GOVERNMENT SPENDING

By Roxanne Salinas and Reisa Widjaja

STA 6853 Group 3

Table of Contents

Introduction

Subjects in a GSS were asked their opinions about government spending on the environment, health, assistance to big cities, and law enforcement. Each participant evaluated the government spending on the four categories by answering whether the government is spending not enough, about right, or too much. The data in this study collected opinions from 597 participants. From this data, we performed a homogeneous association model to analyze the association between the opinions on environment spending and health spending (Required from problem 9.5 on page 370 from "Categorical Data Analysis" 3rd ed. by Alan Agresti shown in Figure 1 below). We then looked to see what other relevant associations we can find in the data.

9.5 Subjects in a GSS were asked their opinions about government spending on the environment (E), health (H), assistance to big cities (C), and law enforcement (L). The data are shown at the text website, with outcome categories 1 = too little, 2 = about right, 3 = too much. For the homogeneous association model, Table 9.19 shows some results, including the two-factor estimates for the EH association for coding by which estimates at category 3 of each variable equal 0.

a. Test the model goodness of fit, and interpret.

b. Report the estimated EH conditional odds ratio for the (i) too much and too little categories, (ii) too much and about right categories, and (iii) about right and too little categories.

c. Table 9.20 reports $\{\hat{\lambda}_{eh}^{EH}\}$ when parameters sum to zero within rows and within columns, and when parameters are zero in the first row and first column. Show how these yield the estimated EH conditional odds ratio for the too much and too little categories. Construct a confidence interval for that odds ratio. Interpret.

Figure 1: Textbook Problem 9.5

Parameter Information

In this study, subjects were asked their opinions on government spending. The categories of interest were the environment, health, big city assistance, and law enforcement. For each category, they were asked to rate the amount of spending as either "too little", "about right", or "too much". This information is represented in table 1 below.

Category	Rating
$E = environment$	$1 = too\ little$
$H = health$	$2 = about\ right$
$C = big\ city\ assistance$	$3 = too\ much$
$L = law\ enforcement$	

Table 1: Parameter Information

The homogeneous association model

We are interested in the association between any two categories. This model requires no three-factor interaction term or four-factor interaction term. The form of the model is shown in figure 2.

$$log(\mu_{hijk}) = \lambda + \lambda_h^E + \lambda_i^H + \lambda_j^C + \lambda_k^L + \lambda_{hi}^{EH} + \lambda_{hj}^{EC} + \lambda_{hk}^{EL} + \lambda_{ij}^{HC} + \lambda_{ik}^{HL} + \lambda_{jk}^{CL}$$

Figure 2: Homogeneous Association Model

In this model, the term with double subscript reflects conditional dependence of the variables indicated given the two variables that are not present are held constant. Each category has 3 levels; therefore, we used the 3rd level as our reference for our model. As shown in figure 3 the total amount of non-redundant parameters is 33.

$$\begin{aligned} HIJK &= 1 + (H-1) + (I-1) + (J-1) + (K-1) + (H-1)(I-1) + (H-1)(J-1) \\ &\quad + (H-1)(K-1) + (I-1)(J-1) + (I-1)(K-1) + (J-1)(K-1) \\ &= 1 + (3-1) + (3-1) + (3-1) + (3-1) + (3-1)(3-1) \\ &\quad + (3-1)(3-1) + (3-1)(3-1) + (3-1)(3-1) + (3-1)(3-1) \\ &\quad + (3-1)(3-1) \\ &= 1 + 2 + 2 + 2 + 2 + (2)(2) + (2)(2) + (2)(2) + (2)(2) + (2)(2) + (2)(2) \\ &= 33 \end{aligned}$$

Figure 3: Non-redundant parameters

The full list of parameter estimates can be found in the appendix. We used the maximum likelihood parameter estimates given by SAS.

Testing the model goodness of fit (9.5a)

After fitting our model to the data, we want to ensure that our model accurately represents the data. To test the goodness of fit of the model, we used the hypothesis indicated below.

