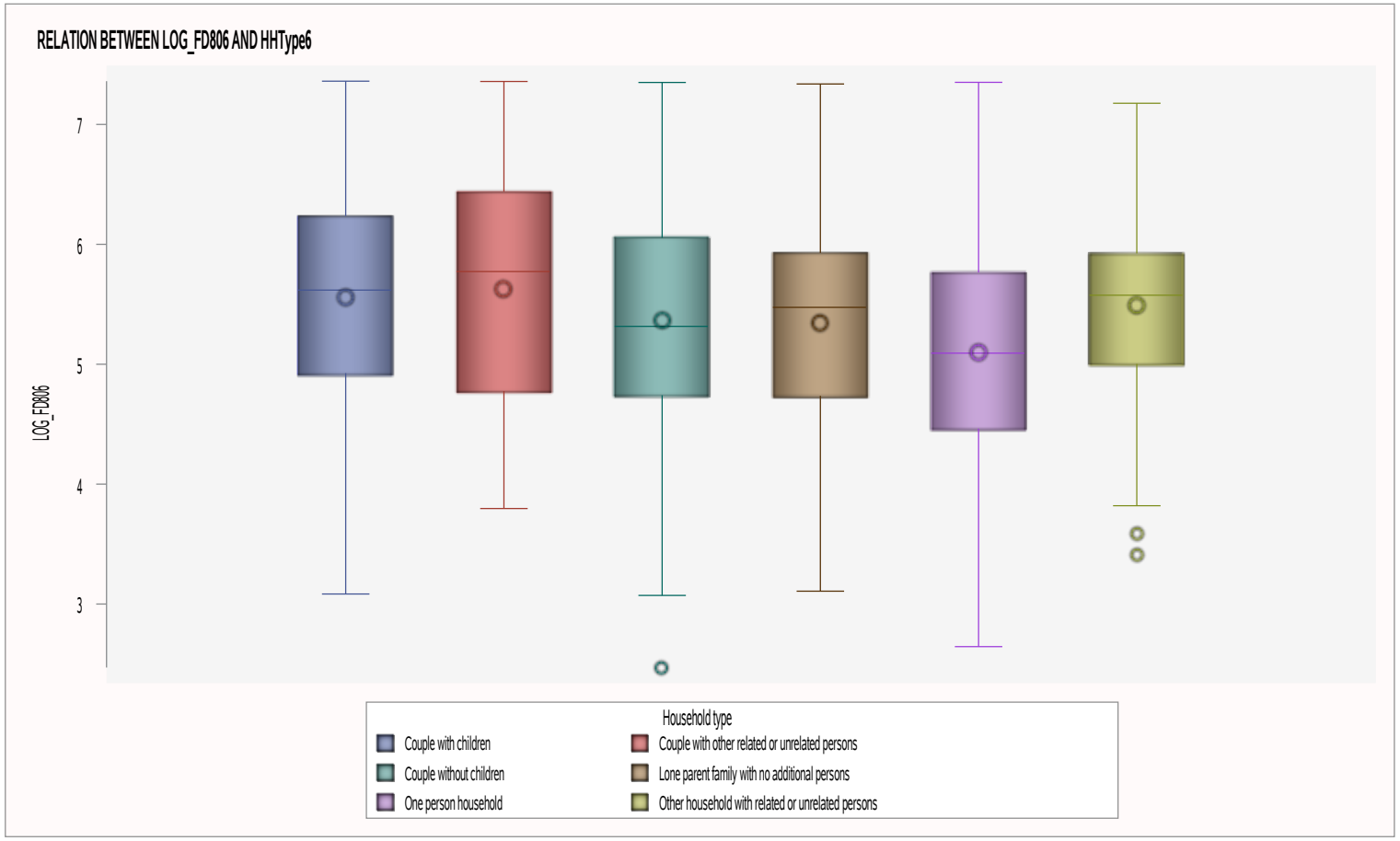
BIVARIATE ANALYSIS OF HHType6 AND LOG_FD806 FOR ANA.MODEL2
RELATION BETWEEN LOG_FD806 AND HHType6

