 2. The median bid, relative to a mean bid, is examined to mitigate the effect of outliers. To provide additional insight into subject bidding behavior we mapped the minimum submitted bid by auction.

Experiment 1 was the experimental base case, where all seller values were fixed at "1." In approximately one-half of the auctions, the reserve prices were fixed across subjects, with subjects randomly assigned variable reserve prices in the remaining auctions. Figure 3H-2 reflects the median difference in the composite reserve prices and composite bids. In the initial auctions the median difference in the subject groups is opposite of what we had expected and surprisingly large. One would expect that the median bid in the "same

group" would be relatively lower than the median bid in the "different reserve group," particularly, since the reserve prices for two-thirds of the subjects was significantly larger than the "same reserve group." Over successive auctions, however, the median difference between the two groups begins to converge.

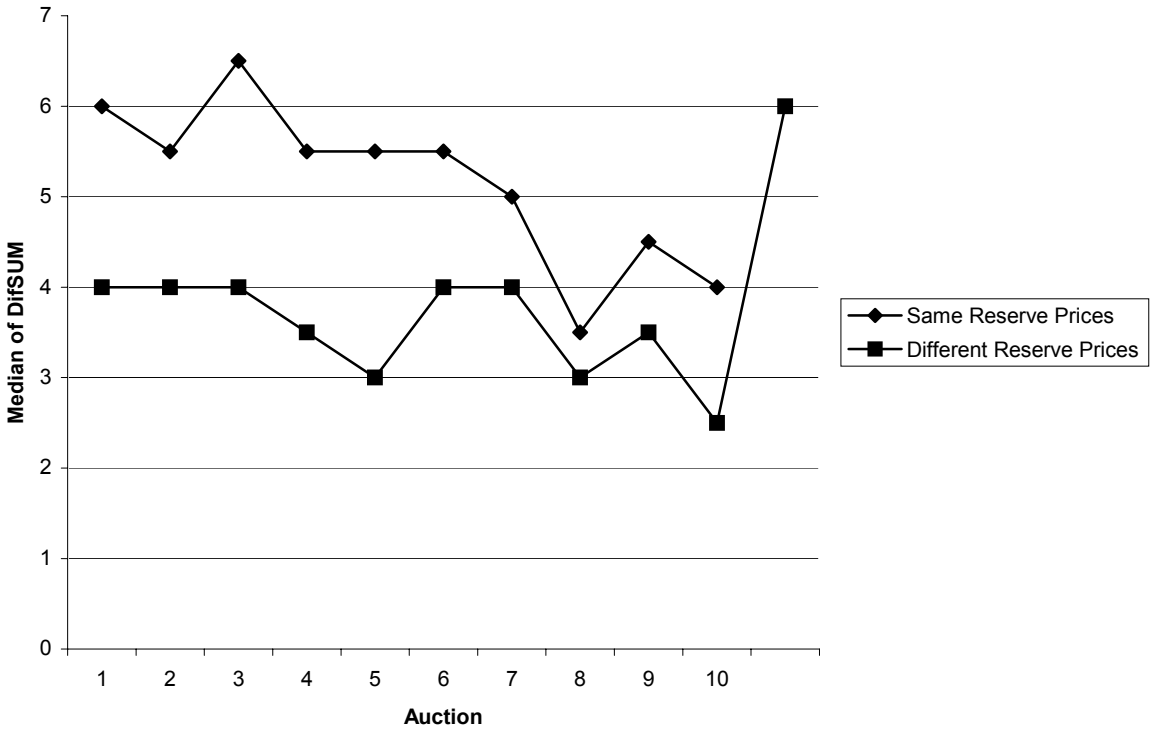

Figure 3H-2: Experiment 1 Median of Difference in Composite Bids.

Figure 3H-3 shows the minimum difference in the composite bids and reserve prices. For the "same reserve price" group at least one subject submitted a composite bid that equaled his or her reserve price. Bids submitted at reserve may imply that subjects placed a greater value on simply winning the auction relative to any earnings potential; alternatively, subjects may be searching for information on possible winning bid combinations. In contrast, the minimum bid in the "different reserve group" was consistently above the minimum composite reserve of the "same reserve group" (see Figure 3H-3). This effect could be attributed to biases in the experimental design. In Experiment 1, there was no randomization in the order subjects played the auctions. In all cases, subjects participated in the "same reserve price" auctions and then the "different reserve price" auctions. This non-randomization of experiments may have provided information to the subjects regarding the distribution of reserve prices across subjects, thereby influencing the bids upward.

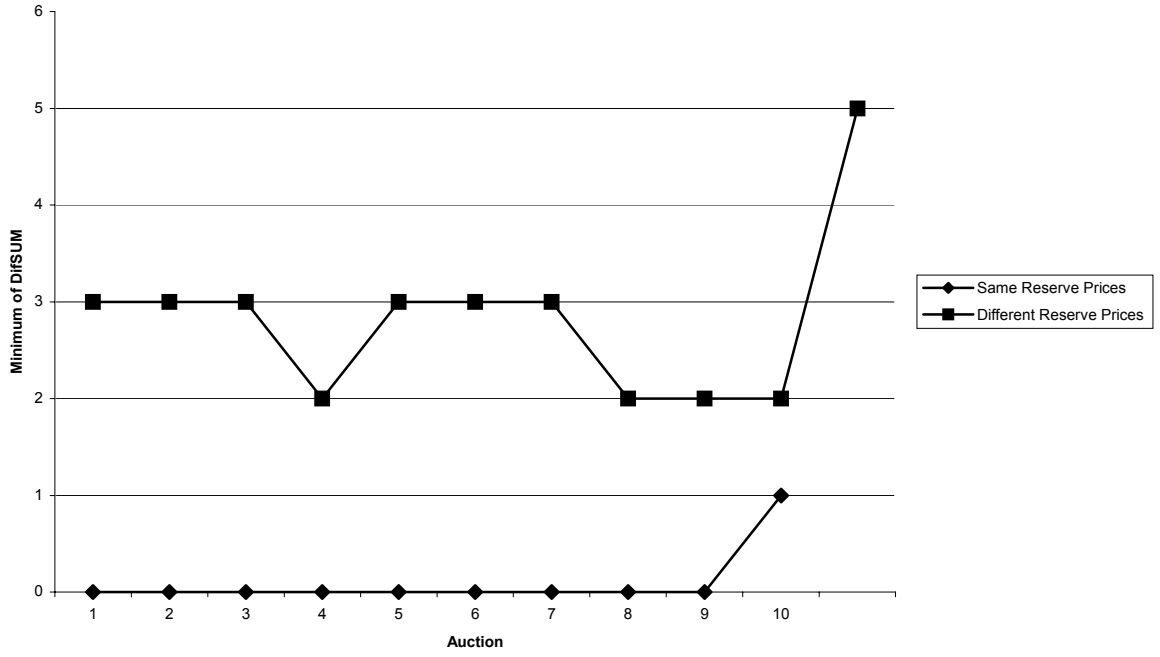

Figure 3H-3: Experiment 1 Minimum of Differences in the Composite Bid.

In Experiment 2, four groups of subjects participated in three auction sessions.[14] Each auction session consisted of 9 auctions for a total of 108 auctions completed. In each auction, two units of the hypothetical good or service were available; therefore, for each auction the two highest ranked bidders were awarded the units. In contrast to Experiment 1, subjects in Experiment 2 were assigned varying seller values; with each seller value assigned a corresponding set of reserve prices.

In Figures 3H-4 and 3H-5 we look at the differences in the median and minimum composite bids. A pattern of convergence to the composite reserve prices is observed across all four groups. In general, the minimum composite bid allowed subjects to potentially earn $1 Experimental. Interestingly, at least one subject per auction submitted his or her reserve price, inferring that some subjects may have been searching for information on winning bid combinations.

[14] Group 1 participated in auctions 1 – 27, Group 2 participated in auctions 28 – 53, Group 3 participated in auctions 54 – 80, and Group 4 participated in auctions 81 – 108. The quadrants, denoted by vertical lines, in the Experiment 2 graphs are mapped to group specific bidding behaviors.

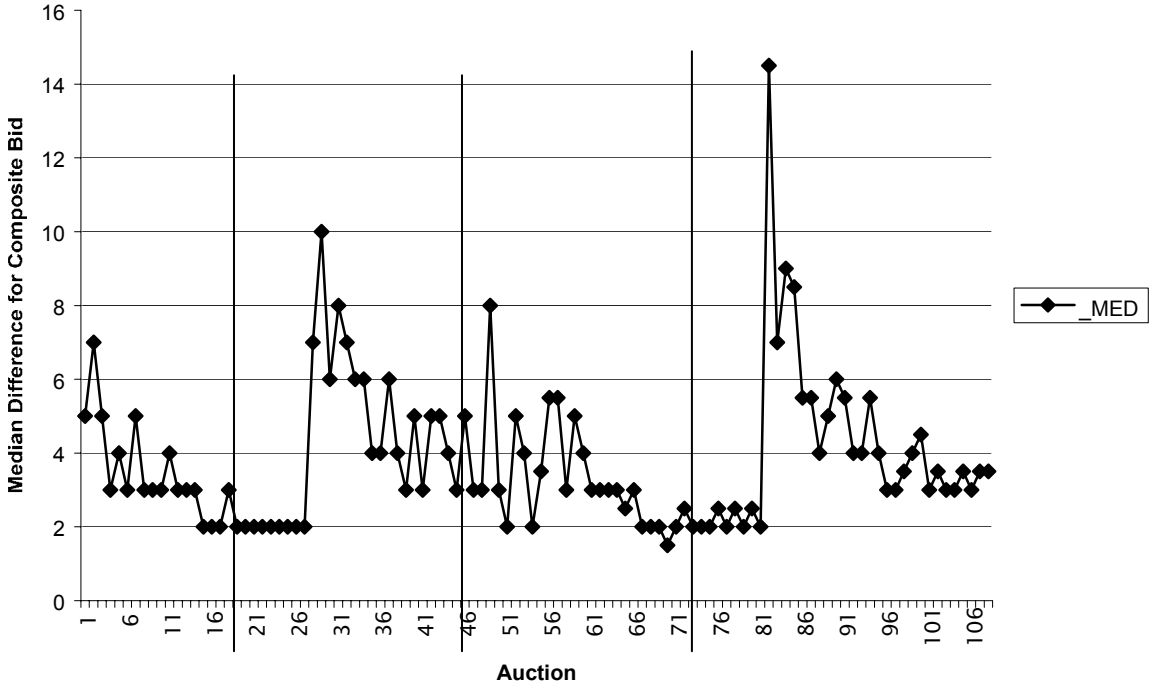

Figure 3H-4: Experiment 2 Median Difference of Composite Bid.