H_0: *the fitted model fits well* $\quad$ H_A: *the fitted model does not fit well* $\quad$ $\alpha = 0.05$

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	48	31.6695	0.6598
Scaled Deviance	48	31.6695	0.6598
Pearson Chi-Square	48	26.5224	0.5526
Scaled Pearson X2	48	26.5224	0.5526
Log Likelihood		1284.9404	
Full Log Likelihood		-119.2515	
AIC (smaller is better)		304.5030	
AICC (smaller is better)		352.2477	
BIC (smaller is better)		383.5198	

Table 2: Criteria for assessing goodness of fit (SAS)

According to table 2, the Deviance $G^2 = 31.6695$ and the Pearson Chi-square $\mathcal{X}^2 = 26.5224$ both with 48 degrees of freedom. This gives a p-value for the G^2 statistic of 0.97 and a p-value for the $\mathcal{X}^2$ statistic of 0.995. Both p-values are greater than the common significance level of $\alpha = 0.05$. Thus, we fail to reject the null hypothesis and conclude the fitted model fits well.

Estimated EH conditional Odds ratio (9.5b)

In exploring our data, we are interested the association between a subjects rating and on government spending on the environment and health. To do this, we looked at the conditional odds ratio for EH. Under the homogeneous association model, the odds ratio between any pair of variables are the same at every level of the other two variables. Using level 3 as our reference, we are able to obtain estimates for λ_{11}^{EH}, λ_{12}^{EH}, λ_{21}^{EH}, and λ_{22}^{EH}. All other parameters are redundant and therefore are zero. The parameters are shown in table 3 below.

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
E*H	1	1	1	2.1425	0.5566	1.0132	3.2209	14.81	0.0001
E*H	1	2	1	1.4221	0.6034	0.2148	2.6069	5.55	0.0184
E*H	1	3	0	0.0000	0.0000	0.0000	0.0000	.	.
E*H	2	1	1	0.7294	0.5667	-0.4194	1.8265	1.66	0.1980
E*H	2	2	1	0.3183	0.6211	-0.9248	1.5346	0.26	0.6084
E*H	2	3	0	0.0000	0.0000	0.0000	0.0000	.	.
E*H	3	1	0	0.0000	0.0000	0.0000	0.0000	.	.
E*H	3	2	0	0.0000	0.0000	0.0000	0.0000	.	.

Table 3: Estimated parameters for EH association (SAS)

Quick notes about notation

In an IxJ table there are $\frac{IJ(I-1)(J-1)}{4}$ possible odds ratios. This means for our data, as shown in figure 4, there are 9 possible odds ratios, some of which are redundant.

$$\frac{(3)(3)(3-1)(3-1)}{4} = 9$$

Figure 4: Possible odds ratios

When referencing the conditional odds ratios as θ_{ij}, we are referring to using both i and j as row indicators and column indicators. For example, θ_{13} refers to the conditional odds ratio for using the 1st and 3rd level of i and 1st and 3rd level of j. Table 4 shows the highlighted cells that are used to calculate the conditional odds ratio for θ_{13}, θ_{12}, and θ_{23} .

Cells used to calculate θ_{13}			Cells used to calculate θ_{12}			Cells used to calculate θ_{23}		
π_{11}	π_{12}	π_{13}	π_{11}	π_{12}	π_{13}	π_{11}	π_{12}	π_{13}
π_{21}	π_{22}	π_{23}	π_{21}	π_{22}	π_{23}	π_{21}	π_{22}	π_{23}
π_{31}	π_{32}	π_{33}	π_{31}	π_{32}	π_{33}	π_{31}	π_{32}	π_{33}

Table 4: Highlighted 3X3 table cells used to calculated conditional odds ratios

Estimated EH conditional odds ratio for the too much (3) and too little (1) (9.5bi)

First, we want to compare opinions on the government spending too little on environment and health with those who spent too much. The calculation for the conditional log odds ratio is shown in figure 5. Essentially, the only parameter estimate we need is $\hat{\lambda}_{11}^{EH}$ because the other estimates are zero.

