

**PREDICTING EMPLOYEE HEALTH CARE
DECISIONS IN A FLEXIBLE BENEFITS
ENVIRONMENT**

A Preliminary Report

Melissa Barringer
Cornell University

George T. Milkovich
Cornell University

Working Paper #90-23

Center for Advanced Human Resource Studies
New York State School of Industrial & Labor Relations
Cornell University
393 Ives Hall
Ithaca, New York 14951-0952
(607) 255-5427

This paper has not undergone formal review or approval of the faculty of the ILR School. It is intended to make results of Center research, conferences, and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions.

INTRODUCTION

The purpose of this study is to identify the determinants of employees' health care selections in a flexible benefits environment. The goal is to develop a model which will enable managers to predict the health care selections of employees. The research tasks required to accomplish this goal are extensive, and are in progress. The following report will summarize the results of analyses completed to date, the analyses that are in progress, the data required to complete these analyses, and the outcomes that can be expected when the study is done.

Currently, evidence regarding the factors that may influence individuals' decisions about health insurance is sparse. Yet, since the provision of health insurance benefits can comprise a significant portion of labor costs, the ability to predict participation rates among multiple health care options could substantially enhance a firm's financial forecasting and planning capacity. The results from this study will enable managers to estimate the probability that an employee will select a particular option, based on the characteristics of the employee and of the option. When applied to a group of employees, the model developed from this study will estimate how many will select each of the options available.

The fact that a prediction model has not yet been developed by other researchers may be due, at least in part, to the extensive methodological requirements associated with such a task. Decisions about health insurance are complex, and may be influenced by numerous factors. Data on all of these factors is required in order to fully explain health care decisions. Moreover, because the outcome (i.e., health care choice) can take on only a limited number of values, and because the relationships among the factors are complex, the analyses required to accurately assess the effects of predictor variables on the outcome are extensive and complex. We believe that we have developed a viable method for effectively investigating the factors that influence health care decisions. With complete

data, we believe that the results of our analyses will enable us to build a model that will accurately predict these decisions.

NCR DATA

The data provided by NCR contains most of the information required for the analyses. The data set contains information on approximately 28,000 employees. The information is employee-specific, rather than aggregated, and includes the employee's age, gender, tenure, marital status, exempt status, union status, salary, occupation, education, state of residence, NCR division where employed, health plan name, eligibility status, benefits credits, dependent status, all Personal Choice selections, and the cost (to NCR) of each. Variables which are not in the data set, and which may be related to decisions about health insurance, include the number of an employee's dependents (covered and uncovered), health status of the employee and family, and the availability of alternative health insurance coverage (e.g., through spouse's employer). The latter information can only be obtained by surveying employees directly.

PRELIMINARY ANALYSES

Basic Observations: Initial analyses have involved an examination of simple relationships among the variables in the data set. The attached tables summarize the findings from these analyses. The tables suggest that there is a relationship between the health care option that the employee selects and his/her age, tenure, salary, gender and marital status. **As age, tenure, and salary increase, the tendency to select the highest coverage indemnity plan also increases** (refer to Tables 1, 2, and 3). Similarly, the

tendency to select lower levels of indemnity plan coverage decreases as age, tenure and salary increase. However, an examination of tables 4 and 5 reveals that, **within groups of employees of similar age or tenure, the relationship between salary and health care selection does not hold.** For example, within any particular age or tenure group, the percent of employees choosing the high coverage option shows no discernible pattern as salary increases. Within salary groups, on the other hand, the patterns for age and tenure do seem to hold. That is, **holding salary constant, the tendency to choose the high coverage option increases with age and tenure.**

There is also evidence that health care decisions are influenced by the employee's gender and marital status. A higher percentage of men than women have selected the high coverage indemnity plan, while the percentage of women choosing the low coverage indemnity plan or an HMO is greater than that of men (refer to Table 6). These differences are more substantial among married employees (see Table 7). Among single employees, however, the pattern is somewhat different. **Thus, the relationship between gender and health care choice may vary according to the individual's marital status.**

Regression Results. Regression analyses provide information about the independent effect of each of the predictor variables on the health care choice. Results of these analyses can be applied to employee and plan data to estimate the probability that a particular health care option will be selected. Initial regression analyses have been necessarily simple, and have only involved the factors that vary across individuals (i.e., age, tenure, etc.), rather than those that vary across choice (i.e., option price).