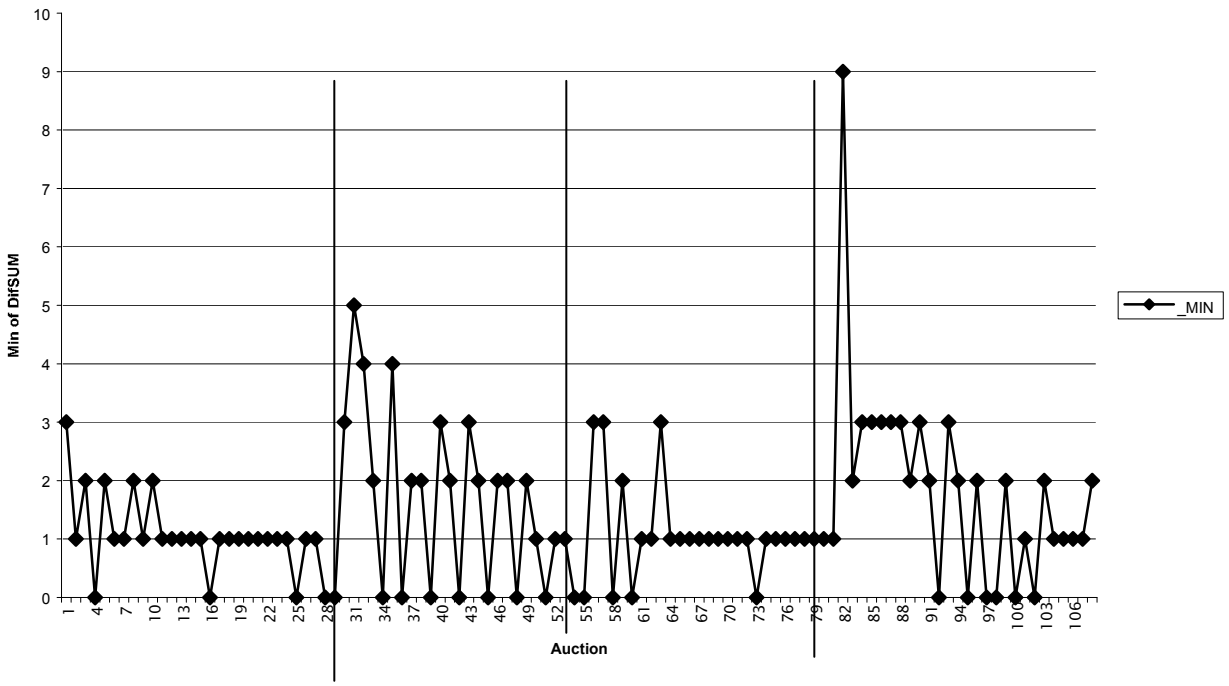

Figure 3H-5: Experiment 2 Minimum of Composite Bid Difference.

In Figures 3H-6 through 3H-11 we decompose the bidding behavior of the four groups by attributes. There appears to be a distinct demarcation between Groups 1 and 3 and Groups 2 and 4. Groups 1 and 3 have a more frequent relative tendency to bid their reserve on at least one attribute. As compared to the median bids for Groups 2 and 4, where, in general, the median was above the attribute reserve price across all attributes.

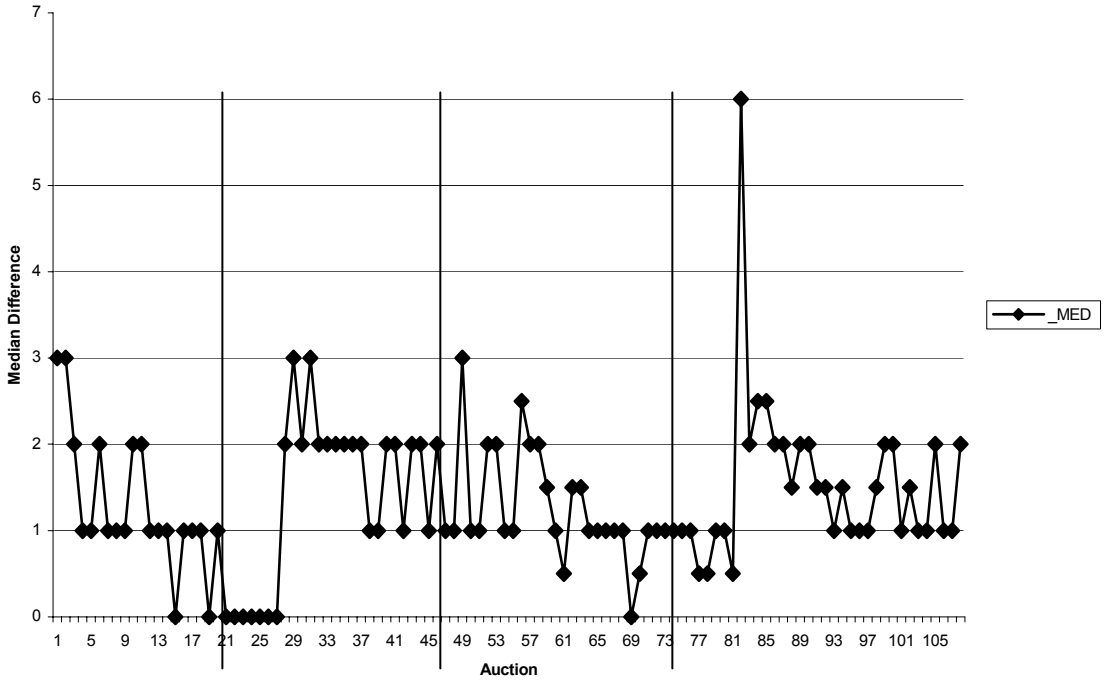

Figure 3H-6: Experiment 2 Median Difference between Reserve Price and Bid for A1.

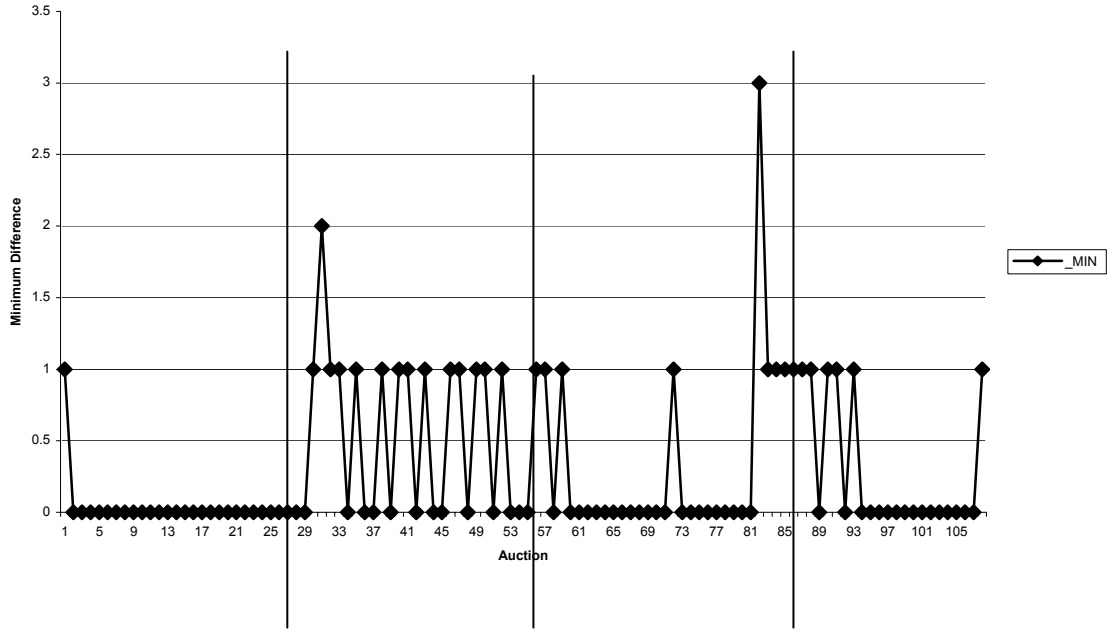

Figure 3H-7: Experiment 2 Minimum Difference between Reserve Price and Bid for A1.

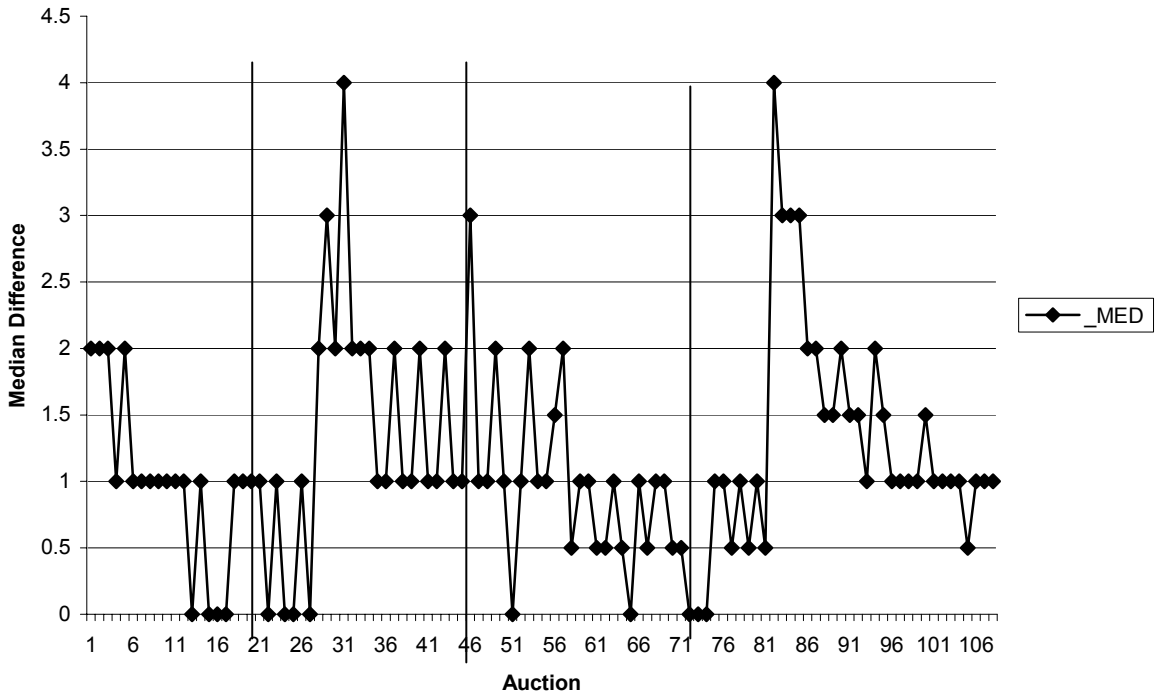

Figure 3H-8: Experiment 2 Median Difference between Reserve Price and Bid for A2.