$$
\begin{aligned}
\log(\hat{\theta}_{13}) = \log\left(\frac{\hat{\mu}_{11jk}\hat{\mu}_{33jk}}{\hat{\mu}_{13jk}\hat{\mu}_{31jk}}\right) &= \log(\hat{\mu}_{11jk}) + \log(\hat{\mu}_{33jk}) - \log(\hat{\mu}_{13jk}) - \log(\hat{\mu}_{31jk}) \\
&= (\hat{\lambda} + \hat{\lambda}_1^E + \hat{\lambda}_1^H + \hat{\lambda}_j^C + \hat{\lambda}_k^L + \hat{\lambda}_{11}^{EH} + \hat{\lambda}_{1j}^{EC} + \hat{\lambda}_{1k}^{EL} + \hat{\lambda}_{1j}^{HC} + \hat{\lambda}_{1k}^{HL} + \hat{\lambda}_{jk}^{CL}) \\
&\quad + (\hat{\lambda} + \hat{\lambda}_3^E + \hat{\lambda}_3^H + \hat{\lambda}_j^C + \hat{\lambda}_k^L + \hat{\lambda}_{33}^{EH} + \hat{\lambda}_{3j}^{EC} + \hat{\lambda}_{3k}^{EL} + \hat{\lambda}_{3j}^{HC} + \hat{\lambda}_{3k}^{HL} + \hat{\lambda}_{jk}^{CL}) \\
&\quad - (\hat{\lambda} + \hat{\lambda}_1^E + \hat{\lambda}_3^H + \hat{\lambda}_j^C + \hat{\lambda}_k^L + \hat{\lambda}_{13}^{EH} + \hat{\lambda}_{1j}^{EC} + \hat{\lambda}_{1k}^{EL} + \hat{\lambda}_{3j}^{HC} + \hat{\lambda}_{3k}^{HL} + \hat{\lambda}_{jk}^{CL}) \\
&\quad - (\hat{\lambda} + \hat{\lambda}_3^E + \hat{\lambda}_1^H + \hat{\lambda}_j^C + \hat{\lambda}_k^L + \hat{\lambda}_{31}^{EH} + \hat{\lambda}_{3j}^{EC} + \hat{\lambda}_{3k}^{EL} + \hat{\lambda}_{1j}^{HC} + \hat{\lambda}_{1k}^{HL} + \hat{\lambda}_{jk}^{CL}) \\
&= \hat{\lambda}_{11}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{13}^{EH} - \hat{\lambda}_{31}^{EH}
\end{aligned}
$$

Figure 5: Log odds Ratio calculation (1,3)

$$\hat{\theta}_{13} = e^{\hat{\lambda}_{11}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{13}^{EH} - \hat{\lambda}_{31}^{EH}} = e^{2.1425+0-0-0} = e^{2.1425} = 8.5207$$

Figure 6: Estimated Odds Ratio calculation (1,3)

Using the log odds found in figure 5, we can calculate the estimated conditional odds ratio as shown in figure 6. While holding the subject's opinions on government spending for big city assistance and law enforcement constant, participants are 8.5207 times more likely to agree that the government has been spending too little on the environment and health issue than the participants who think the government was spending too much.

Estimated EH conditional odds ratio for the too much (3) and about right (2) (9.5bii)

We can repeat the same process we used in the previous section to calculate the conditional log odds ratio for too much and about right. The calculation is shown in Figure 7 below.

$$\log(\hat{\theta}_{23}) = \log\left(\frac{\hat{\mu}_{22jk}\hat{\mu}_{33jk}}{\hat{\mu}_{23jk}\hat{\mu}_{32jk}}\right) = \log(\hat{\mu}_{22jk}) + \log(\hat{\mu}_{33jk}) - \log(\hat{\mu}_{23jk}) - \log(\hat{\mu}_{32jk})$$
$$= \hat{\lambda}_{22}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{23}^{EH} - \hat{\lambda}_{32}^{EH}$$

Figure 7: Log Odds Ratio calculation (2,3)

$$\hat{\theta}_{23} = e^{\hat{\lambda}_{33}^{EH}+\hat{\lambda}_{22}^{EH}-\hat{\lambda}_{23}^{EH}-\hat{\lambda}_{32}^{EH}} = e^{0+0.3183-0-0} = e^{0.3183} = 1.3748$$

Figure 8: Estimated Odds Ratio calculation (2,3)

Using the log odds found in figure 7, we can calculate the estimated conditional odds ratio as shown in figure 8. While holding the subject's opinions on government spending for big city assistance and law enforcement constant, participants are 1.3748 times more likely to agree that the government has been spending the right amount on the environment and health issue than the participants who think the government was spending too much.

