UNIVERSITY OF LIMERICK
OLLSCOIL LUIMNIGH
FACULTY OF ARTS, HUMANITIES AND SOCIAL SCIENCES

END OF TERM ASSESSMENT PAPER

MODULE CODE: SO5041 **SEMESTER**: Autumn 2014

MODULE TITLE: Quantitative Research Methods I (MA Sociology) **EXAM DURATION**: Two hours

LECTURER: Dr. Brendan Halpin **% OF TOTAL MARKS**: 40%

EXTERNAL EXAMINER: Prof. Emer Smyth

INSTRUCTIONS TO CANDIDATES:

- Answer question 1 (40%) and two others (30% each).
- Calculators allowed: Yes
- Dictionaries allowed: Yes

---

1 **Data interpretation**: (Compulsory, 40%; answer this question and two others)
See Table 1, drawn from Buchmann and DiPrete (2006). This is a relatively complex table, looking simultaneously at several factors that affect college completion in the US.

(i) Clearly state the outcome variable that the table is summarising.

(ii) Identify the variables that the table is using to classify the outcome.

(iii) In the sample, approximately how many women in the 1938-65 cohort, whose mother and father both had high school or less, completed college?

(iv) Create a 2×2 table of cohort and gender, summarising how many respondents had a mother with "some college or more".

(v) Table 2 contains four panels than summarise aspects of the data in Table 1. Use both tables to write a short report on the determinants of college completion, and how they change over time. Have reference both to the data in the tables, and substantive ideas about the likely mechanisms behind the patterns.

Table 1: Table for data interpretation question (Q2)

Table 1. College Completion Rates by Parents' Education and Family Structure, GSS data

		Father's Education					
		High school or less		Some college or more		Father not present	
Mother's Education		Male	Female	Male	Female	Male	Female
A. White 1938–1965 Birth Cohorts							
High school or less	%	19.5	14.3	41.6	34.6	19.3	13.5
	N	1,349	1,649	330	365	197	281
Some college or more	%	37.7	23.9	60.2	63.1	35.1	30.0
	N	183	239	374	429	77	70
B. White 1966–1977 Birth Cohorts							
High school or less	%	12.4	16.6	38.8	33.6	13.7	13.1
	N	241	271	103	110	73	84
Some college or more	%	26.2	40.8	61.5	63.6	26.1	37.0
	N	65	71	174	195	46	54

Source: Authors' calculations of 1972–2002 General Social Survey data.
Note: Table shows percent of white males and females ages 25–34 that have completed college. GSS = General Social Surveys.

Note: From Buchmann and DiPrete, 2006, "The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement", *American Sociolgical Review* 71, pp515-541

Table 2: Summary tables from Table 1: percent completing college

Mother's educ	Cohort and Gender					
	1938-65			1966-77		
	Male	Female	Total	Male	Female	Total
High school	23.4	17.4	20.1	19.1	20.0	19.6
College	50.7	47.3	48.8	47.7	54.1	51.1
Total	30.3	24.7	27.2	30.8	33.9	32.4

Father's educ	Cohort and Mother's educ					
	1938-65			1966-77		
	High school	College	Total	High school	College	Total
High school	16.6	29.9	18.3	14.6	33.8	18.7
College	37.9	61.7	50.7	36.1	62.6	52.9
Absent	15.9	32.7	19.8	13.4	32.0	20.6
Total	20.1	48.8	27.2	19.6	51.1	32.4

Father's educ	Cohort and Gender					
	1938-65			1966-77		
	Male	Female	Total	Male	Female	Total
High school	21.7	15.5	18.3	15.3	21.6	18.7
College	51.5	50.0	50.7	53.1	52.8	52.9
Absent	23.7	16.8	19.8	18.5	22.5	20.6
Total	30.3	24.7	27.2	30.8	33.9	32.4

Mother's educ	Father's educ and Gender								
	High school			College			Absent		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
High school	18.4	14.6	16.3	40.9	34.4	37.5	17.8	13.4	15.3
College	34.7	27.8	30.8	60.6	63.3	62.0	31.7	33.0	32.4
Total	20.6	16.5	18.3	51.9	50.8	51.3	22.2	18.4	20.1

Table 3: Couples' total work time, mins/day

Wife	Husband	Difference
395	185	-210
406	348	-58
426	340	-86
428	421	-7
378	396	18
397	619	222
402	232	-170
384	478	94
384	300	-84
395	331	-64
391	278	-113
412	325	-87

Table 4: Independent sample t-test on total work time by gender

```
. gen totwk = outwk + domwk

. ttest totwk if outwk>0, by(wgtsex)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    male |     365    350.3425    11.22604    214.4731    328.2664    372.4185
  female |     382    400.4843    10.18982    199.1582     380.449    420.5196
---------+--------------------------------------------------------------------
combined |     747    375.9839    7.616198    208.1606    361.0322    390.9357
---------+--------------------------------------------------------------------
    diff |           -50.14183    15.13549               -79.85511   -20.42855
------------------------------------------------------------------------------
    diff = mean(male) - mean(female)                              t =  -3.3129
Ho: diff = 0                                     degrees of freedom =      745

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0005         Pr(|T| > |t|) = 0.0010          Pr(T > t) = 0.9995
```

2 **T-tests**: (30%)

(i) Twelve couples report their total work time (domestic plus paid) in minutes per day. See Table 3 for their data. Conduct a test of the hypothesis that husbands' and wives' total work-time differs.