The analyses were conducted using either the full sample or one of several subsets of the sample (i.e., employees who chose one of the three indemnity plans, exempt

employees, non-exempt employees, employees who selected self-only coverage, or single employees who selected self-only coverage). Furthermore, due to the nature of the analysis, the outcome variable had to be constrained to fall into one of two categories. In one set of analyses, a model for predicting whether the employee will select the high coverage indemnity plan versus one of the other two indemnity options was estimated. In another set of analyses, the outcomes being predicted are HMO versus all other options. The results from the latter set of analyses suggest that **only two factors are related to the decision to select an HMO: gender and salary. Women seem to have a lower probability of joining an HMO than men, while an increase in salary seems to increase the probability of choosing this option.** The effect of salary was only significant in the sample of non-exempt employees. The effect of gender was significant in both subsamples.

The results of the first set of analyses indicate that **there is a significant relationship between the decision to choose the high coverage indemnity plan and employee gender, tenure, salary, credits, marital status, and age. The strength and direction of the relationships between this decision and gender and salary variables, however, may vary according to the subset of employees being analyzed.** For example, results of the analyses using the full sample, and the sample of all exempt employees, suggests that women have a lower probability of choosing the high coverage indemnity plan than men. Results from the analyses using exempt employees who have selected self only coverage, and single employees who have selected self only coverage, suggest that women have a higher probability of choosing the high plan. The effect of salary is also ambiguous. The effects of the other variables, however, do not seem to vary across analyses. Consistent with the findings (described above) from analyses of simple

relationships among variables, **the probability of choosing the high coverage indemnity plan seems to increase with age and tenure. Not surprisingly, the probability of choosing the high plan also increases with benefit credits. Finally, being married seems to decrease the probability of choosing the high plan.**

LIMITATIONS

These findings suggest that the relationships between health care choices and predictor variables are complex. Clearly, more extensive analyses will be required to disentangle the independent effects of each of these variables on employees' health care decisions. We will be utilizing a procedure called logistic regression. This type of analysis is required when, as here, the outcome that is being predicted can only take on a limited number of values. This procedure does not require the outcome variable to fall into one of two categories, and it allows the inclusion of predictor variables that vary across choice as well as those that vary across individuals. Nevertheless, the computer programming required to estimate a logistic model which includes both choice-specific and individual-specific factors is extensive and highly complex.

A further limitation of our initial findings is that, while interesting, they are not enough to build a model which will help predict employee health care selections with any degree of accuracy. This is because all of these employee-specific variables, taken together, can explain no more than six percent of the observed variation in employee health care selection. Thus, it appears that there are other factors, not yet analyzed, which are related to, and can help predict, employees' health care selections. It is our hope that the inclusion of choice-specific factors (i.e., premium charged to the employee) will improve the explanatory power of the model. We also believe that the inclusion of supplemental

information about employees' perceptions of their health benefits needs will be essential to building a model that will accurately predict health care selection.

ADDITIONAL DATA NEEDS

As discussed above, most of the data needed for this study is in the data set provided by NCR. We do not have information regarding the number of an employee's dependents (covered and uncovered), health status of the employee and family, and the availability of alternative health insurance coverage (e.g., through spouse's employer). These factors are all related to employees' perceptions of their health benefits needs and are therefore crucial to the development of an accurate model for predicting health care choice. As we understand it, this information is not available in NCR information systems and the only way to obtain it is through a survey. We therefore propose conducting a survey of a subset of about 5,000 employees who are currently participating in the Personal Choice program.

PROJECT STATUS

The statistical package we have chosen to run the logistic regression procedures has the broadest computing capacity currently available for this procedure. Because of the extensive and complex programming requirements, we have been consulting with the individual in New York City who wrote the package. We expect to have the program set up to run analyses by January, 1991.

PROJECTED RESULTS

The results we expect to obtain from these analyses will provide information about the strength and direction of relationships between health care choice and the predictor variables, and allow us to estimate the probability that a particular option will be chosen. We might find, for example, that being female increases the probability of choosing the high coverage indemnity plan; or that as the price (to the employee) of an option increases, the probability of choosing it decreases. All of this information can then be used to estimate the probability that an employee with a certain set of characteristics will choose a particular option. Thus, for example, the probability that an HMO option will be selected would be estimated based on the premium (credits) charged to the employee for this option and the characteristics of the employee to whom the option is available.