Estimated EH conditional odds ratio for the about right (2) and too little (1) (9.5biii)

To look at the last conditional odds ratio for EH, we can follow the same process as the previous two sections. The calculation for the conditional log odds ratio is shown in figure 9.

$$\log(\hat{\theta}_{12}) = \log\left(\frac{\hat{\mu}_{11jk}\hat{\mu}_{22jk}}{\hat{\mu}_{12jk}\hat{\mu}_{21jk}}\right) = \log(\hat{\mu}_{11jk}) + \log(\hat{\mu}_{22jk}) - \log(\hat{\mu}_{12jk}) - \log(\hat{\mu}_{21jk})$$
$$= \hat{\lambda}_{11}^{EH} + \hat{\lambda}_{22}^{EH} - \hat{\lambda}_{12}^{EH} - \hat{\lambda}_{21}^{EH}$$

Figure 9: Log Odds Ratio calculation (1,2)

$$\hat{\theta}_{12} = e^{\hat{\lambda}_{11}^{EH}+\hat{\lambda}_{22}^{EH}-\hat{\lambda}_{12}^{EH}-\hat{\lambda}_{21}^{EH}} = e^{2.1425+0.3183-1.4221-0.7294} = e^{0.3093} = 1.362$$

Figure 10: Estimated Odds Ratio calculation (1,2)

Using the log odds found in figure 9, we can calculate the estimated conditional odds ratio as shown in figure 10. While holding the subject's opinions on government spending for big city assistance and law enforcement constant, participants are 1.362 times more likely to agree that the government has been spending too little on the environment and health issue than the participants who think the government the right amount.

More on the estimated conditional odds ratios (9.5b)

From the same homogeneous association model, we can find the estimated conditional odds ratio between any pairs of variables, while holding the other variables as constant. For example, table 5 shows part of the estimated parameter results on the association between the government spending on big city assistance (C) and law enforcement.

	Zero for Third Level		
	L		
C	1	2	3
1	1.8741	1.0366	0
2	1.9371	1.823	0
3	0	0	0

Table 5: CL conditional log odds table

The conditional odds ratio for too much (3) and too little (1), about right (2) and too much (3), too little (1) and about right (2) are still able to calculate in a similar way, while now focusing on the estimated conditional odds ratio on the CL association with holding the government spending on the Environment (E) and Health (H) is constant. This is shown in figure 11.

$$\hat{\theta}_{13} = e^{\hat{\lambda}_{11}^{EH}+\hat{\lambda}_{33}^{EH}-\hat{\lambda}_{13}^{EH}-\hat{\lambda}_{31}^{EH}} = e^{1.8741+0-0-0} = e^{1.8741} = 6.515$$

$$\hat{\theta}_{23} = e^{\hat{\lambda}_{22}^{EH}+\hat{\lambda}_{33}^{EH}-\hat{\lambda}_{23}^{EH}-\hat{\lambda}_{32}^{EH}} = e^{1.823+0-0-0} = e^{1.8741} = 6.1904$$

$$\hat{\theta}_{12} = e^{\hat{\lambda}_{11}^{EH}+\hat{\lambda}_{22}^{EH}-\hat{\lambda}_{12}^{EH}-\hat{\lambda}_{21}^{EH}} = e^{1.8741+1.823-1.0366-1.9371} = e^{0.7234} = 2.0801$$

Figure 11: CL conditional odds ratio calculation (1,3) (2,3) & (1,2)

More on the estimated EH odds ratios (9.5c)

The sum to constraints table shown in table 6 below, reports the estimates for λ_{eh}^{EH} when parameters sum to zero within rows and within columns.

	Sum to Zero Constraints		
	H		
E	1	2	3
1	0.509	0.166	−0.676
2	−0.065	−0.099	0.163
3	−0.445	−0.068	0.513

Table 6: Sum to Zero Constraints

These estimates can be used to calculate the conditional odds ratio for EH. The calculation for the conditional odds ratio for too much and too little is shown in figure 12. This is the same conditional odds ratio calculated before.

$$\log(\hat{\theta}_{13}) = \hat{\lambda}_{11}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{13}^{EH} - \hat{\lambda}_{31}^{EH} = 0.509 + 0.513 - (-0.676) - (-0.445) = 2.143$$
$$\hat{\theta}_{13} = e^{2.143} = 8.5207$$

Figure 12: Estimated Odds Ratio calculation using sum to zero constraints (1,3)

We can repeat this process for $\hat{\theta}_{23}$ and $\hat{\theta}_{12}$. These calculations are shown in Figure 13 below. Just as before, the estimated conditional odds ratio calculated with the sum to zero constraints are the same as those calculated with the parameter estimates.

$$\log(\hat{\theta}_{23}) = \hat{\lambda}_{22}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{23}^{EH} - \hat{\lambda}_{32}^{EH} = -0.099 + 0.513 - 0.163 - (-0.068) = 0.319$$
$$\hat{\theta}_{23} = e^{0.319} = 1.3748$$

$$\log(\hat{\theta}_{12}) = \hat{\lambda}_{11}^{EH} + \hat{\lambda}_{22}^{EH} - \hat{\lambda}_{12}^{EH} - \hat{\lambda}_{21}^{EH} = 0.509 + (-0.099) - 0.166 - (-0.065) = 0.309$$
$$\hat{\theta}_{12} = e^{0.309} = 1.362$$

Figure 13: Estimated Odds Ratio calculation using sum to zero constraints (2,3) & (1,2)

Table 7 shows the estimates for λ_{eh}^{EH} when parameters are zero in the first row and first column. This is equivalent to setting too little (level 1) as our reference rather than too much (level 3).

Zero for First Level H		
1	2	3
0	0	0
0	0.309	1.413
0	0.720	2.142

Table 7: Zero for First Level

The diagonal of table 7 gives us the conditional log odds of EH when we hold $e = 1$. This means we can calculate the odds ratio for $\hat{\theta}_{13}$ and $\hat{\theta}_{12}$ as shown in figure 14.

$$\hat{\theta}_{13} = e^{\hat{\lambda}_{33}^{EH}} = e^{2.142} = 8.5207$$
$$\hat{\theta}_{12} = e^{\hat{\lambda}_{22}^{EH}} = e^{0.309} = 1.362$$

Figure 14: Estimated Odds Ratio calculation zero for first level (1,3) & (1,2)

We can also use this table to calculate $\hat{\theta}_{23}$; however, this is not as straightforward. Recall that $\hat{\theta}_{23}$ refers to the estimated conditional odds ratio for the 2nd and 3rd level of environment and health. Referencing a 3x3 table with cell probabilities π_{ij}, $\hat{\theta}_{23}$ can be calculated as shown in figure 15.

$$\hat{\theta}_{23} = \frac{\hat{\pi}_{22}\hat{\pi}_{33}}{\hat{\pi}_{23}\hat{\pi}_{32}}$$

Figure 15: conditional odds ratio in terms of probabilities (2,3)

Since table 7 gives us four estimated local conditional log odds ratios, we can use them to calculate $\hat{\theta}_{23}$. We can rewrite table 7 to create table 8 below.

0	0	0
0	$\log\left(\frac{\hat{\pi}_{22}\hat{\pi}_{33}}{\hat{\pi}_{23}\hat{\pi}_{32}}\right)$	$\log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{23}}{\hat{\pi}_{13}\hat{\pi}_{21}}\right)$
0	$\log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{32}}{\hat{\pi}_{12}\hat{\pi}_{31}}\right)$	$\log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{33}}{\hat{\pi}_{13}\hat{\pi}_{31}}\right)$

Table 8: Zero first level in terms of probabilities

If we use these values in table 8, we can calculate $\hat{\theta}_{23}$ (shown in figure 16). As we can see, the conditional odds ratio is the same value as before. The minor difference in values is due to rounding error.

$$\log\left(\frac{\hat{\pi}_{22}\hat{\pi}_{33}}{\hat{\pi}_{23}\hat{\pi}_{32}}\right) + \log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{33}}{\hat{\pi}_{13}\hat{\pi}_{31}}\right) - \log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{32}}{\hat{\pi}_{12}\hat{\pi}_{31}}\right) - \log\left(\frac{\hat{\pi}_{11}\hat{\pi}_{23}}{\hat{\pi}_{13}\hat{\pi}_{21}}\right) = \log\left(\frac{\hat{\pi}_{22}\hat{\pi}_{33}}{\hat{\pi}_{23}\hat{\pi}_{32}}\right)$$

$$\hat{\theta}_{23} = e^{\hat{\lambda}_{22}^{EH} + \hat{\lambda}_{33}^{EH} - \hat{\lambda}_{23}^{EH} - \hat{\lambda}_{32}^{EH}} = e^{0.309+2.142-1.413-0.720} = 1.374$$