(ii) Independent sample t-test:
Refer to the Stata output in Table 4. Note that the variable `outwk` represents daily minutes of paid work and travel outside the home, and `domwk` represents daily minutes of domestic work. The variable `wgtsex` is gender, and the data is drawn from the ESRI 2005 Time Use Study.

a. By reading the two commands, state exactly what the `ttest` command is testing.

b. Interpret the t-test, specifically the hypothesis being tested by `Ha: diff != 0`.

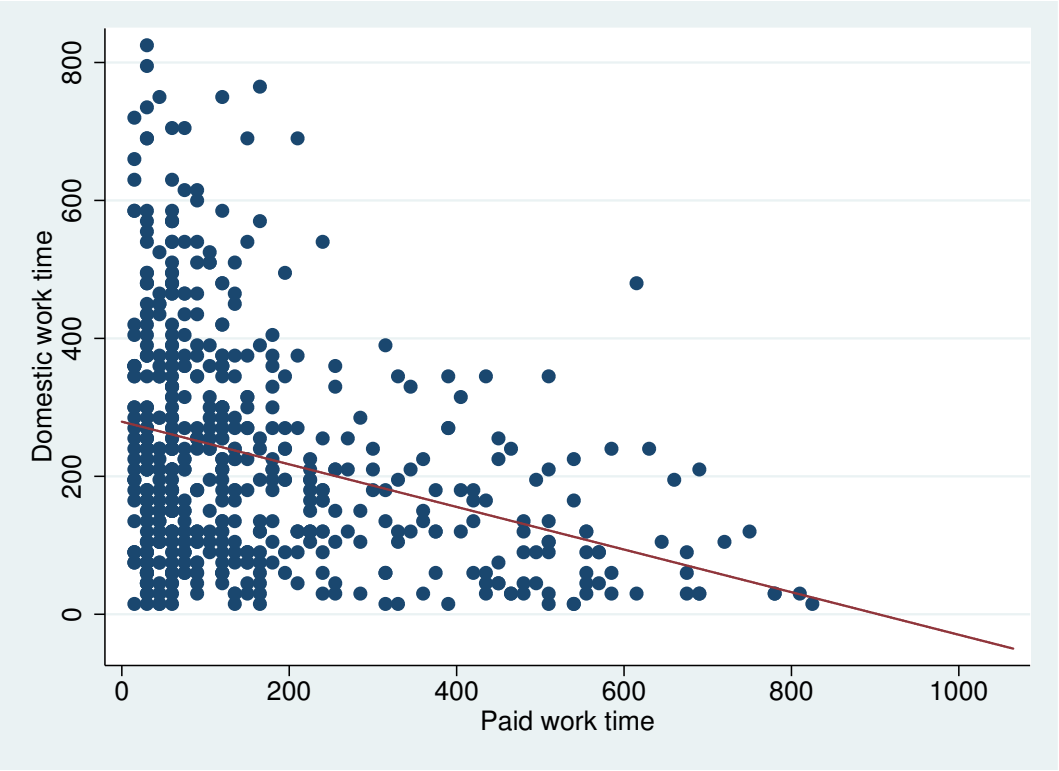

Figure 1: Scatter plot and regression line, domestic work time explained by paid work time

3 **Regression analysis**: (30%)
See Figure 1 and Table 5 (next page): these show the relationship between domestic work time and paid work time (including travel time) as a scatter plot and as a regression analysis.

(i) Describe the relationship evident in the scatterplot.
(ii) Report the regression equation.
(iii) Test the hypothesis that there is a linear relationship between paid and domestic work time.
(iv) Report and interpret the goodness of fit of the regression analysis.
(v) Calculate the predicted value of domestic work time for values of paid work time of 100, 500 and 1000. Interpret the results.
(vi) Do you think a regression analysis properly captures the relationship between these two time uses?

4 **Essay**: (30%)
"Sample survey practice is the key strength, and greatest weakness, of the quantitative method". Discuss.