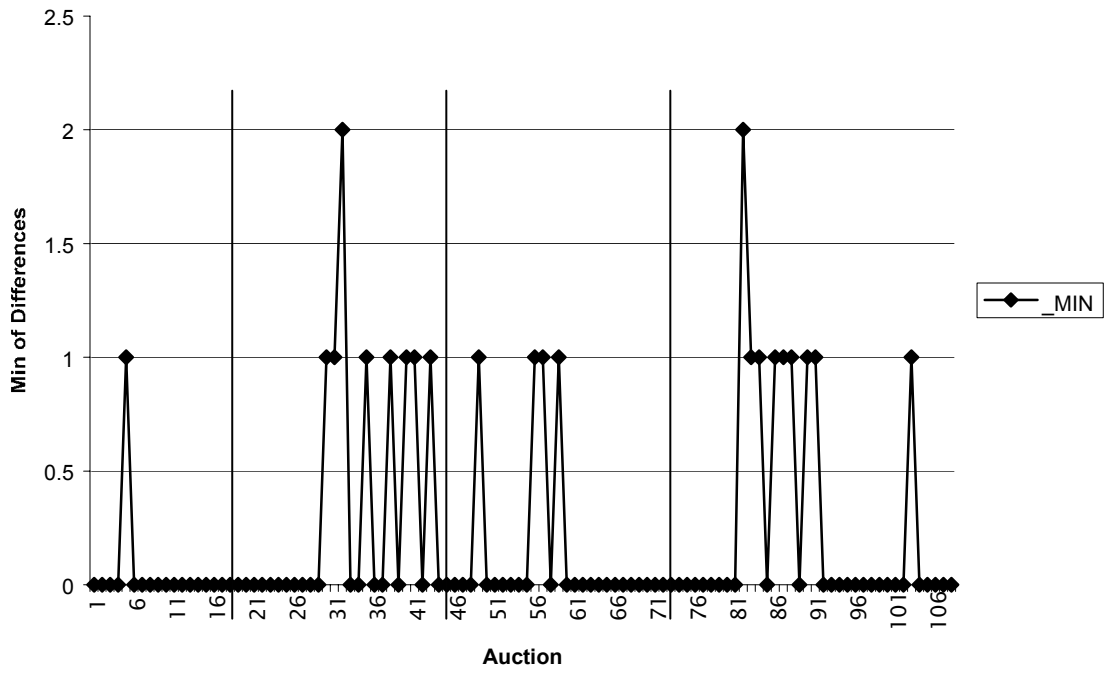

Figure 3H-9: Experiment 2 Minimum Value of Differences between Reserve Price and Bid for A2.

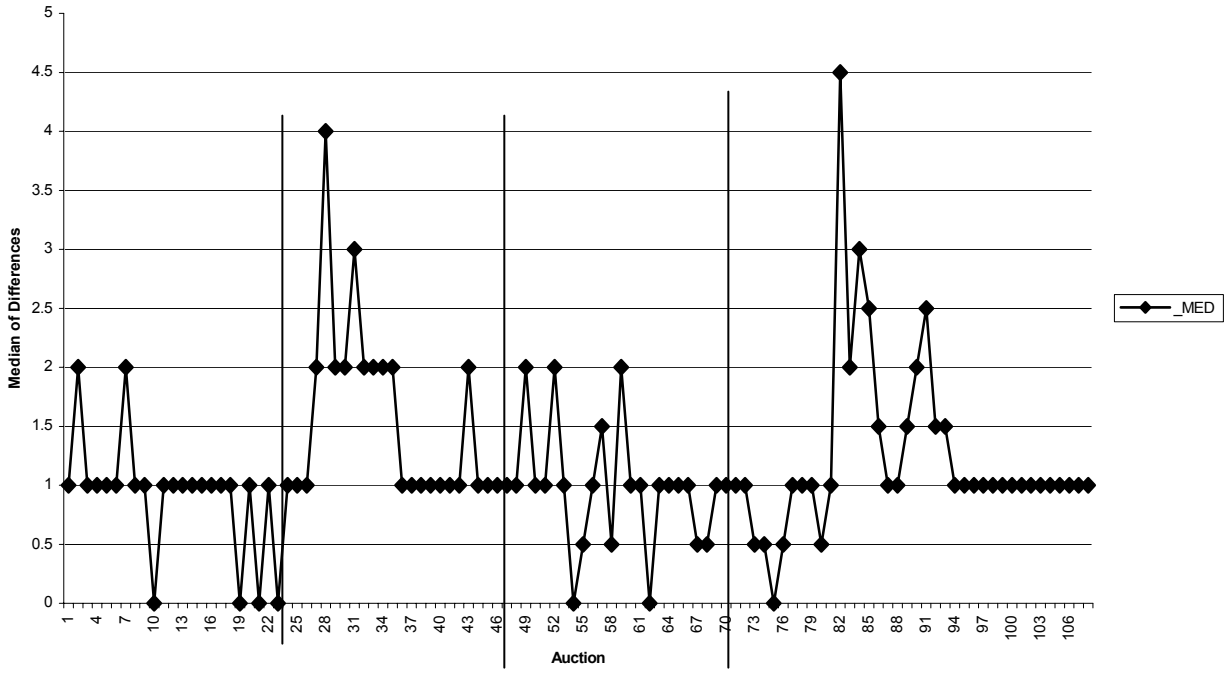

Figure 3H-10: Experiment 2 Median of Differences between Reserve Price and Bid for A3.

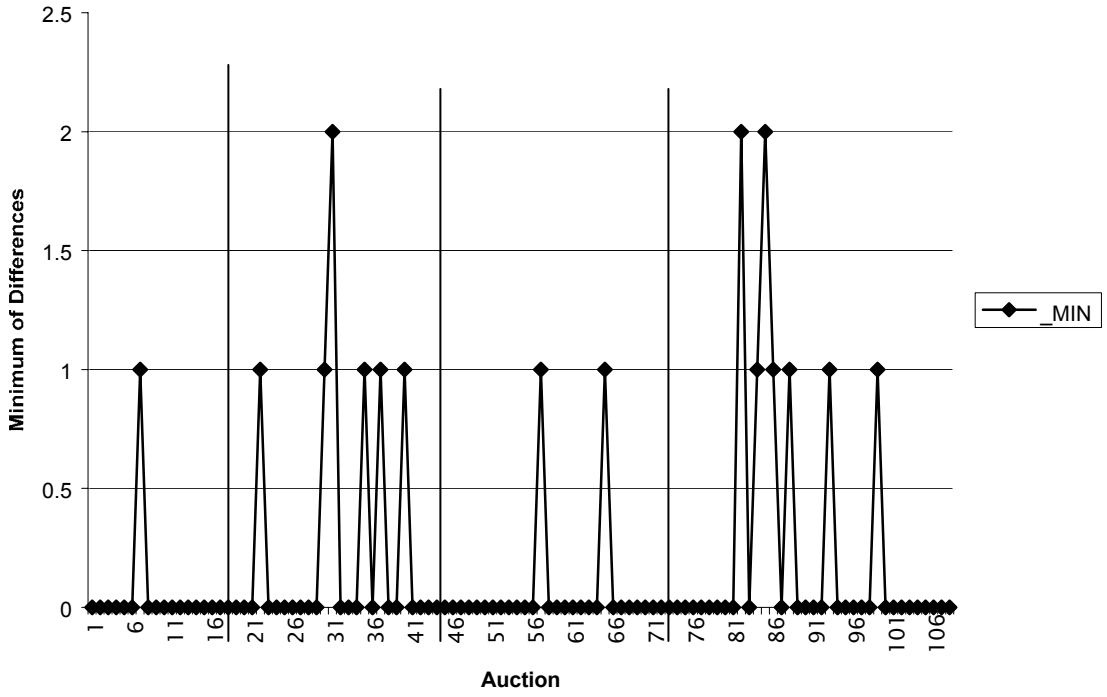

Figure 3H-11: Experiment 2 Minimum of Differences between Reserve Price and Bid for A3.

A comparison of the median bid between A1 and A2 indicates that in the initial auctions subjects were more likely to bid relatively lower on A2 than A1, however, the median bid for A3 was remained on average $1 Experimental above the subjects reserve. The subject focus on A1 can perhaps be attributed to relative weight of A1's reserve price. It may be the case that subjects perceive that attributes with higher reserve are given more weight in determining the auction winner. If subjects believe that A1 is the influencing attribute, then a winning bidding strategy of bidding relatively lower on A1 and higher on A2 and A3 may exist. In Experiments 1 and 2 subjects were not (de)briefed on possible winning bidding strategies nor did the experiments allow for experience bidders.

Submission of a bid equal to the reserve, the minimum, occurred more frequently in Groups 1 and 3. Subjects in Group 2 appeared to be less likely to submit a minimum bid on A1 and A2, but more likely to submit a minimum bid on A3. Of the three attributes, Group 2 was more likely to bid the minimum on A1 and A3, but in only one auction, auction 50, do we observe a minimum bid on A2. Overall Group 2 had higher earnings relative to the other groups. We postulate that the higher relative earnings of Group 2 and hence their bidding behavior may have resulted from a break in protocol during the training sessions.

The training sessions are designed to familiarize the subjects with the auction software and rules. For purposes of training, the buyer accepts the two highest ranked feasible bids, regardless of the bid value. In the Group 2 training session, the group as a whole submitted very high bids, with the experimenters announcing the earnings. The announcement of a high earnings may have induced the group to bid relatively more aggressively.

The combination of seller value and associated reserve price was established through successive testing of the auction parameters. The seller value-reserve price combination was set such that each seller in the auction had an equal opportunity of winning. In order to give sellers with low values an equal opportunity to win auctions the high value sellers were restricted to having higher reserve values.

The difference in composite bids to reserves by seller value for the composite bid is illustrated in Figure 3H-12. As expected, the greater the seller value the greater the difference in the composite bid, as the high value bidders attempt to exploit their high productivity. However, the relatively larger difference in the composite bid is observed only in the early auctions. Regardless of the seller value, over successive auctions the difference in composite bids begins to converge towards the reserve.