The MEANS Procedure

Analysis Variable : LOG_FD806														
Household type	N Obs	N	N Miss	Minimum	Lower Quartile	Median	Mean	Upper Quartile	Maximum	Quartile Range	Coeff of Variation	Lower 95% CL for Mean	Upper 95% CL for Mean	Skewness
Couple with children	2457237	2457237	0	3.08	4.91	5.62	5.56	6.24	7.36	1.33	16.53	5.56	5.56	-0.29
Couple with other related or unrelated persons	557180	557180	0	3.80	4.76	5.77	5.63	6.44	7.36	1.68	17.07	5.62	5.63	-0.38
Couple without children	2206478	2206478	0	2.47	4.73	5.32	5.37	6.05	7.35	1.33	17.05	5.37	5.37	-0.08
Lone parent family with no additional persons	513225	513225	0	3.11	4.72	5.47	5.34	5.93	7.34	1.21	16.86	5.34	5.34	-0.24
One person household	1964999	1964999	0	2.64	4.45	5.09	5.10	5.77	7.35	1.32	18.78	5.10	5.10	0.08
Other household with related or unrelated persons	429757	429757	0	3.41	4.99	5.58	5.49	5.92	7.18	0.94	14.77	5.49	5.49	-0.35



One-way ANOVA Assumptions
In order to run a one-way ANOVA the following assumptions must be met:
1.The response of interest is continuous and normally distributed for each treatment group:
Normality test: PROC UNIVARIATE NORMAL and QQPlot for each group.
2.Treatment groups are independent of one another. Experimental units only receive one treatment,and they do not overlap.
3.There are no major outliers.
4.A check for unequal variances will help determine which version of a one-way ANOVA is most appropriate
(Levene's test, Null hypothesis: variances are equal between groups):
A .If variances are equal, then the assumptions of a standard one-way ANOVA are met.
B. If variances are unequal, then a Welch's one-way ANOVA is appropriate.

Normal Distribution?
Null hypothesis: sample has a normal distribution
CLT :
a.If it looks normal and each group have more than 30 observations
b.If moderately skewed, each group must have more than 100 observations
*rule of thumb: If skewness is between -1 and -0.5 or between 0.5 and 1, the distribution is moderately skewed.
*if the sample size is over 2000, the Kolmgorov test should be used. If the sample size is less than 2000, the Shapiro test is better.

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with children

Moments			
N	2457237	**Sum Weights**	2457237
Mean	5.55694286	**Sum Observations**	13654725.6
Std Deviation	0.91854558	**Variance**	0.84372598
Skewness	-0.2884932	**Kurtosis**	-0.4780882
Uncorrected SS	77951763.8	**Corrected SS**	2073233.85
Coeff Variation	16.5296927	**Std Error Mean**	0.00058597

Basic Statistical Measures			
Location		**Variability**	
Mean	5.556943	**Std Deviation**	0.91855
Median	5.618479	**Variance**	0.84373
Mode	5.680343	**Range**	4.27700
		Interquartile Range	1.32507

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with children

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	9483.284	**Pr > |t|**	<.0001
Sign	**M**	1228619	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.51E12	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.041799	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	1011.795	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	7695.114	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.36074
99%	7.29261
95%	7.04319
90%	6.66675
75% Q3	6.23566
50% Median	5.61848
25% Q1	4.91059
10%	4.25873
5%	3.94623
1%	3.33506
0% Min	3.08374

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with children

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
3.08374	3250	653	7.33995	1345	483
3.15274	132	22	7.34003	776	396
3.21727	186	161	7.34357	669	680
3.25810	1093	172	7.35361	491	456
3.26805	1144	636	7.36074	43	306

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with other related or unrelated persons

Moments			
N	557180	**Sum Weights**	557180
Mean	5.62580255	**Sum Observations**	3134584.66
Std Deviation	0.96058476	**Variance**	0.92272307
Skewness	-0.3818358	**Kurtosis**	-0.9360414
Uncorrected SS	18148676.3	**Corrected SS**	514121.919
Coeff Variation	17.0746262	**Std Error Mean**	0.00128688

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with other related or unrelated persons

Basic Statistical Measures			
Location		**Variability**	
Mean	5.625803	**Std Deviation**	0.96058
Median	5.773122	**Variance**	0.92272
Mode	6.488141	**Range**	3.56182
		Interquartile Range	1.67842