Figure 16: Estimated Odds Ratio calculation zero for first level (2,3)

Confidence interval for EH conditional odds ratio for the too much (3) and too little (1) (9.5c continued)

A 95% confidence interval for θ_{13} is calculated by using the confidence interval for λ_{11}^{EH}. This is calculated in figure 17 below.

$$A\ 95\%\ interval\ for\ \log(\theta_{13}) = \langle 1.0132, 3.2209\rangle$$
$$This\ implies\ that\ A\ 95\%\ interval\ for\ \theta_{13} = \langle e^{1.0132}, e^{3.220*}\rangle = \langle 2.7544, 25.0507\rangle$$

Figure 17: 95% confidence interval for true odds ratio

This means that we are 95% confident the true conditional odds ratio between too much and too little lies between 2.7544 and 25.0507. Since the entire interval is greater than one, this means that the subjects are more likely to say the government is spending too little on the environment and health than say they are spending too much.

Estimated Marginal Odds ratio for EH using the marginal table with fitted values.

The marginal table can show us how a subject's opinion on how much the government spends on the environment relates to how much the government spends on health given their opinions on government spending for big city assistance and law enforcement is held constant. Table 9 shows the fitted counts for the marginal table between environment and health.

		Health		
		Too Little	About Right	Too Much
Environment	Too Little	335	95	14
	About Right	81	33	14
	Too Much	18	10	7

Table 9: Fitted Marginal EH table

Looking at the estimated marginal odds ratio between too little and too much shown in figure 18, we can see that the odds of a participant claiming the government spent too little on health would also claim the government spent too little on the environment is 20.94 times the odds of claiming the government spent too little on health but spent too much on the environment.

$$\hat{\theta}_{EH_{13}} = \frac{335(7)}{14(18)} = \frac{2345}{112} = 20.94$$

Figure 18: EH odds ratio using fitted marginal table (1,3)

This is more than the marginal odds ratio between those who think the government spends the right amount on the environment and health and those who think they spend too much (calculation shown in figure 19). This means that the odds of a participant claiming the government spent right amount on health would also claim the government the right amount on the environment is 1.65 times the odds of claiming the government spent the right amount on health but spent too much on the environment.

$$\hat{\theta}_{EH_{23}} = \frac{33(7)}{14(10)} = \frac{231}{140} = 1.65$$

Figure 19: EH odds ratio using fitted marginal table (2,3)

The marginal odds ratio between too little and about right (shown in figure 20) states that the odds of a participant claiming the government spent too little on health would also claim the government spent too little on the environment is 1.44 times the odds of claiming the government spent the too little on health but spent the right amount on the environment.

$$\hat{\theta}_{EH_{12}} = \frac{335(33)}{95(81)} = \frac{11055}{7695} = 1.44$$

Figure 20: EH odds ratio using fitted marginal table (1,2)

Looking for a better model

We were asked to use the homogenous association model. To make sure that this is a good model to use, we looked at other possible models. When looking for a simpler model, we used backward elimination to remove two factor interactions. We also looked at a no interaction model which would yield an interaction only logistic model giving us little information. We also checked the three-factor interaction model adding in all interaction terms. A summary of the goodness of fit criteria is shown in table 10 below.

Model	Homogenous Association -EL			Homogenous Association -EL -HC			No interaction			Three factor interaction		
Criterion	DF	Value	Value/DF	DF	Value	Value/DF	DF	Value	Value/DF	DF	Value	Value/DF
Deviance	52	34.1026	0.6558	56	39.4112	0.7038	72	124.3443	1.7270	16	8.5237	0.5327
Scaled Deviance	52	34.1026	0.6558	56	39.4112	0.7038	72	124.3443	1.7270	16	8.5237	0.5327
Pearson Chi-Square	52	29.1564	0.5607	56	34.0851	0.6087	72	281.9375	3.9158	16	6.8176	0.4261
Scaled Pearson X2	52	29.1564	0.5607	56	34.0851	0.6087	72	281.9375	3.9158	16	6.8176	0.4261
Log Likelihood		1283.7238			1281.0695			1238.6030			1296.5132	
Full Log Likelihood		-120.4680			-123.1224			-165.5889			-107.6786	
AIC (smaller is better)		298.9361			296.2447			349.1778			345.3573	
AICC (smaller is better)		333.0537			319.8811			351.7130			917.3573	
BIC (smaller is better)		368.3751			356.1059			370.7279			500.9965	