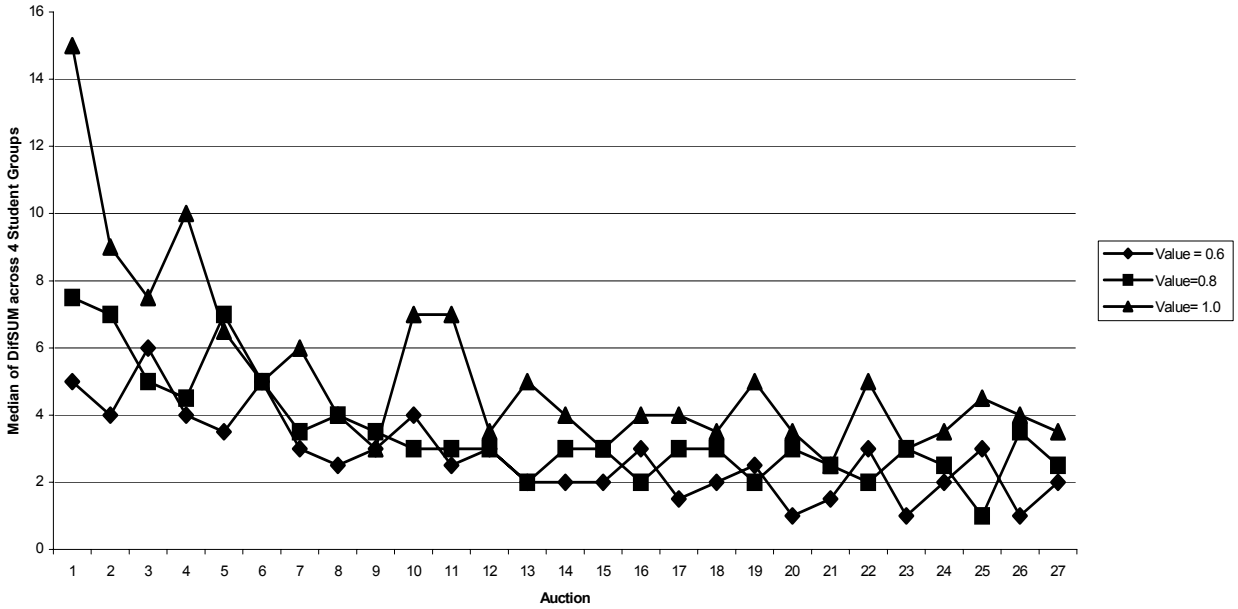

Figure 3H-12: Median of Difference of Composite Bid by Seller Value.

This research only discusses preliminary data analysis on the experiments performed to date. Two experiments, Experiments 1 and 2, were executed to observe bidding behavior in a multi-attribute auction setting over varying reserve prices and seller values. On average 7 subjects participated in each experimental session or 27 auctions. Interestingly, convergence of subject bids to individual reserve prices generally occurs within five auctions and even with as few subjects/bidders as seven.

While theoretically it can be shown that a first price open out cry auction quickly converges to subject reserve prices, we have shown that the first price sealed bid auction design addressed in this paper also quickly converges to the subject reserve. Interestingly, the rapid convergence occurs despite the fact that information used in the current auction is based on past auction winning bids. Future experimentation outlined in Annex B will provide further insight with respect to consistent performance of the auction designed discussed in this research.

3H.7 OTHER COMPLETED EXPERIMENTS

Testing of additional hypothesis, listed below, has been completed. Final results and conclusions are not discussed in this review.

1) Introduction of asymmetric information into the auction game. Asymmetric information will be introduced into the market by randomly providing one-half of the subjects with information as to the buyer's reserve, while this information is unknown to the remaining sellers. As the intent of the auction experiments is to consider Navy constraints, we propose to examine the effect of asymmetric information when bidding rounds are unlimited as compared to announcing limited bidding rounds.
2) Under the current experimental design, subjects participated in a 3-attribute auction environment. Relative efficiencies and bidding behavior may differ as a function of the attribute number. We propose to examine the effect of bidding behavior in one- and two-attribute auctions.

3) In previous experiments the number of job openings has been unknown to the subjects. Bidding behavior may be affected if the sellers know the number of job openings.

4) In earlier experiments the seller weight or the distribution of seller weights across subjects is unknown to the auction participants. We propose to examine bidding behavior when, at a minimum, the distribution of the seller weights are known.

3H.8 CAFETERIA STYLE COMPENSATION PACKAGES

The standard compensation package offered to military personnel is fixed. There has been substantial discussion on implementing a flexible style benefits package plan, whereby the Department of Defense can set the overall level of compensation and the member chooses from a menu of non-cash and cash benefits. A benefit to offering a flexible compensation package is that 'choice' promotes increased job satisfaction and thereby retention.

Critics of flexible benefits package plans cite potential high administration costs and that too many choices could lead to sub-optimal choices and hence job dissatisfaction. In a study supported by Navy Personnel Research, Studies, and Technology, results indicate that in complex multi-choice environments, individuals, on average, make optimal choices. The research on decision-making in complex environments lends support to the use of flexible compensation plans.

3H.9 DISCRETE CHOICE IN MULTI-ATTRIBUTE DECISION MAKING ENVIRONMENTS[15]

Managers and employers use an array of rewards to attract and retain quality employees. An increasingly significant component of the overall compensation is the employee's benefits package. Flexible benefits plans have the potential of increasing job satisfaction (White, 1983), thereby increasing the likelihood of retaining existing employees. Additionally, the labor supply pool from which the employer can draw may increase if prospective employees place a greater relative value on a flexible compensation package.[16] Nevertheless, in order for the flexibility to result in greater job satisfaction, employees must **perceive** the new plans as more valuable than the existing ones and this requires that employees are able to make optimal choices from the menu of benefits. Increasing the number of options and the flexibility of combinations increases the subjective value of the compensation package in so far as the employee is able to effectively decide on the optimal combination of benefits.

If the choice problem facing the employee becomes too complex, it is possible that the subjective value of the compensation package actually decreases with the number of options. Selecting an optimal benefit plan can be difficult because by nature the attributes of such plans are both multiple and discrete. These two factors

[15] See *Managerial and Decision Economics*, 2006, for full article.

[16] While a flexible compensation scheme may decrease the employer's labor cost, it could increase other costs. Offering a large number of benefit combinations would likely impose substantial monitoring and administrative costs on the employer. Further, employees may attribute very little added value beyond the *kth* option (Bucci and Grant 1995). The employer, therefore, has an incentive to offer just enough options and associated attributes such that the marginal cost of offering an additional option is just equal to its marginal benefit.

complicate the employee's choice and may push the limits of his or her cognitive abilities.[17] The outcome may be a high degree of dissatisfaction with the plan, leading to reduced job satisfaction and job tenure; the benefit package has not been successful. In this paper we report on an experimental investigation of individual decision-making over complex (multiple and discrete) options broadly designed on the choice of an employment benefits compensation packages. Our research is motivated by recruitment and retention issues facing managers as they compete in an increasingly sophisticated labor market.[18]

Many decision scientists argue that individuals have computational or cognitive limits, and that these limits become more critical when decisions involve difficult choices. Individuals may cope by adopting simplifying techniques, known as heuristics. Following Simon's (1955) seminal work, different schools of thought have emerged ranging from the belief that human decision making is intrinsically prone to errors to the belief that it is fundamentally efficient.[19] Three strands of research address decision-making in complex settings (see von Winterfeldt and Edwards, 1986). The first investigates individual assignment of values for commodities with multiple attributes. This strand is largely normative, as the researcher constructs a decision aid to facilitate the "correct" decision. The second strand studies whether individuals are capable of making payoff-maximizing decisions in the presence of multiple attributes. The third looks at whether or not individuals are able to consistently make payoff-maximizing decisions with multidimensional decisions and meeting multiple constraints. The second and third strands are more positive as they assess the ability of individuals from observations of unaided behavior.

Our research focuses on the second issue. The experimental literature in economics and psychology has paid little attention to problems of complexity arising from choices over discrete, multi-attribute objects (McKinney and van Huyck, 2000). To make optimal choices over multi-attribute and discrete alternatives, an individual must, in effect, solve an integer-programming problem. Using data generated through laboratory experiments, we examine the choices of individuals when faced with sets of discrete multi-attribute goods. We vary the relationship between the relative values of the attributes, and the value of an outside (simple) option, in effect varying both the complexity and the potential value of the options package. The main results are that (a) individual subjects respond to the relative trade-off between the attributes and (b) the majority of subjects adopt heuristics that approximate the optimal solution to the programming problem. Further, individuals appear to value the right to make choices since our subjects rarely choose the fixed payoff option,

---

[17] These two factors also complicate the employer's decision on what combinations to offer to the employee. Here, we focus on the employee's decision-making task.

[18] These are especially significant issues for the United States Department of Defense (DOD) and the U.S. Navy in particular. During FY97, the Navy encountered, for the first time in its history, problems meeting recruitment and retention goals (in FY03 this was less of a problem due, in part to worsened economic conditions). To address these issues, the Navy uses pecuniary incentives, which include numerous enlistment and re-enlistment incentives. But the overall compensation package includes non-pecuniary benefits such as dental and health care. Introducing flexible benefit plans may help solve staffing issues, as it would allow employees to tailor the plan according to their own preferences. Under such a scheme, the employer (Navy) chooses the level of compensation and the employee (a sailor) chooses the mix of salary and benefit options in the compensation package. Primarily because of institutional constraints, the Navy has typically ignored this option. The existing benefits package provided to military personnel by the DOD is standardized and relatively inflexible. Currently DOD is prohibited from offering non-standard packages to specific groups.

[19] The literature expressing the view that human decision makers are prone to errors has received recent attention, with reviews in Camerer (1995) and Starmer (2000). Prospect Theory, introduced by Kahneman and Tversky (1979), and theory of Framing Effects (Kahneman and Tversky, 1984), are particularly popular among behavioral economists and psychologists as alternatives to the neo-classical optimization approach. An alternative decision theory view, that human heuristic decision-making is fundamentally efficient, is based on a concept of ecological rationality, which is promoted by Gigerenzer (1996) and Gigerenzer et al. (1999). According to this view, many decision heuristics have evolved through time, affected by the successes and failures of individual decision makers.

with known payoff and low decision cost, even when the fixed payoff is up to 80% of the maximum possible under the complex decision task. (The 80% maximum possible under the fixed option refers to the reward a subject would receive prior to the "fixed deduction." If you take into consideration the fixed deduction then in fact the reward to choosing the fixed payoff option is significantly less than 80% of the reward to search – may want to keep this in mind).

3H.10 THE DECISION MAKING TASK

In this section we present our decision-making task and discuss how it represents a stylized version of an optional benefits package. In this experiment, the subject earns money by either playing a "cell selection" game or by choosing a fixed payoff in lieu the game. In the cell selection game, the subject's choice set is discrete and represented by an $n \times m$ matrix. (Just a thought using the notation n x m implies a rectangular matrix, when in fact in all cases (test cases and game matrices) we used square matrices, may want to say n x n or m x m to indicate the use of a square matrix.) Each cell of the matrix has three attributes: a **cell payoff**, a **cell weight** and a **cell value**. The cell weight and cell payoff are fixed, at the same level for each cell in the matrix, while the cell value is different for each cell. When a subject selects a cell, her[20] reward is the cell payoff plus the product of the cell weight and the cell value. Thus the subject's reward from selecting k cells is the sum of the rewards from those cells:

$$\text{Reward from selecting } k \text{ cells} = \sum_{i=1}^{k} [\text{Cell Payoff} + (\text{Cell Weight} \times \text{Cell Value}_i)] \qquad (1a)$$

$$= k \times \text{Cell Payoff} + [\text{Cell Weight} \times \sum_{i=1}^{k} \text{Cell Value}_i]. \qquad (1b)$$

Additionally, the subject has a **value limit** that constrains her choices. She can continue to select cells as long as the sum of the cell payoffs from the cells she has selected does not exceed her value limit. Specifically, if the subject selects k cells, then her selections must satisfy:

$$\sum_{i=1}^{k} \text{Cell Value}_i \leq \text{Value Limit} \qquad (2)$$

Thus the decision task facing a profit-maximizing subject is a constrained optimization problem where the objective function is equation (1a) or (1b) and the constraint is equation (2). In the context of neo-classical consumer theory, the selected cells are analogous to the market basket, and the value limit is analogous to the budget constraint – but the goods are discrete, not continuous, and so optimization at the margin is not strictly available. In effect, the subject must solve the "knapsack problem" (Greenberg, 1971). While this type of problem may be intuitively simple, the process of identifying the optimal solution can be extraordinarily complex, requiring an integer -programming algorithm.[21]

Figure 3H-13 shows an example of the computerized interface viewed by our experimental subjects. All of the decision matrices we presented to subjects were selected from the set of problems where integer

[20] For simplicity, we refer to a generic subject with a female pronoun. Both male and female subjects participated in the experiment.

[21] The experimenter interface of our computer program includes a choice of algorithms for calculating the optimal solution; we used the branch-and-bound method (see Garfinkel and Nemhauser, 1972 or Parker and Rardin, 1988). We calibrated the program's calculations with a separate exhaustive search algorithm.

r-programming algorithms calculate an exact optimum. The box labeled "You could earn this or more" in the upper middle of Figure 3H-13 identifies this optimum; the "or more" designation is for the more generic cases where integer-programming algorithms provide only an approximate solution. We explicitly told subjects that the number shown in the box was the maximum they could earn (in experimental dollars) for the given matrix. One of our research questions is what (if any) heuristics subjects would utilize. We determined **a priori** that an essential piece of information they needed was the maximum possible earnings. Additionally, subjects need this information to decide whether or not to opt for the fixed payoff, to which we now turn.

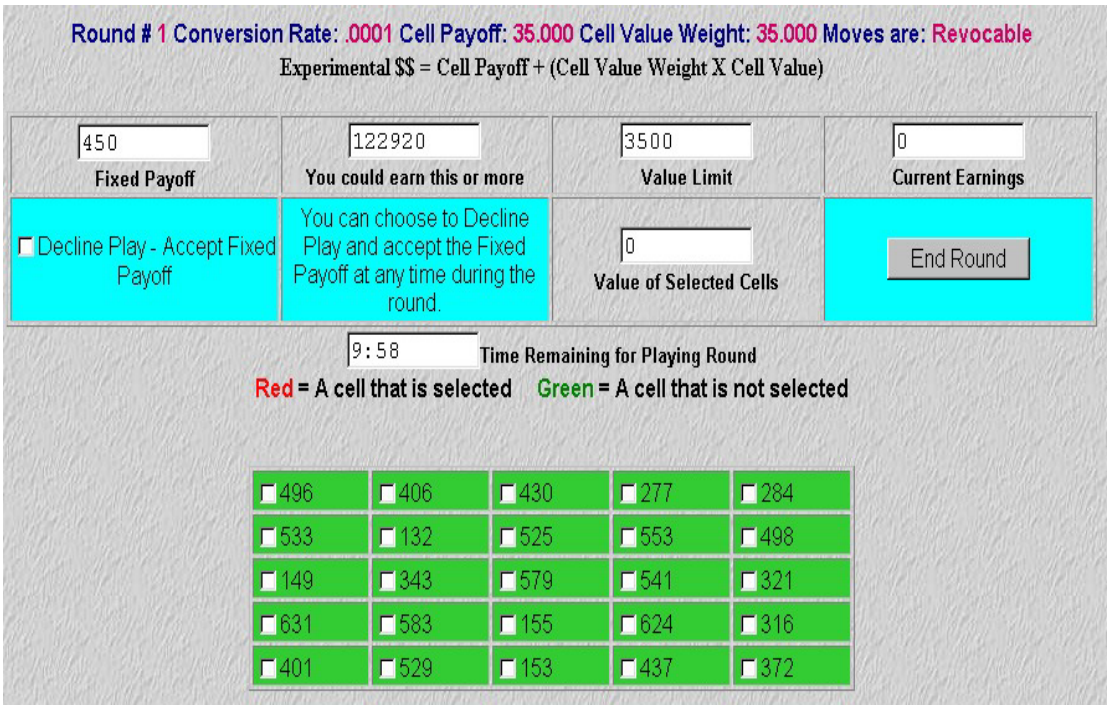

Figure 3H-13: Subject Screen Display.

As an alternative to selecting cells, the subject may choose the **fixed payoff** shown in the upper right hand corner of Figure 3H-13. That is, instead of "playing the game" where she is rewarded for selecting cells, the subject can "opt out" and receive the fixed payoff. So as to have a benchmark by which to decide whether or not to take a given fixed payoff, the maximum possible earnings from selecting cells is displayed on the subject's computerized interface.

Our decision-making task is designed to loosely represent the selection of components of a flexible benefits package. The matrix represents the benefits options offered by the employer. Each cell constitutes a component of the package, such as medical, dental and childcare policies, and the combination of cells that the subject selects is the package chosen by an employee. The value of a selected compensation package consists of two parts. First, the cell value is the dollar value of coverage for each policy while the cell weight determines the economic discount (or premium) at which the employee purchases that coverage. For example, in this experiment, we use a cell weight of 1.2. In order to receive a reward of 100 from selecting a cell (ignoring for the moment the cell payoff), the subject need only chose a cell with a cell payoff of 83. Alternatively, $83 of expenditure translates into $100 of benefits coverage for the employee. Second, in addition to the value of the coverage, employees place values on coverage, independent of the amount, and this is represented with the cell

payoff. The total reward in equation (1a) represents the value that the employee places on the package she has selected. The value limit represents the total amount of coverage that the employer is willing to offer the employee; it is the employee's benefit "budget." Finally, the fixed payment option represents the monetary value the employee places on an alternative non-flexible benefits package.

3H.11 HEURISTICS

Given the computational complexity of the optimal solution, we expected subjects to develop heuristics or "rules of thumb" to simplify the decision-making task. The effectiveness of a particular heuristic can be evaluated relative to the optimal solution as well as relative to other heuristics.

Here we identify three heuristics that stand out because of their (apparent) simplicity. These heuristics assume that a subject chooses to focus on a range of cell values: high (H), medium (M) or low (L). Simple versions of these heuristics often leave a substantial amount of unspent budget (i.e., the sum of the selected cells is substantially less that the value limit). Thus we expected subjects to use slightly sophisticated versions. As described in 3H.13 below, our cell values range from a minimum of 100 and a maximum of 1000:

- **Heuristic H.** Select cells with values in the 700 – 1000 range that (nearly) exhaust the value limit, and if the value limit is not reached, select one or two cells with low cell values to (nearly) reach the value limit.
- **Heuristic M.** Select cells with cell values in the 350 – 700 range that (nearly) exhaust the value limit, and if the value limit is not reached, select cells with high or low cell values to (nearly) reach the value limit.
- **Heuristic L.** Select cells with cell values in the 100 – 350 range that (nearly) exhaust the value limit, and if the value limit is not reached, select one or two cells with higher cell values to (nearly) reach the value limit.

We conducted "simulations" utilizing the actual matrices used in the experiment to obtain number-of-cells-selected and earnings estimations for each of the above heuristics. [22] Based on these simulations, heuristic H results in the selection of two or three cells, heuristic M results in the selection of four or five, and heuristic L results in the selection of six or more (most often seven or eight cells). To the extent that heuristic H involves selecting the fewest cells, it has the lowest decision cost of the three. Heuristic L involves selecting the greatest number of cells, so arguably it has the highest decision cost. While heuristic M focuses on a specific range, it has equal emphasis high and low values outside the range, so in that sense it is the most flexible of the three and has the intermediate decision cost.

Because we are interested in what heuristics subjects might employ when facing a computationally difficult problem, we utilize decision matrices that have several cells in the optimal solution. The subjects in this experiment view a total of thirty-eight matrices (see 3H.13 below), and thirty-six have optimal solutions comprised of six or more cells. Thus the experiment may appear biased towards heuristic L, as that heuristic selects six or more cells – but this is by design, as we want to see if subjects can identify a useful (i.e., profitable) heuristic in a complex setting. While heuristic L has the best **a priori** chance of being the

[22] We conducted our simulations using a spreadsheet with cell values ranked in one column from high to low (versus the randomly populated matrix 5 x 5 matrix viewed by subjects) for each of the thirty-eight decision matrices. This alternative format makes it easier to identify the most profitable choice(s) under any of the three heuristics, and our analysis did not face a time constraint, as did our subjects. Our number-of-cells-selected estimations and earnings estimations are meant as benchmarks for comparison, rather than behavioral predictions **per se**.

most profitable because of its close alliance with the optimal solution, it has the highest decision cost. Our simulations suggest that heuristic H, with the lowest decision cost, should yield of over 90% of the maximum possible earnings and heuristic M should yield around 95%. Thus it would not be surprising to see subjects adopt these less onerous rules.[23]

3H.12 THE EXPERIMENT

3H.12.1 Design

The experiment is a 2 × 2 design with the cell payoff and fixed payoff option as the treatment variables. The cell payoff is either 20 or 100. Holding the cell weight constant, the higher the cell payoff the less the relative contribution of the cell value to the reward from selecting a given cell (see equation (1a) above). The fixed payoff is either 80% or 50% of the maximum possible reward from selecting cells; both the maximum payoff and the fixed payoff are displayed to the subject (see Figure 3H-13 above). The lower the fixed payment percentage, the higher the opportunity cost of foregoing the option to select cells.

The experimental treatments allow us to examine two important aspects of this decision-making task. First, we can examine how the relative importance of the cell value versus the cell payoff (holding cell weight constant) affects the subject's choice of cells. This corresponds to an employee's choice between having a particular benefit in the package (a cell's reward is determined primarily by the cell payoff) versus the amount of the coverage (a cell's reward is determined primarily by the cell value).

Second, we can examine how variations in the fixed payoff affect the subject's decision to select cells versus taking the fixed payoff. Our cell selection game represents a flexible package and the fixed payoff represents a fixed package. Fixed benefits are generally less than those under a flexible package, (if the statement 'fixed benefits are generally less than those under a flexible package' refers to the experimental design then fine. However, if this statement is meant to refer to real world benefits plans, I am not certain that one could argue that fixed plans are generally less than flexible plans at least to the employee in terms of real money. Flexible plans are offered because either (1) the employer can receive some tax savings (I believe employer taxes/federal/state contributions are greater on wages relative to say medical benefits) if employees on average chose more medical vice salary in the their compensation packages or the employer saves on recruiting costs if flex plans are a retention tool and (2) from the employees point of view the monetary value of the two types of plans may be the same, but the employees combined explicit and implicit value of a flex plan may be greater as he derives utility from choice. Bottom line is not sure you can make an absolute statement that money value of fixed plan is less than money value of flex plans, but a flexible package requires that the beneficiary incur the subjective cost of identifying the preferred package. If the subjective cost is high,

[23] A fourth but somewhat complicated heuristic k' is based on information displayed on the subject interface (see Figure 1). If a subject can infer equation (1b) from the version of equation (1a) given in the instructions, she can substitute the maximum possible earnings for "the earnings from selecting k cells" in (1b), treat equation (2) as an equality and substitute the value limit for the sum of the cell payoffs in (1b), and then solve for k to get:

$$k' = \begin{matrix}\text{Optimal}\\ \text{number of}\\ \text{cells to select}\end{matrix} = \frac{\text{Max Possible Earnings} - (\text{Cell Weight} \times \text{Value Limit})}{\text{Cell Payoff}}$$

This k' calculation provides an excellent approximate (and when the optimal solution exhausts the value limit, an exact) count of the number of cells in the optimal solution. In our simulations, heuristics L and k' result in the same number of cells being selected in 35 of the 38 matrices. Given its complexity, we did not expect many subjects to use heuristic k'. Nonetheless, our experiment is not designed to distinguish between heuristics L and k'.

an employee might choose the fixed package with lower total benefit if the difference between benefits is relatively small. In our experiment, the subject incurs her subjective decision-making cost in the cell selection game; if that cost is high, she may opt for the fixed payoff.

Tables 3H-9 and 3H-10 report our experimental parameters. To avoid confusion with the cells of the game matrix, the four cells of the experimental design are referred to as "sessions" S1, S2, S3, and S4.[24] The subject completes all four sessions in one trip to the laboratory (see 3H-12B below). Table 3H-9 shows those parameters that are varied across sessions, and Table 3H-10 shows those that are held constant across sessions. As described below, various considerations resulted in our choice of experimental parameters. Table 3H-9 shows that session S1 has a cell payoff of 20 and fixed payoff percentage of 80%, S2 has 100 and 80%, S3 has 20 and 50%, and S4 has 100 and 50%, respectively. The fixed payoff is a percentage of the maximum possible earnings from selecting cells; the subject sees the amount of the fixed payoff each round, but is otherwise not informed as to how that amount is determined.

Table 3H-9: Experimental Parameters Varied across Sessions

	Session			
Parameter	S1	S2	S3	S4
Cell Payoff	20	100	20	100
Fixed Payoff Percentage	80%	80%	50%	50%
Number of Rounds	9	8	10	11
Fixed Deduction (US$)	$17.00	$18.00	$20.00	$27.00

Table 3H-10: Experimental Parameters Constant across Sessions

	Cell Values						
Matrix Dimensions	**Lower Bound**	**Upper Bound**	**Cell Weight**	**Value Limit**	**Revocable Moves?**	**Seconds per Round**	**Conversion Rate**
5 × 5	100	1000	1.2	2000	Yes	240	0.001

Each session is a series of decision-making "rounds" where the given cell payoff and fixed payoff percentage applies to each round. The subject sees a different decision matrix and fixed payoff each round. For example, session S1 has eight rounds, so the subject sequentially views eight different decision matrices, each with a cell payoff of 20 and a fixed payoff that is 80% of the maximum possible earnings from the selecting cells in the given matrix. The number of rounds per session is varied across sessions to control for "end of experiment" and/or "end of sequence" effects (see Davis and Holt, 1998).

Each subject completes all four sessions (i.e., thirty-eight rounds). As a control for order effects, we use four different randomly determined sequences. A total of 80 subjects participated in the experiment – 21 subjects

[24] In the experiment, subjects are told the session's labels are S17, S18, S19, and S20, respectively.

received the sequence S4-S1-S2-S3, 20 received the sequenced S1-S3-S4-S2, 19 received the sequence S3-S2-S1-S4 and 20 receive the sequence S2- S4-S3-S1.[25]

Table 3H-9 also shows a "fixed deduction" that varies with each session. At the end of each session, a deduction is subtracted from the subject's earnings. This deduction is explained to the subject as part of the instructions, and the amount of the deduction is revealed to the subject as she begins each session. This deduction serves to maintain salient incentives while at the same time keeping overall earnings within the research budget.[26]

Table 3H-10 shows the parameters that are identical in each of the four sessions. All decision matrices are square 5×5 matrices. This dimensionality has a sufficiently rich optimal solution while making the comparison of individual cells relatively easy for the subject. Both the combination and location of cells that comprise the optimal solution vary considerably each round.[27]

Each of the twenty-five cells of a given matrix is randomly populated with a cell value from the range 100 to 1000 (inclusive). We chose this range because the variety of choices is sufficiently rich while the computational difficulty is significant but not overwhelming. Our expectation was that the subject could develop a useful but non-trivial decision rule (possibly a heuristic) for deciding which cells to select.

The cell weight is fixed at 1.2 in all rounds of all sessions. The value limit is also constant across all rounds of all four sessions. In any given round, the subject may continue to select cells as long as the sum of the cell values that have been selected does not exceed 2000. We chose these values in conjunction with, and for similar reasons as, the matrix size and cell value range.

We allow subjects to make revocable choices when selecting cells (i.e., if a subject selects a cell by clicking on it with her mouse, and she can "deselect" the cell by clicking it a second time). She can deselect a cell at any time, even after selecting other cells, and she can select and deselect cells as many times as she wishes (provided time has not expired). One of our research questions is what (if any) decision rules are used by subjects, and we expected that at least some subjects would experiment before settling on a decision rule. Thus we included this deselect feature to increase the chance of identifying systematic strategies in the subjects' final choices of cells. The computerized interface allows the experimenter to choose whether or not the subject's cell selection decisions are revocable.

Each round lasts up to four minutes (240 seconds). When a subject is finished making decisions in a round, she mouse-clicks either the "accept fixed payoff" button to take the fixed payoff, or an "end round" button to indicate that she is done selecting cells. If she does not click one of these two buttons before time expires,

[25] We intended to randomly assign twenty subjects to each sequence. After completing the experiment, we realized that one subject was inadvertently given sequence S4-S1-S2-S3 instead of S3-S2-S1-S4.

[26] Subjects typically earned \$10 – \$20 for participation in a two-hour session. Without the fixed deduction, a subject would earn around \$60 simply by choosing the fixed payoff each round. Although she could earn an additional \$5 – \$10 selecting cells, she might very well opt for fixed payoff because that option has both a substantially lower time cost and substantially lower decision cost. Other methods of achieving the desired earnings levels would have involved modifying parameter values in ways that would have reduced incentives For example, small cell weights and cell payoffs may lead a subject to interpret equation (1a) as essentially zero, irrespective of the cell values.

[27] Based on pre-testing, 4×4 or smaller matrices often have trivial or easily identifiable solutions, and 6×6 or larger matrices have excessively burdensome solutions so that subjects almost always opt for the fixed payoff.

the subject's reward for that round is zero.[28] This time limit does not appear to be binding – only 23 (0.8%) of the 3040 total rounds have a zero payoff (i.e., end before one of the two buttons is clicked).

The subject's rewards are expressed in "experimental dollars" or E$, which are converted into U.S. dollars or US$ at the rate of E$1 = US$0.001. The computer interface automatically converts the subject's experimental rewards into U.S. currency at the end of each session. The subject then records her US$ earnings from the screen onto a paper record sheet, and then subtracts the appropriate fixed deduction. We chose this conversion rate in conjunction with our other parameters, so as to yield an expected payout in the $15 – $20 (US) range per subject for participation in a (roughly) two-hour period. (See footnote #26 – In the footnote you state 'subjects typically earned $10 – $20 for a two hour session. In the above sentence you use the word 'expected' payout. A question might arise as to what explains the difference between expected payout and actual payout.)

3H.12.2 Procedure

We conducted the experiment at the Mississippi Experimental Research Laboratory (MERLab) at the University of Mississippi and at the Business, Economics, Accounting and Marketing Laboratory (BeamLab) at the University of South Carolina. The computerized application is comprised of an experimenter interface and a subject interface, which includes computerized instructions. The subjects were recruited from undergraduate business courses at the respective universities. Participation by the subject involved two visits to the laboratory, called Part I and Part II. Part I is instructional training that lasts approximately one-and-a-half hours. Part II is the data-collection period, and lasts approximately two hours. The subject was paid her earnings from both parts upon completion of Part II.[29] This included a $6 participation fee for each part.

The training in Part I familiarized (or familiarizes?) the subject with the computerized interface, record keeping, and the sequence of events (multiple sessions, fixed deductions, etc.). Upon arrival, the subject is assigned a subject number, given a blank earnings record sheet, and seated in a private computer carrel. She completes the computerized instructions and two practice rounds with 2 × 2 matrices. An experimenter then visits her carrel, and she is prompted for questions and/or clarification regarding the task. After any discussion, the subject is presented with some brief printed instructions and a consent form, and given a few moments to review that material. After she signs the consent form, an experimenter writes the first session number and corresponding fixed deduction on her printed record sheet. While the experimenter watches, the subject restarts the program and enters the session and subject number. The subject then proceeds through all rounds of the given session. After completing the session, the subject raises her hand, receives another session number and fixed deduction, and restarts the program while the experimenter watches. This process continues until all Part I sessions are completed. After the final Part I session, the subject totals her earnings on the record sheet, signs up for Part II (a least one day later, and no more than seven), and is excused.

The decisions from Part II are used in the data analysis. Upon returning for Part II, the subject is again seated at a private computer carrel and given her record sheet from Part I (she uses the same subject number both parts). The Parts I and II the procedures are identical[30], but the Part II parameters and decision matrices are different from those used in Part I. Upon completion of the final session of Part II, the subject is privately paid

[28] The time remaining in the round is shown on the subject's screen (see Figure 3H-13). Also, the "zero payoff if no decision is made" was reiterated in the subjects' instructions.

[29] This payment schedule is explicitly explained to subjects when they are recruited.

[30] Subjects were explicitly recruited for participation in two parts, with each part on a separate day. The consent form is signed prior to participation in Part I. Prior to participation in Part II, each subject is reminded that she signed the consent form during her previous visit, and offered an opportunity to review the form at her request.

her cash earnings and excused. In our experiment, cash earnings in Part I average $14.35, with a low of about $4 and a high of about $25. In Part II, the average is $20.67, with a low of about $3 and a high of about $28 (exclusive of the participation fees). (This statement seems to be contradictory to footnote #26 where we state that subjects earned $10 – $20 for participation in a 2-hour session).

3H.13 RESULTS

3H.13.1 V.A. Summary

Our eighty subjects completed a total of 3040 rounds in Part II of the experiment. Table 3H-11 reports summary data on the round-level outcomes, both overall and by individual session. In those rounds where the subjects choose the cell selection game, earnings are calculated as a percentage of the maximum possible reward for the respective round. The table shows the percent of rounds where subjects chose the fixed payoff and the percent of rounds that fall into a given earnings category. Rounds where a subject "times out" (i.e., fails to click either the "choose the fixed payoff" button or the "end round" button before time expires) is categorized as 0% payoff. We make four general observations based on the data shown in Table 3H-11:

- First, subjects predominantly choose to play the cell selection game, as the fixed payoff option is chosen in only 4% of the rounds.[31] However, the fixed payoff is chosen significantly more often (7% v. 1% of the time) in sessions S1 and S2 where the fixed payoff is 80% of the maximum possible earnings, than sessions S3 and S4 where the fixed payoff percentage is 50%. This is consistent with a decision-cost model of behavior, and with our expectations, although we are somewhat surprised it only occurred in 7% of the rounds where the fixed payoff percentage is 80%. (Why would you be surprised by '7%'? This statement implies that you would expect to observe subjects' frequency of fixed payoff to be higher. Do you need to justify why you were 'surprised' at only 7%?)
- Second, the 240-second time limit is not binding, as subjects "time out" in only 1% of the rounds, and the occurrences are roughly evenly dispersed across sessions.[32]
- Third, when subjects choose to play the game, they do quite well. In almost half (48%) of the rounds earnings are in the 97 – 99% range, and in 85% of the rounds earnings are 90% or more of the maximum possible. Inspection of the individual data reveals that low earnings are confined to a subset of subjects.[33]
- Fourth, although subjects consistently earn over 90% of the maximum possible, there are differences across the cell value treatment. In sessions S1 and S3 where the cell payoff is 20, subjects earn

[31] Inspection of the individual data reveals that the fixed payoff choice is confined to less than a third of the subjects (24/80). By session, the fixed payoff was chosen 47 times in S1 by ten different subjects, 46 times in S2 by eighteen different subjects, 10 times in S3 by six different subjects, and 4 times in S4 by three different subjects.

[32] Inspection of the data reveals that twenty different subjects time out across all four sessions. Three of those subjects time out twice, two of them in two separate sessions and one of them twice in the same session (S2). By session, time outs occur five times in S1, seven times in S2 (by six different subjects), four times in S3 and seven times in S4. The maximum number of times this occurs for any one subject is twice; there are three such instances and only one of those is occurs in the same session (7407, S3). Other summary data on the zero payoffs: Five of the zero payoffs are in S1, seven in S2, four in S3 and seven in S4. Thus 61% (14/23) occur when the cell payoff is 100. Fifty-seven percent (13/23) occur in either rounds 1, 2 or 3. By random-order sequence, four zero payoffs occur in the first random sequence, five in the second sequence, six in the third, and eight in the fourth.

[33] The 339 rounds with earnings less than 90% are confined to less than half of the subjects (38/80), and 208 of those rounds are confined to 18% (14/80) of the subjects. The 41 rounds with earnings below 80% are confined to six subjects, and two subjects account for 32 of those 41 rounds.

97 – 99% of the maximum possible in 60% and 69% of the rounds, respectively, and they rarely earn less that 90% of the maximum (2% of the time in each session). This contrasts with sessions S2 and S4 where the cell payoff is 100. There, subjects earn in the 97 – 99% range in only 31% and 30% of the rounds, respectively, and earn less than 90% in 18% and 24% of the rounds, respectively. In general, when the cell payoff is 20 earnings tend to be higher and less disperse than when the cell payoff is 100.[34] A relatively low cell payoff implies that the cell value is more important in determining the reward from selecting a particular cell. Perhaps this makes subjects concentrate relatively more on the cell value, and thus more closely approximate the optimal solution. Further discussion on subjects' strategies is presented below.

Table 3H-11: Summary of Round-Level Outcomes

	Experimental Parameters			Chose cell selection game and earned (% of maximum possible earnings)					
	Fixed Payoff Pct.	Cell Value	Chose fixed payoff	0% (timed out)	Less than 80%	80 – 89%	90 – 96%	97 – 99%	100%
Total	–	–	4%	1%	1%	10%	28%	48%	9%
n = 3040			(110)	(20)	(41)	(299)	(855)	(1451)	(264)
S1	80%	20	7%	1%	1%	1%	25%	60%	6%
n = 720			(47)	(5)	(4)	(5)	(181)	(433)	(45)
S2	80%	100	7%	1%	2%	16%	34%	31%	10%
n = 640			(46)	(7)	(11)	(99)	(216)	(198)	(63)
S3	50%	20	1%	1%	0%	2%	23%	69%	5%
n = 800			(10)	(4)	(0)	(12)	(180)	(553)	(41)
S4	50%	100	1%	1%	3%	21%	32%	30%	13%
n = 880			(7)	(4)	(26)	(183)	(278)	(267)	(115)

3H.13.2 Fixed Effects Estimation

We estimate a fixed effects (or dummy variable) regression for three different dependent variables: Earnings Ratio, Cell Ratio and Search Ratio.[35] Referring to these three dependent variables generically as Y, our fixed effects model is given by:

[34] But curiously, subjects are able to earn 100% the maximum possible in a greater percentage of the rounds in S2 (9.8%) and S4 (13.1%) than in S1 (6.3%) and S3 (5.1%).

[35] We examined random effects models as well but this estimation was rejected, on the basis of a Hausman test, in favor of the fixed effects model.

$$Y = \beta_0 + \beta_F \text{Fixed} + \beta_T \text{Timeout} + \sum_{i=2}^{4} \beta_{Si} \text{Session}_i + \sum_{j=2}^{11} \beta_{Rj} \text{Round}_j + \sum_{k=2}^{80} \beta_{Subk} \text{Subject}_k + \varepsilon \quad (3)$$

The three dependent variables and the right-hand side fixed effects dummy variables are summarized in Table 3H-12 and discussed below.

Table 3H-12: Variables in Equation (3) Fixed Effects Regression Model

Variable	Definition
Dependent Y	
Earnings Ratio	Subject's per-round earnings, as a percent of the maximum possible in the round
Cell Ratio	Number of cells in subject's final per-round choice, as a percent of the number of cells in the round's optimal solution
Search Ratio	Total number of cells subject selects per round (including those not part of subject's final choice), as a ratio of the number of cells in subject's final choice for the round
Fixed Effects	
Fixed	= 1 if subject chooses the fixed payoff option in the round = 0 otherwise
Timeout	= 1 if time expires before subject is finished in the round = 0 otherwise
Session $_i$ i **= 2, …,4**	= 1 if Y observation from session Si = 0 otherwise
Round $_j$ j **= 2, …,11**	= 1 if Y observation from round Rj = 0 otherwise
Subject $_k$ k **= 2, …,80**	= 1 if Y observation from subject Sub$k$ = 0 otherwise

The first two dependent variables measure the subject's performance relative to the optimal solution. The Earnings Ratio measures the subject's per-round earnings, as a percentage of the maximum possible earnings in the given round of the given session. The Cell Ratio measures the number of cells in the subject's final choice each round, as a percentage of the number of cells in the optimal solution for the given round of the given session. The number of cells in the subject's final choice is the number of cells that the subject has selected when she clicks the "end round" button.

The Search Ratio measures the search activity of the subject while controlling for the heuristic used by the subject. The numerator of this ratio is a count of **all** cells that the subject selects during the course of a round, including those that the subject subsequently deselects (i.e., those that are not part of the subject's final choice). Recall that with the revocable choice option, the subject is able to search out and try different combinations of cells before clicking the "end round" button. The denominator of Search Ratio is the count of the number of cells in the subject's final choice in the round. Recall from 3H-12 that the subject's heuristic can be categorized according to the number of cells that in the subject's final choice. Thus this measure normalizes search activity so that it is comparable across subjects.[36]

The 94 independent variables are fixed effects dummy variables, as shown in Table 3H-12. The Fixed and Timeout variables control for those rounds where, respectively, the subject chooses the fixed payoff option or she "times out" (i.e., fails to click either the "choose the fixed payoff" button or the "end round" button before time expires). If Fixed = 1, then the Earning Ration equals the amount of the fixed payoff (divided by the maximum possible for the round) and both Cell Ratio and Search Ratio equal zero. When Timeout = 1, all three dependent variables have a value of zero. The three Session dummy variables control for individual session effects, and are used to test for the effect of the experimental treatments outlined in Table 3H-12 above. The ten Round dummy variables control for individual round effects or variation over time, such as learning. The seventy-nine Subject dummy variables control for a variety of subject-specific effects, including (but not limited to) risk aversion.[37]

Table 3H-13 presents the hypothesis tests from the fixed effects regressions. The models fit the data fairly well, with adjusted R^2s of 0.85, 0.77 and 0.49 for the Earnings Ratio, Cell Ratio and Search Ratio, respectively. There is strong evidence of a session or treatment effect in the Earnings Ratio and the Cell Ratio regressions ($p < 0.001$ and $p = 0.002$, respectively, in the Session Effect row); there is marginal evidence of a session effect in the Search Ratio ($p = .087$). There does not appear to be a round effect in the Earnings Ratio ($p = .260$ in the Round Effect row), but there are significant round effects in both the Cell Ratio and Search Ratio ($p < .05$ for each). As we expect, there are strong subject effects in all three ratios.

[36] Consider the following example. Two subjects both have three cells in their final selection so that their decisions are consistent with heuristic H (see above). Suppose that the second subject selects and then deselects two additional cells in the process of making her final choice (i.e., she searches a total of five cells while the first subject only searches three). The first subject has a Search Ratio of 3/3 = 1.0, and the second has a ratio of 5/3 = 1.67, thus the subject who searches more has a higher Search Ratio. Now consider a third subject who has five cells in her final choice but does not deselect any cells during the round. This subject, whose final decision is consistent with heuristic M, has a Search Ratio of 5/5 = 1.0. Thus according to the this measure of search activity, the first and third subjects search an equal amount, after accounting for their different heuristics, and the second subject searches more.

[37] A dummy variable for session sequence is not included as the order of sessions is varied across subject with one of four random sequences (3H.13) in order to control for sequence effects.

Table 3H-13: Hypotheses Tests from Fixed Effects Regressions

Test	Dependent Variable		
	Earnings Ratio	**Cell Ratio**	**Search Ratio**
Overall Model	$R^2_{Adj} = 0.85$	$R^2_{Adj} = 0.77$	$R^2_{Adj} = 0.49$
H_0: $\beta_{S2} = \ldots = \beta_{Sub80} = 0$	F = 196.9	F = 111.2	F = 29.7
	$p < .001$	$p < .001$	$p < .001$
Session Effect	F = 174.0	F = 5.07	F = 2.19
H_0: $\beta_{S2} = \beta_{S3} = \beta_{S4} = 0$	$p < .001$	$p = .002$	$p = .087$
Estimated Coefficients			
S1 β_0	1.007	0.989	1.291
S2 $\beta_0 + \beta_{S2}$	0.978	0.998	1.247
S3 $\beta_0 + \beta_{S3}$	1.011	1.012	1.258
S4 $\beta_0 + \beta_{S4}$	0.971	0.998	1.249
Round Effect	F = 1.24	F = 4.37	F = 2.13
H_0: $\beta_{R2} = \ldots = \beta_{R11} = 0$	$p = .260$	$p < .001$	$p = .020$
Subject Effect	F = 28.1	F = 70.9	F = 13.6
H_0: $\beta_{Sub2} = \ldots = \beta_{Sub80} = 0$	$p < .001$	$p < .001$	$p < .001$

Note: n = 3040 for each regression.