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4371.662	**Pr > |t|**	<.0001
Sign	**M**	278590	**Pr >= |M|**	<.0001
Signed Rank	**S**	7.761E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.110996	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	1932.028	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	11781.56	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35843
99%	7.28909
95%	6.92992
90%	6.65368

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple with other related or unrelated persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	6.43948
50% Median	5.77312
25% Q1	4.76106
10%	4.07923
5%	3.95124
1%	3.79661
0% Min	3.79661

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
3.79661	12589	725	7.04349	1539	767
3.95124	21675	739	7.08256	60	709
3.95623	36	745	7.16671	1523	759
4.01530	16745	769	7.28909	12702	747
4.07923	20541	765	7.35843	4907	734

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple without children

Moments			
N	2206478	**Sum Weights**	2206478
Mean	5.36827009	**Sum Observations**	11844969.9
Std Deviation	0.91508788	**Variance**	0.83738583
Skewness	-0.0773114	**Kurtosis**	-0.6184178
Uncorrected SS	65434670	**Corrected SS**	1847672.57
Coeff Variation	17.046234	**Std Error Mean**	0.00061605

Basic Statistical Measures			
Location		**Variability**	
Mean	5.368270	**Std Deviation**	0.91509
Median	5.316059	**Variance**	0.83739
Mode	5.248812	**Range**	4.87950
		Interquartile Range	1.32880

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	8714.076	**Pr > |t|**	<.0001
Sign	**M**	1103239	**Pr >= |M|**	<.0001
Signed Rank	**S**	1.217E12	**Pr >= |S|**	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple without children

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.039008	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	612.9778	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	4703.488	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35014
99%	7.17749
95%	6.84050
90%	6.62263
75% Q3	6.05477
50% Median	5.31606
25% Q1	4.72597
10%	4.08665
5%	3.89467
1%	3.38912
0% Min	2.47064

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Couple without children

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.47064	3940	1110	7.26582	1379	1108
3.07177	347	952	7.30361	5912	1186
3.08374	1372	1370	7.32095	5649	1341
3.22764	463	1187	7.32214	2134	1094
3.23789	987	983	7.35014	1360	889

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Lone parent family with no additional persons

Moments			
N	513225	**Sum Weights**	513225
Mean	5.34240093	**Sum Observations**	2741853.72
Std Deviation	0.90069822	**Variance**	0.81125728
Skewness	-0.2365959	**Kurtosis**	-0.5129508
Uncorrected SS	15064438.6	**Corrected SS**	416356.705
Coeff Variation	16.8594276	**Std Error Mean**	0.00125726

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Lone parent family with no additional persons

Basic Statistical Measures			
Location		**Variability**	
Mean	5.342401	**Std Deviation**	0.90070
Median	5.474285	**Variance**	0.81126
Mode	3.990464	**Range**	4.23016
		Interquartile Range	1.20732

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4249.238	**Pr > |t|**	<.0001
Sign	**M**	256612.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	6.585E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.103706	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	942.7026	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	5252.118	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.33743
99%	7.10462
95%	6.74412
90%	6.54276

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Lone parent family with no additional persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	5.92831
50% Median	5.47429
25% Q1	4.72100
10%	3.99046
5%	3.80355
1%	3.39786
0% Min	3.10727

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
3.10727	381	1693	7.13159	990	1666
3.21727	1553	1652	7.16410	1554	1690
3.39786	5372	1658	7.27891	505	1684
3.58129	14896	1655	7.28345	442	1687
3.70549	629	1558	7.33743	836	1713

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=One person household

Moments			
N	1964999	**Sum Weights**	1964999
Mean	5.09706631	**Sum Observations**	10015730.2
Std Deviation	0.95714818	**Variance**	0.91613264
Skewness	0.07979068	**Kurtosis**	-0.3473517
Uncorrected SS	52851039.9	**Corrected SS**	1800198.81
Coeff Variation	18.7784134	**Std Error Mean**	0.00068281

Basic Statistical Measures			
Location		**Variability**	
Mean	5.097066	**Std Deviation**	0.95715
Median	5.092277	**Variance**	0.91613
Mode	5.958166	**Range**	4.70633
		Interquartile Range	1.31732

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	7464.87	**Pr > |t|**	<.0001
Sign	**M**	982499.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	9.653E11	**Pr >= |S|**	<.0001