As stated earlier, the results of this study will be used to develop a model which will enable managers to predict the health care selections of employees. It is our hope that this model will prove to be a valuable planning tool for human resource managers at NCR.

Table 1
Health Care Choice by Age

Health Care Choice	Age						Total
	18-25	26-35	36-44	45-54	55-64	65-73	
High	49.2%	55.8%	57.4%	71.3%	78.8%	86.5%	61.5%
Medium	30.9	18.9	17.7	12.3	6.9	5.6	17.1
Low	5.7	8.2	7.5	5.0	2.9	2.2	6.5
HMO	14.2	17.1	17.5	11.4	11.3	5.6	14.9
	100.0	100.0	100.0	100.0	100.0	100.0	100.0

N = 24,516

Table 2
Health Care Choice by Tenure

Choice	<u>Years with NCR</u>										Total
	0-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	
High	48.7%	60.7%	61.3%	68.3%	71.8%	78.6%	82.8%	86.9%	87.7%	100.0%	61.5%
Medium	26.5	14.7	13.8	12.2	12.4	8.9	5.5	5.0	3.9	0	17.1
Low	8.5	7.4	7.0	5.3	4.0	3.2	2.6	0.6	0	0	6.5
HMO	16.2	17.2	17.9	14.2	11.8	9.3	9.2	7.5	8.4	0	14.9
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

N = 24,516

Table 3
Health Care Choice by Salary

Choice	<u>Salary (in thousands)</u>							Total
	<20	20-30	30-40	40-50	50-75	75-100	>100	
High	50.2%	61.3%	65.1%	64.0%	64.9%	66.5%	83.1%	61.5%
Medium	21.0	18.7	16.8	14.2	12.7	12.2	9.9	17.1
Low	8.9	6.2	6.3	5.9	5.0	3.3	0	6.5
HMO	19.9	13.8	11.9	15.9	17.5	18.0	7.0	14.9
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

N = 24,516

Table 4
Health Care Choice by Salary & Tenure

		<u>YEARS WITH NCR</u>				
		<u><1</u>	<u>1-5</u>	<u>6-10</u>	<u>11-15</u>	<u>>16</u>
<u>SALARY</u>	<u>OPTION</u>					
<20	High	44.2%*	47.0%	54.3%	61.9%	60.2%
	Medium	35.4	22.9	15.7	11.5	11.1
	Low	7.3	9.0	10.9	10.0	7.0
	HMO	13.1	21.2	19.1	16.5	21.6
	Cell N	(398)	(2056)	(681)	(260)	(171)
20-30	High	35.9%	55.6%	61.0%	59.3%	67.8%
	Medium	52.7	24.0	17.0	12.6	10.1
	Low	6.0	7.3	6.4	9.4	4.9
	HMO	5.4	13.1	15.7	18.7	17.2
	Cell N	(184)	(1819)	(1664)	(562)	(1077)
30-40	High	40.8%	46.6%	64.3%	62.4%	78.1%
	Medium	43.7	26.9	14.1	16.2	10.7
	Low	6.3	10.5	6.4	5.9	3.7
	HMO	9.2	16.0	15.3	15.6	7.5
	Cell N	(174)	(2450)	(1251)	(545)	(3850)
40-50	High	37.5%	50.1%	59.7%	62.8%	74.3%
	Medium	42.5	26.3	12.8	13.1	9.9
	Low	2.5	8.1	8.1	6.0	3.5
	HMO	17.5	15.6	19.4	18.1	12.3
	Cell N	(40)	(791)	(898)	(487)	(1559)
50-75	High	35.7%	55.9%	61.9%	61.7%	74.1%
	Medium	35.7	22.9	13.0	14.5	9.3
	Low	21.4	5.7	7.6	5.9	3.0
	HMO	7.1	15.4	17.5	17.8	13.6
	Cell N	(14)	(279)	(709)	(674)	(1410)
75-100	High		47.6%	60.5%	51.9%	77.6%
	Medium		42.9	13.2	13.0	7.2
	Low			5.3	7.4	1.3
	HMO		9.5	21.1	27.8	13.8
	Cell N	(0)	(21)	(38)	(54)	(152)
>100	High		42.9%	85.7%	90.9%	89.8%
	Medium	100.0	57.1		9.1	2.0
	Low					
	HMO			14.3		8.2
	Cell N	(2)	(7)	(7)	(11)	(49)
<u>COLUMN N</u>		812	7423	5248	2593	8268

*Percent of Cell Total

Table 5
Health Care Choice by Salary & Age

SALARY	OPTION	AGE					
		18-25	26-35	36-44	45-54	55-64	65-73
<20	High	49.5%*	48.1%	45.2%	56.0%	68.4%	60.0%
	Medium	28.4	20.7	22.5	13.5	12.9	20.0
	Low	4.2	9.4	12.0	13.1	7.0	13.3
	HMO	18.0	21.8	20.3	17.4	11.7	6.7
	Cell N	(796)	(1435)	(706)	(443)	(171)	(15)
20-30	High	55.8%	58.8%	53.9%	64.9%	75.1%	84.6%
	Medium	29.8	19.4	18.6	13.5	7.8	7.7
	Low	4.0	6.8	8.8	7.3	3.3	
	HMO	10.4	15.0	18.7	14.3	13.8	7.7
	Cell N	(669)	(2164)	(1249)	(770)	(398)	(26)
30-40	High	44.1%	55.7%	62.2%	76.4%	82.9%	100.0%
	Medium	33.6	19.7	17.4	12.0	5.9	
	Low	8.4	8.9	6.9	4.2	2.5	
	HMO	13.9	15.7	13.6	7.4	8.7	
	Cell N	(934)	(2388)	(1498)	(2674)	(759)	(17)
40-50	High	46.0%	56.0%	61.2%	70.2%	79.4%	86.7%
	Medium	32.4	18.8	15.7	11.9	6.5	
	Low	5.4	8.3	6.2	4.4	1.7	
	HMO	16.2	16.9	17.0	13.6	12.4	13.3
	Cell N	(37)	(1235)	(990)	(1144)	(354)	(15)
50-75	High	66.7%	63.7%	60.2%	70.3%	77.9%	100.0%
	Medium	33.3	11.9	16.1	12.3	6.8	
	Low		8.4	5.6	3.1	2.3	
	HMO		16.0	18.1	14.3	12.6	
	Cell N	(3)	(688)	(959)	(1022)	(398)	(16)
75-100	High		57.1%	59.6%	72.1%	78.3%	
	Medium		9.5	18.1	11.5	2.2	
	Low		9.5	3.2	1.0	4.4	
	HMO		23.8	19.2	15.4	15.2	
	Cell N	(0)	(21)	(94)	(104)	(46)	(0)
>100	High		50.0%	75.0%	85.0%	92.9%	
	Medium		50.0	20.0	7.5		
	Low						
	HMO			5.0	7.5	7.1	
	Cell N	(0)	(2)	(20)	(40)	(14)	(0)
COLUMN N		2469	7933	5516	6197	2140	89

*Percent of Cell Total

Table 6
Health Care Choice by Gender

Health Care Choice	Gender		Total
	Male	Female	
Indemnity - High Coverage	65.1%	52.2%	61.5%
Indemnity - Medium Coverage	16.9	17.7	17.1
Indemnity - Low Coverage	4.6	11.5	6.5
HMO	13.4	18.7	14.9
	100.0	100.0	100.0

N = 24,487

Table 7
Health Care Choice
 by
Marital Status, Dependent Category, & Gender

MARITAL/GENDER	OPTION	<u>DEPENDENT CATEGORY</u>		
		Ee Only	Ee +1	Ee +2
Married Female	High	50.9%*	52.8%	43.1%
	Medium	13.0	16.1	18.5
	Low	21.8	17.8	17.8
	HMO	14.3	13.3	20.6
	Cell N	(440)	(1478)	(1818)
Married Male	High	62.7%	72.9%	65.8%
	Medium	14.7	12.8	14.2
	Low	11.3	5.8	3.4
	HMO	11.3	8.6	16.7
	Cell N	(381)	(4497)	(8425)
Single Female	High	62.0%	46.2%	43.6%
	Medium	17.8	16.0	28.1
	Low	2.2	8.1	2.8
	HMO	18.1	29.7	25.5
	Cell N	(2218)	(407)	(353)
Single Male	High	57.6%	55.4%	51.9%
	Medium	26.0	24.5	32.2
	Low	5.1	8.8	2.5
	HMO	11.3	11.4	13.5
	Cell N	(3303)	(511)	(513)
COLUMN N		6342	6893	11,109

*Percent of Cell Total