Table 3H-13 also shows the individual coefficient estimates for the Session fixed effects. For the Earnings Ratio, the nature of the significant treatment effect is apparent. Sessions S1 and S3 are virtually identical; they both have a cell payoff of 20. Estimates for S2 and S4 are substantially less, and those two sessions have a cell payoff of 100. However, all four sessions have estimates are reasonably close to 1.0 (or 100%). The estimated coefficients are consistent with data in Table 3H-11, where most of the rounds have earnings of 90% or higher, but more are in the 97% – 100% range in S1 and S3 than in S2 and S4. Collectively, Tables 3H-11 and 3H-12 suggest that a lower cell payoff, and thus a relatively higher importance of cell value, results in higher earnings. The coefficient estimates in Table 3H-13 also suggest that S4 has (slightly) lower earnings than S2. In S4, the fixed payoff was 50% of the maximum possible, while in S2 it is 80%. Recall from Table 3H-11 that the fixed payment option is chosen in 7% of the rounds in S2, but is chosen in only 1% of the round in S4s. One interpretation is that 80% fixed payment option provides a cushion in those rounds where the subject does (relatively) poorly, but the 50% option does not provide such a cushion. However, no such effect is apparent in the comparison of S1 and S3, which have 80% and 50% fixed payment options, respectively.

The estimated coefficients for the Cell Ratio tell a very different story than those of the Earnings Ratio. The Cell Ratio also exhibits a significant treatment effect, but for (or 'but with respect to this ratio') this ratio, S1 and S3 appear dissimilar – of the four Cell Ratio estimates, S1 has the lowest and S3 has the highest – while S2 and S4 are virtually identical and lie between S1 and S3. This is somewhat paradoxical: S1 is the

furthest from the optimal solution in terms of the Cell Ratio, but it is the **closest** to the optimal solution in terms of the Earnings Ratio. We note that while there is a significant treatment effect, all four sessions have coefficient estimates close to 1.0 (or 100%). (It might be important to note that in a handful of a cases the subject found (submitted in his/her final solution) a combination of cells that actually received a higher payoff than the optimal payoff posted. This would affect the earnings ratio and cell ratio for that subject).

The estimated coefficients for the Search Ratio suggest that the search activity is higher in S1 and S3 than in S2 and S4, but the marginal session effect (p-value = .087) appears to be due to the difference of S1 from the other three sessions, as the S2, S3 and S4 are relatively close to one another. There is, however, substantial search occurring in all sessions. The estimated coefficients are in the 1.25 to 1.30 range, suggesting that on average subjects click 25 – 30% more cells than they include in their final decision.

We make a brief observation about the Round effects, as this effect is statistically significant for the Cell and Search Rations. Based on our inspection of the individual coefficient estimates (not shown in Table 3H-13), the Cell Ratio exhibits an upward trend and Search Ratio exhibits a downward trend over the first several rounds. Both of these trends are consistent with learning – subjects pick more cells over time, but do so with less searching. (The Earnings Ratio exhibits a very slight upward trend, and an insignificant Round Effect.) However, the trends for both the Cell Ratio and the Search Ratio also exhibit a significant downward "spike" in round 6, and then a return to (approximately) the previous level in round 7. This unexpected movement may be the primary reason for the significant Round Effect. One possible explanation is a "training effect." In Part I of the experiment, subjects complete four instructional training sessions, each of which lasts five rounds. Upon returning for Part II, subjects may incorrectly anticipate that round 5 is again the last round. But why subjects systematically choose and/or search substantially fewer cells when confronted with an unexpected round 6 remains a mystery for which we have no parsimonious explanation. We note that this "spike" phenomenon is not observed in the Earnings Ratio.

3H.13.3 Observed Heuristics

We ***Ex Post*** categorize a subject's decisions according to the heuristics discussed in 3H-12. The categorization is a three-step process:

- First, the subject is assigned a round-level categorization in each of her thirty-eight rounds. In a given round, if her final submission choice for a round includes three or less cells then she is assigned to heuristic H, if it includes four or five cells she is assigned to heuristic M, and if it includes six or more cells she is assigned heuristic L.
- Second, the subject is assigned a session-level categorization for each of the four sessions, based on her round-level categorizations. We use four session-level categories: H, M, L or Unable to Categorize. The session-level categorization is subjective, but generally individual subjects exhibited a great deal of consistency in their choices across the rounds of a given session. In only five instances do we designate the subject as "Unable to Categorize" at the session level?[38]
- Third, the subject is assigned an overall categorization. If she was assigned a session-level H categorization in all four session, then she is categorized as H at the overall level; the same rule is used to for categories M and L. If the subject is categorized as using different heuristics at the session level, then she is categorized as Mixed at the overall level. The overall categorization is also

[38] In session S1, subject 68 chose the fixed payoff in all nine rounds and subject 79 chose the fixed payoff in eight rounds. In session S2, subject 47 chose the fixed payoff in seven of the eight rounds. In session S3 subject 50 and in session S4 subject 49 appeared to employ all three heuristics.

subjective, but like the session-level categorization, the exceptions are relatively few. Nonetheless, because of the subjectivity at the session and overall levels, we do not conduct formal statistical tests on our categorization data.

Table 3H-14 shows the results of our categorization at the overall and session level. For simplicity, when summarizing these data, we use terminology like "the subject uses heuristic X" when the more accurate terminology is "the subject's choices are consistent with heuristic X." Two regularities are the frequent use of heuristic L and the infrequent use of heuristic H, both at the session level and the overall level. Approximately half (51%) of the subjects use heuristic L at the overall level, and heuristic L is used over 60% of the time at the session level. By contrast, the corresponding percentages are 5% and about 10% for heuristic H. Interestingly, heuristic M is used about 25% of the time at the session level, but only 16% of the subjects used heuristic M overall. Inspection of the individual level data indicates that 60% (13/22) of the subjects who use heuristic M in some sessions use heuristic H in other sessions.[39]

Table 3H-14: *Ex Post* Categorization of Subject Heuristics

Heuristic	Overall Level	Session Level			
		S1	S2	S3	S4
H: chose 3 cells or less	5% (4/80)	10% (8/80)	10% (8/80)	14% (11/80)	9% (7/80)
M: chose 4 or 5 cells	16% (13/80)	26% (21/80)	28% (22/80)	25% (20/80)	28% (22/80)
L: chose 6 or more cells	51% (41/80)	61% (49/80)	61% (49/80)	60% (48/80)	63% (50/80)
Mixed: used multiple strategies	28% (22/80)	n/a	n/a	n/a	n/a
Unable to Categorize	n/a	3% (2/80)	1% (1/80)	1% (1/80)	1% (1/80)
Column Total	100% (80/80)	100% (80/80)	100% (80/80)	100% (80/80)	100% (80/80)

In Table 3H-14 above, heuristic L is identified as having the highest expected payoff but the highest decision cost, while heuristic H has the lowest expected payoff and the lowest decision cost. Our interpretation of the data in Table 3H-14 is as follows. The majority (51%) of the subjects opt to incur the subjective decision cost in order to obtain the higher payoff; while very few (5%) opt to avoid the subjective decision cost in lieu of lower earnings – but to the extent that thirteen subjects use heuristic M at the overall level, and that thirteen of the subjects who use a Mixed heuristic at the overall level use a combination of H and M, a significant minority (44% = 26/80) undertake a strategy designed to avoid at least some of the decision cost.

[39] Of the remaining nine, six use a combination of L and M, two use L and H, and one uses L, M and H.

3H.14 CONCLUSION

Although wage and salary compensation is an important factor in recruitment and retention of employees, the scope and mix of the fringe benefits is an oft-cited factor. For managers and employers, the effectiveness of the benefit package as a recruitment and retention tool depends on the ability of the workers to make optimal choices from the available options. There is considerable evidence from the naturally occurring world that discrete, multi-attribute goods complicate the decision-making task. The present research uses experimental methods to examine individual decision-making in a stylized discrete multi-attribute goods setting.

Two main results emerge:

- First, the relative trade-off between the attributes of the discrete good is a significant treatment variable. The discrete good has two attributes; one that is fixed across choices and one that varies across choices. When the variable attribute has relatively more weight in the overall reward function, subjects earn a higher reward on average. One hypothesis is that that with greater weight on the variable attribute, subjects focus on the "essence of the problem" and develop heuristics that closely approximate the optimal solution. Overall, subjects do quite well, as most of the eighty subjects consistently earn 90% of the maximum possible reward. Further analysis reveals the majority of experimental subjects adopt heuristics that approximate the optimal solution to the complex linear programming problem.
- Second, subjects rarely choose a fixed payoff option with a known payoff and low decision cost, even when the fixed payoff is 80% of the maximum possible under the decision-making task. This suggests that the subjects place a high implicit valuation on the flexibility in making choices, and that they appear confident in their ability to exceed the fixed payoff.

Collectively, the results suggest that individuals (at least, financially motivated experimental subjects) can indeed handle difficult decision-making tasks like those involving discrete multi-attribute goods – and these individuals systematically reveal a preference for the task, as they rarely opt for a fixed payoff option. This indicates that a flexible benefits package may be strongly preferred to a pre-defined benefits package.

It is interesting that the subjects choose to apply the heuristic that most closely approximates the marginal setting of neo-classical consumer theory. This suggests that individuals prefer settings that allow for marginal decisions, but it may also suggest that individuals prefer, *ceteris paribus*, consumption bundles with larger numbers of goods. This ambiguity suggests that unanswered questions concerning individual behavior remain for further investigation.

3H.14.1 Summary

Compensation and incentives are used as a retention tool. Two general categories of incentives have been discussed, selective re-enlistment bonuses and auctions. While SRB has been shown to be effective in retaining individuals in critical skill categories, the fixed nature of SRB may lead to overpayment or underpayment of military personnel. In contrast, auctions can mitigate the overpayment or underpayment of bonuses to the individuals. Preliminary research shows that individuals tend to bid their reserve price for any given job were the reserve price is the amount of compensation just necessary to induce the individual to continue military employment. However, the effectiveness of auctions, such as relative cost savings and retention effects, is contingent on auction rules. Effectiveness of a real world auction in a military labor market, subject to constraints, is yet to be determined.

Private sector adoption of flexible compensation plans and their effectiveness in retaining employees is evident. Alternatively, while the military has discussed offering a cafeteria-style benefits package institutional constraints of the military are likely to deter adoption of such policies. However, continued shortages in critical skill areas may necessitate a revisit by DOD of offering a flexible compensation plan.