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=One person household

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.037467	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	481.3685	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	3328.717	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.35109
99%	7.11486
95%	6.81928
90%	6.39533
75% Q3	5.76525
50% Median	5.09228
25% Q1	4.44793
10%	3.92593
5%	3.53777
1%	3.02237
0% Min	2.64476

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=One person household

Extreme Observations					
Lowest			Highest		
Value	Freq	Obs	Value	Freq	Obs
2.64476	15313	2018	7.20510	5467	2127
2.71337	1520	2154	7.21127	1683	2058
3.02237	4062	1917	7.27547	1311	1818
3.03495	8060	1722	7.27783	46	1943
3.05777	7665	1892	7.35109	759	1975

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Other household with related or unrelated persons

Moments			
N	429757	**Sum Weights**	429757
Mean	5.48983025	**Sum Observations**	2359292.98
Std Deviation	0.81070602	**Variance**	0.65724425
Skewness	-0.352329	**Kurtosis**	0.19184912
Uncorrected SS	13234572.6	**Corrected SS**	282454.662
Coeff Variation	14.7674151	**Std Error Mean**	0.00123666

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Other household with related or unrelated persons

Basic Statistical Measures			
Location		**Variability**	
Mean	5.489830	**Std Deviation**	0.81071
Median	5.575570	**Variance**	0.65724
Mode	4.952300	**Range**	3.77018
		Interquartile Range	0.93585

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4439.223	**Pr > |t|**	<.0001
Sign	**M**	214878.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	4.617E10	**Pr >= |S|**	<.0001

Tests for Normality				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.0806	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	391.7784	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2667.195	**Pr > A-Sq**	<0.0050

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	7.17670
99%	7.14419
95%	6.96876
90%	6.56021

The UNIVARIATE Procedure
Variable: LOG_FD806

Freq: WeightD

Household type=Other household with related or unrelated persons

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	5.92383
50% Median	5.57557
25% Q1	4.98798
10%	4.35002
5%	3.97594
1%	3.40652
0% Min	3.40652

Extreme Observations					
Lowest			**Highest**		
Value	**Freq**	**Obs**	**Value**	**Freq**	**Obs**
3.40652	12248	2308	6.98258	2800	2216
3.58269	838	2230	6.98764	644	2241
3.81991	6607	2322	6.99027	10294	2260
3.95124	1422	2313	7.14419	1518	2317
3.97594	8041	2224	7.17670	3082	2325

Null hypothesis: equal variances
a.If variances are equal, then a pooled t-test is appropriate
b.If variances are unequal, then a Satterthwaite (also known as Welch's) test is appropriate

The GLM Procedure

Class Level Information		
Class	**Levels**	**Values**
HHType6	6	Couple with children Couple with other related or unrelated persons Couple without children Lone parent family with no additional persons One person household Other household with related or unrelated persons

Number of Observations Read	2327
Number of Observations Used	2327
Sum of Frequencies Read	8128876
Sum of Frequencies Used	8128876

The GLM Procedure

Dependent Variable: LOG_FD806

Frequency: WeightD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	274072.502	54814.500	64259.8	<.0001
Error	8.13E6	6934038.512	0.853		
Corrected Total	8.13E6	7208111.014			

R-Square	Coeff Var	Root MSE	LOG_FD806 Mean
0.038023	17.16007	0.923587	5.382190

Source	DF	Type I SS	Mean Square	F Value	Pr > F
HHType6	5	274072.5021	54814.5004	64259.8	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
HHType6	5	274072.5021	54814.5004	64259.8	<.0001

The GLM Procedure

Levene's Test for Homogeneity of LOG_FD806 Variance ANOVA of Absolute Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
HHType6	5	10401.9	2080.4	7276.31	<.0001
Error	8.13E6	2324137	0.2859		

Welch's ANOVA for LOG_FD806			
Source	**DF**	**F Value**	**Pr > F**
HHType6	5.0000	61587.9	<.0001
Error	1932349		

The GLM Procedure

Level of HHType6	**N**	**LOG_FD806**	
		Mean	**Std Dev**
Couple with children	2457237	5.55694286	0.91854558
Couple with other related or unrelated persons	557180	5.62580255	0.96058476
Couple without children	2206478	5.36827009	0.91508788
Lone parent family with no additional persons	513225	5.34240093	0.90069822
One person household	1964999	5.09706631	0.95714818
Other household with related or unrelated persons	429757	5.48983025	0.81070602