Table 5: Regression of domestic work time on travel plus paid work time, ESRI 2005 Time Use Study

```
      Source |       SS       df       MS              Number of obs =     573
-------------+------------------------------           F(  1,   571) =   62.83
       Model |  1543522.82     1  1543522.82           Prob > F      =  0.0000
    Residual |  14027900.2   571  24567.2508           R-squared     =  0.0991
-------------+------------------------------           Adj R-squared =  0.0975
       Total |    15571423   572  27222.7675           Root MSE      =  156.74

------------------------------------------------------------------------------
       domwk |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       outwk |  -.3089137   .0389726    -7.93   0.000    -.3854608   -.2323666
       _cons |   279.2214   9.056075    30.83   0.000     261.4342    297.0087
------------------------------------------------------------------------------
```

Formulas and Tables

(a) Standard deviation:

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

(b) z-score: If X is drawn from a normal distribution, with mean μ, and standard deviation σ, its corresponding z-score is:

$$z = \frac{X - \mu}{\sigma}$$

(c) Standard deviation for a proprortion, π:

$$\sigma_\pi = \sqrt{\pi(1-\pi)}$$

(d) Sample standard error, SE, depends on sample standard deviation, s, and sample size, n:

$$SE = \frac{s}{\sqrt{n}}$$

(e) Confidence interval around point estimate, ε, where z is the z-score for the required level of confidence, and SE the standard error (note: z-score may be derived from standard normal distribution or t-distribution, as appropriate):

$$\varepsilon \pm z \times SE$$

(f) Chi-squared statistic for a table,

$$X^2 = \sum \frac{(O-E)^2}{E}$$

where O is the observed value and E the expected value.

(g) Expected value under independence in a table:

$$E = \frac{rc}{T}$$

where r is the row total and c the column total for that cell, and T the grand total for the table.

(h) Predicted value from a bi-variate regression, where a is the constant and b the slope coefficient:

$$\hat{Y} = a + bx$$

(i) Standard error for comparing means of two sub-samples, whose variance may not be the same:

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where s_i is the standard deviation for group i, and n_i the number of cases in group i.

Table of the Standard Normal Distribution

Right tail (probability of $X > z$)

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.00	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.10	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.20	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.30	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.40	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.50	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.60	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.70	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.80	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
.90	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.00	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.10	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.20	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.30	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.40	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.50	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.60	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.70	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.80	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.90	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.00	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.10	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.20	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.30	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.40	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.50	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.60	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.70	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.80	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.90	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.00	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.10	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.20	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.30	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.40	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.50	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.60	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.70	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.80	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.90	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
4.00	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Table of Student's *t* Distribution

Two-tailed probability

Degrees of Freedom	Probability level (Area under both tails)				
	0.10	0.05	0.025	0.01	0.005
1	6.314	12.706	25.452	63.657	127.321
2	2.920	4.303	6.205	9.925	14.089
3	2.353	3.182	4.177	5.841	7.453
4	2.132	2.776	3.495	4.604	5.598
5	2.015	2.571	3.163	4.032	4.773
6	1.943	2.447	2.969	3.707	4.317
7	1.895	2.365	2.841	3.499	4.029
8	1.860	2.306	2.752	3.355	3.833
9	1.833	2.262	2.685	3.250	3.690
10	1.812	2.228	2.634	3.169	3.581
11	1.796	2.201	2.593	3.106	3.497
12	1.782	2.179	2.560	3.055	3.428
13	1.771	2.160	2.533	3.012	3.372
14	1.761	2.145	2.510	2.977	3.326
15	1.753	2.131	2.490	2.947	3.286
16	1.746	2.120	2.473	2.921	3.252
17	1.740	2.110	2.458	2.898	3.222
18	1.734	2.101	2.445	2.878	3.197
19	1.729	2.093	2.433	2.861	3.174
20	1.725	2.086	2.423	2.845	3.153
21	1.721	2.080	2.414	2.831	3.135
22	1.717	2.074	2.405	2.819	3.119
23	1.714	2.069	2.398	2.807	3.104
24	1.711	2.064	2.391	2.797	3.091
25	1.708	2.060	2.385	2.787	3.078
26	1.706	2.056	2.379	2.779	3.067
27	1.703	2.052	2.373	2.771	3.057
28	1.701	2.048	2.368	2.763	3.047
29	1.699	2.045	2.364	2.756	3.038
30	1.697	2.042	2.360	2.750	3.030
35	1.690	2.030	2.342	2.724	2.996
40	1.684	2.021	2.329	2.704	2.971
50	1.676	2.009	2.311	2.678	2.937
60	1.671	2.000	2.299	2.660	2.915
75	1.665	1.992	2.287	2.643	2.892
100	1.660	1.984	2.276	2.626	2.871
500	1.648	1.965	2.248	2.586	2.820
1000	1.646	1.962	2.245	2.581	2.813
100000	1.645	1.960	2.241	2.576	2.807

Table of the χ^2 distribution (chi-sq)

Values of the χ^2 statistic for various degrees of freedom and areas under the right tail

Degrees of Freedom	Area under right tail				
	0.100	0.050	0.025	0.010	0.005
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28.300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.144	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.289	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.652	40.646	44.314	46.928