Table 10: Additional models

By comparing the AIC, removing a few interaction terms only made the model marginally better than the homogenous association model. Using no two-factor interaction terms made the model worse as well did adding in the three factor interaction terms. Since the problem required us to look at the EH conditional association, using the homogenous association model is appropriate.

Conclusion

To sum up, for the data that collects 597 opinions about whether the government is spending on 4 categories with 3 different choices. We first fitted the homogeneous association model. We focused on the conditional odds ratio between levels on the association of government spending between Environment and Health. The model set zeros to the 3rd levels, while we also compared the conditional odds ratios when the parameter estimates were calculated when the model sums the parameters to zero constraints or set zeros to the 1st levels. We found out that in either methods of parameter estimations, the conditional odds ratios for the certain level pairs in the association of the government spending between Environment and Health would not change. Moreover, we also calculated the 95% confidence interval for the given conditional odds ratio according to the confidence interval of the estimated parameters. Finally, we made the marginal table between government spending in Environment and Health and calculated the corresponding marginal conditional odds ratio on the opinions of whether the government is spending too little or too much.

Appendix

A: full list of maximum likelihood parameter estimates for the homogenous association model

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept			1	0.5842	0.5757	-0.7268	1.5692	1.03	0.3102
e	1		1	0.1436	0.6555	-1.0854	1.5430	0.05	0.8266
e	2		1	0.3190	0.6704	-0.9488	1.7421	0.23	0.6341
h	1		1	-0.5028	0.6040	-1.6762	0.7163	0.69	0.4052
h	2		1	-1.2827	0.7181	-2.7344	0.1147	3.19	0.0741
c	1		1	-2.0699	0.6968	-3.5401	-0.7834	8.82	0.0030
c	2		1	-3.0381	0.7169	-4.5341	-1.7086	17.96	<.0001
l	1		1	0.2099	0.6474	-1.0064	1.5919	0.11	0.7457
l	2		1	-0.3776	0.7104	-1.7551	1.0944	0.28	0.5950
e*h	1	1	1	2.1425	0.5566	1.0132	3.2209	14.81	0.0001
e*h	1	2	1	1.4221	0.6034	0.2148	2.6069	5.55	0.0184
e*h	2	1	1	0.7294	0.5667	-0.4194	1.8265	1.66	0.1980
e*h	2	2	1	0.3183	0.6211	-0.9248	1.5346	0.26	0.6084
e*c	1	1	1	1.2000	0.5177	0.2599	2.3297	5.37	0.0204
e*c	1	2	1	1.3896	0.4774	0.5117	2.4154	8.47	0.0036
e*c	2	1	1	0.6917	0.5605	-0.3491	1.8868	1.52	0.2171
e*c	2	2	1	1.3767	0.5024	0.4446	2.4436	7.51	0.0061
e*l	1	1	1	-0.1328	0.6378	-1.5189	1.0257	0.04	0.8350
e*l	1	2	1	0.3739	0.6975	-1.0938	1.6961	0.29	0.5919
e*l	2	1	1	-0.2630	0.6796	-1.7036	1.0130	0.15	0.6987
e*l	2	2	1	0.4250	0.7361	-1.0974	1.8433	0.33	0.5637
h*c	1	1	1	-0.1865	0.4547	-1.0639	0.7385	0.17	0.6817
h*c	1	2	1	0.7464	0.4808	-0.1632	1.7487	2.41	0.1206
h*c	2	1	1	-0.4675	0.4978	-1.4348	0.5332	0.88	0.3476
h*c	2	2	1	0.7293	0.5023	-0.2252	1.7688	2.11	0.1466
h*l	1	1	1	1.8741	0.5079	0.8560	2.8658	13.61	0.0002
h*l	1	2	1	1.0366	0.5262	-0.0147	2.0672	3.88	0.0489
h*l	2	1	1	1.9371	0.6226	0.7368	3.2080	9.68	0.0019
h*l	2	2	1	1.8230	0.6355	0.5981	3.1188	8.23	0.0041
c*l	1	1	1	0.8735	0.4604	0.0172	1.8469	3.60	0.0578
c*l	1	2	1	0.5707	0.4863	-0.3433	1.5863	1.38	0.2406
c*l	2	1	1	1.0793	0.4326	0.2682	1.9854	6.23	0.0126
c*l	2	2	1	1.2058	0.4462	0.3661	2.1355	7.30	0.0069
Scale			0	1.0000	0.0000	1.0000	1.0000		

B: Project code

```
data spending;
input E H C L count;
datalines;
1 1 1 1 62
1 1 1 2 17
1 1 1 3 5
1 1 2 1 90
1 1 2 2 42
1 1 2 3 3
1 1 3 1 74
1 1 3 2 31
1 1 3 3 11
1 2 1 1 11
1 2 1 2 7
1 2 1 3 0
1 2 2 1 22
1 2 2 2 18
1 2 2 3 1
1 2 3 1 19
1 2 3 2 14
1 2 3 3 3
1 3 1 1 2
1 3 1 2 3
1 3 1 3 1
1 3 2 1 2
1 3 2 2 0
1 3 2 3 1
1 3 3 1 1
1 3 3 2 3
1 3 3 3 1
2 1 1 1 11
2 1 1 2 3
2 1 1 3 0
2 1 2 1 21
2 1 2 2 13
2 1 2 3 2
2 1 3 1 20
2 1 3 2 8
2 1 3 3 3
2 2 1 1 1
2 2 1 2 4
2 2 1 3 0
2 2 2 1 6
2 2 2 2 9
2 2 2 3 0
2 2 3 1 6
2 2 3 2 5
2 2 3 3 2
2 3 1 1 1
2 3 1 2 0
2 3 1 3 1
2 3 2 1 2
2 3 2 2 1
2 3 2 3 1
```

```
2 3 3 1 4
2 3 3 2 3
2 3 3 3 1
3 1 1 1 3
3 1 1 2 0
3 1 1 3 0
3 1 2 1 2
3 1 2 2 1
3 1 2 3 0
3 1 3 1 9
3 1 3 2 2
3 1 3 3 1
3 2 1 1 1
3 2 1 2 0
3 2 1 3 0
3 2 2 1 2
3 2 2 2 1
3 2 2 3 0
3 2 3 1 4
3 2 3 2 2
3 2 3 3 0
3 3 1 1 1
3 3 1 2 0
3 3 1 3 0
3 3 2 1 0
3 3 2 2 0
3 3 2 3 0
3 3 3 1 1
3 3 3 2 2
3 3 3 3 3
;
run; quit;

proc genmod data=spending;
class E H C L/param=ref;
model count=E H C L E*H E*C E*L H*C H*L C*L / dist=poi lrci type3 residuals obstats;
title 'Homogenours Association Model';
output out=pred p=exp1;
run;quit;

/*Better AIC*/
proc genmod data=spending;
class E H C L/param=ref;
model count=E H C L E*H E*C H*C H*L C*L / dist=poi lrci type3 residuals obstats;
title 'Homogenours Association Model -EL';
output out=pred p=exp1;
run;quit;

/*Even smaller improvement on AIC*/
proc genmod data=spending;
class E H C L/param=ref;
model count=E H C L E*H E*C H*L C*L / dist=poi lrci type3 residuals obstats;
title 'Homogenours Association Model -EL -HC';
output out=pred p=exp1;
run;quit;
```

```
/*Higher AIC*/
proc genmod data=spending;
class E H C L/param=ref;
model count=E H C L / dist=poi lrci type3 residuals obstats;
title 'No interaction';
output out=pred p=exp1;
run;quit;

/*Higher AIC*/
proc genmod data=spending;
class E H C L/param=ref;
model count=E H C L E*H E*C E*L H*C H*L C*L E*H*C E*C*L E*H*L H*C*L/ dist=poi
lrci type3 residuals obstats;
title 'Three factor interaction';
output out=pred p=exp1;
run;quit;
```