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

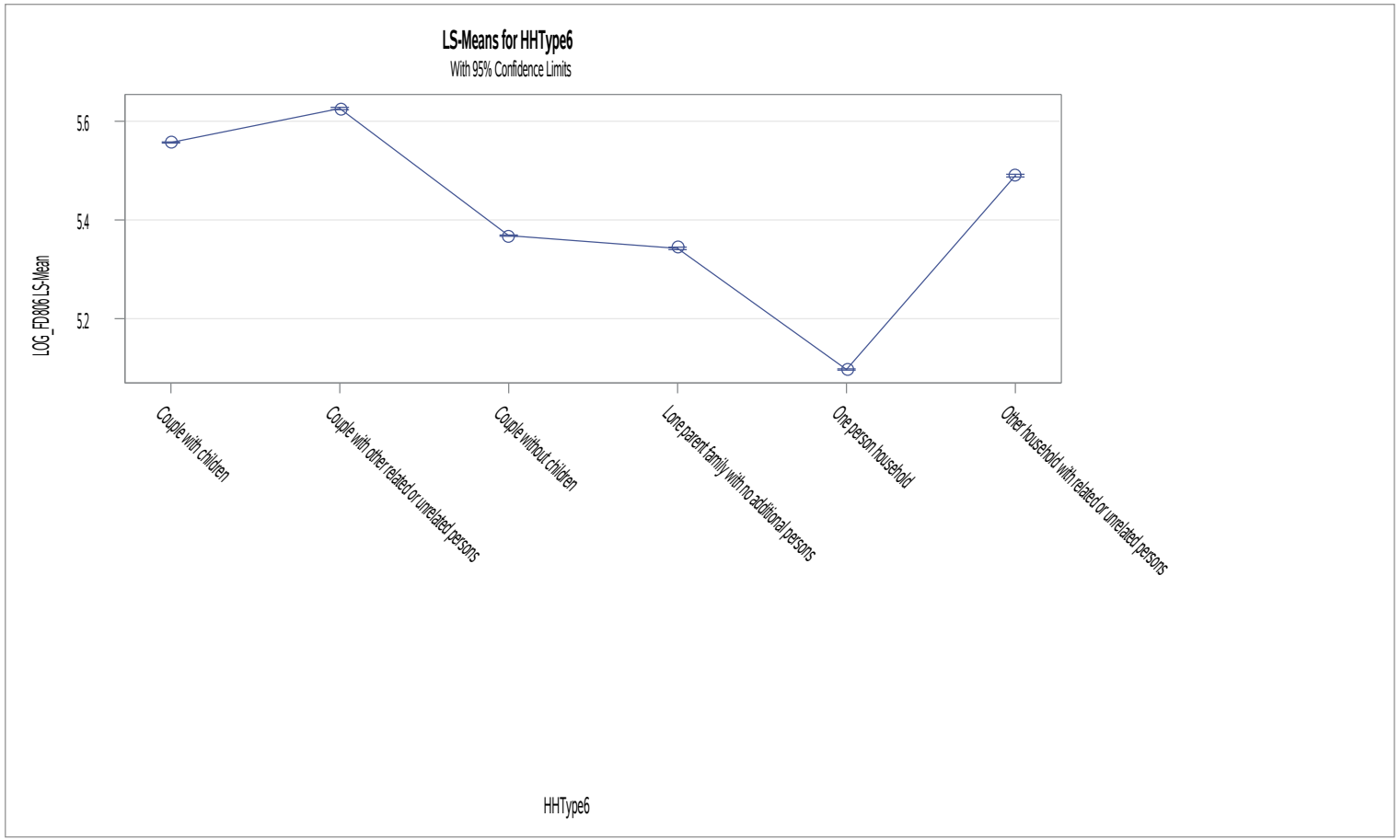
HHType6	LOG_FD806 LSMEAN	LSMEAN Number
Couple with children	5.55694286	1
Couple with other related or unrelated persons	5.62580255	2
Couple without children	5.36827009	3
Lone parent family with no additional persons	5.34240093	4
One person household	5.09706631	5
Other household with related or unrelated persons	5.48983025	6

Least Squares Means for effect HHType6
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: LOG_FD806

i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001	

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer



The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer