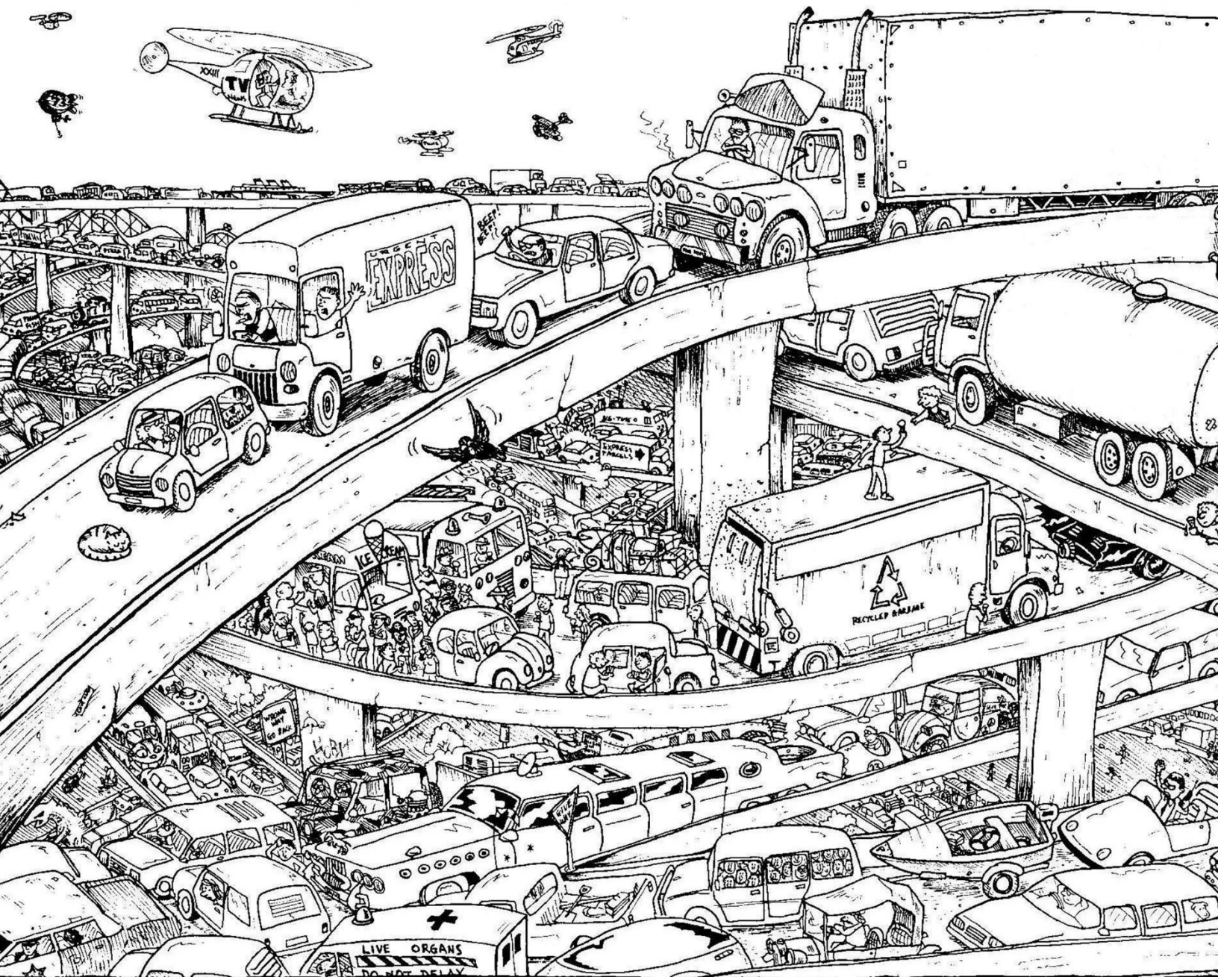**Antonio Jurlina**

Cardiff: Impact of congestion on satisfaction

Cardiff: Impact of congestion on satisfaction

by

Antonio Jurlina

ECO 485

Introductory Econometrics

Schools of Economics

University of Maine

Part I

1. Data

Council authority of Cardiff has, in cooperation with Cardiff Business Partnership, conducted an online survey, yielding 2,094 responses. After data clean-up process, 2,045 observations remained. While most variables were dropped, 17 were deemed significant enough to include (see Figure 1). Missing values were filled with unconditional means, except in the case of the variables describing the primary method of transportation and proximity of a train station, where 19 and 32 observations with missing values were dropped, respectively. Variables regarding age, income and employment duration were acquired via drop-down answer menus on the survey, with each answer constituting a specific bracket. For the sake of continuity and interpretability, these answers were encoded as the midpoint value of each bracket. Since each set of brackets concluded with an open ended one (e.g. more than 10, 65+, etc.), and no middle point could be determined, each of those categories was encoded as the sum of the size of the largest bracket (for the relevant variable) and the bracket's lower bound.

2. Methodology

Literature review, focusing on measuring different sorts of satisfaction (Blanchflower and Oswald, 2002; Fetai et al., 2015; Rayton, 2007), exemplifies various probit models and their appropriate usage. Binary nature of the dependent variables, together with the literature, was the main reason for choosing a probit-model approach. Given that the report attempts to clarify the effect congestion might have on satisfaction with place of residence and working in Cardiff, following two models represent its cornerstone:

$$[job\ satisfaction] = \beta_0 + \beta_1 congestion + x\delta + \varepsilon_1 \quad \text{(Model 1)}$$

$$[satisfaction\ with\ place\ of\ residence] = \beta_0 + \beta_1 congestion + x\delta + \varepsilon_2 \quad \text{(Model 2)}$$

Model 2 dependent variable was binary in the form the data was presented. However, Model 1 dependent variable had to be created from a set of Likert scale variables focusing on job satisfaction. Mean value for the set was created across individuals and compared to "3", which is the neutral option on the answer sheet. Individuals above "3" were categorized as satisfied and the rest were categorized as unsatisfied. This way, a binary variable was generated, appropriate for use in a probit model. Cronbach's alpha was used to determine the scale reliability coefficient (0.8471), indicating that the set of Likert-scale variables used are closely related as a group and representative of the shared concept.

Furthermore, for both Models 1 and 2, β and δ represent coefficients to be estimated (see Figure 1). Control variables were chosen based on two criteria: 1) those causally linked with the dependent variable through congestion and 2) those with a direct, causal link to satisfaction that were deemed as likely sources of severe endogeneity. Explanatory variables of interest consist of demographic characteristics (Richey, 2012; Auh and Cook, 2009), congestion (Lipsetz, 2000; Novaco et al. 1990; Kahneman et al., 2004) and those capturing satisfaction spillover (Dolan & Gosselin, 2000; Ilies et al. 2009). Stochastic terms are noted as ε_1 and ε_2. Also, since the probit models are being used, robust standard errors are reported to correct for underlying (inherent) heteroskedasticity.

3. Hypotheses

H_{01}: $Congestion\ has\ no\ significant\ impact\ on\ job\ satisfaction\ for\ residents\ of\ Cardiff.$
H_{A1}: $Congestion\ has\ a\ negatve\ impact\ on\ job\ satisfaction\ for\ residents\ of\ Cardiff.$

H_{02}: $Congestion\ has\ no\ significant\ impact\ on\ satisfaction\ with\ place\ of\ residence.$
H_{A2}: $Congestion\ has\ a\ negative\ impact\ on\ satisfaction\ with\ place\ of\ residence.$

Models are demonstrating the conditional probability of a specific outcome occurring ($Y_i = 1$ meaning satisfied with job for Model 1 and satisfied with place of residence for Model 2), for

which the marginal effects show how a unit change for *congestion* increases (or decreases) the probability of the given outcome occurring (*et ceteris paribus*).

Part II

1. Satisfaction with working and living in Cardiff

I am quite confident that an increase in perceived level of congestion makes individuals 4% less likely to be satisfied with working in Cardiff.

I am quite confident that an increase in perceived level of congestion makes individuals 5% less likely to be satisfied with living in Cardiff.

Figures 2 and 3 present more detailed results for these claims and state the possibility of a mistake being made. Numbers showing the impact of congestion on satisfaction with working and living in Cardiff are positive. This is a consequence of the way the survey was designed (1 is very bad and 5 is very good) and should be interpreted with an opposite sign. Literature review supports this finding. High levels of traffic congestion are associated with mental and physical stress (Stokols et al. 1978; Novaco & Gonzales, 2009), which are further aggravated with the inability to complete daily routines (Olsson et al. 2013). Furthermore, long work commutes cause residual stress in the workplace (Novaco et al. 1990; Glass & Singer, 1972; Sherrod, 1974). Kahneman et al. (2004) found that work commutes were most frequently associated with negative feelings, out of all daily habits.

Public transportation plays a major role in mitigating the effects of congestion on satisfaction (Kottenhoff & Freij, 2009). However, while it is the most accepted solution to congestion (Schlag & Teubel, 1997), it only works if it is appropriately scaled with the needs of the public (TfL, 2005). In Cardiff, an increase in perceived quality of public transportation makes

individuals 4% (on average) more likely to feel satisfied with their work and 6% (on average) more likely to feel satisfied with where they live, holding everything else the same. Moreover, utilizing public transportation systems decreases the likelihood of being satisfied with place of residence by 6% (on average), holding everything else the same. Those most likely to utilize public transportation live inside the urban area while those living in the suburbs are more likely to drive to work. Therefore, this decrease in likelihood of being satisfied with the place of residence, based on public transport utilization, indicates a preference for living outside of the city. Furthermore, driving decreases the likelihood of being satisfied with working in Cardiff by 8% (on average), holding everything else the same (see Figures 2 and 3 for more detailed results and the likelihood of mistakes being reported). Finally, living close to work increases the likelihood of being satisfied with working in Cardiff by 2% (on average), holding everything else the same.

While there are negative effects of congestion on satisfaction with working and living in Cardiff, those effects are smaller than anticipated. This could be because working during a period of recession or post-recession recovery (ONS, 2014) causes enough satisfaction that lessens the burdens of daily commute (Olsson et al. 2013). Additionally, as noted by Stokols et al. (1978), congestion is only relevant when it significantly differs from the expected traffic levels. If people are exposed to similar congestion levels every day, negative effects are diminished.

2. Policy relevance

Those that live within 2 miles of a train station and those that live further away, are about 20% (on average) less likely to feel satisfied with working in Cardiff, holding everything else the same. Moreover, those that live within 2 miles of a train station are about 20% (on average) more likely to be satisfied with their place of residence, holding everything else the same (see Figures 2 and 3). Since most of survey respondents (74%) live within 2 miles of a train station and more

than half (59%) drive to work (not necessarily overlapping groups), it could be assumed that there is a negative perception of public transportation and a strong tendency to avoid it. This, in addition to reported congestion and public transportation effects, indicates that policy under consideration by the council authority of Cardiff needs to address both public transportation and congestion. Building a new metro system as well as introducing a congestion charge for those driving into the center of Cardiff should be packaged together. Literature review supports this claim, as well (Jaensirisak et al. 2005).

3. Limitations

Variability of answers provided was reduced when most missing values were filled with average values for the category. Further adding to this issue was the design of the survey itself. Answers that were selected from a drop-down menu were presented as brackets, chosen by the surveyor. This eliminated most of the effect that individuals with answers far away from the average would have. Moreover, for privacy reasons, answers stating the specific place of residence were excluded, introducing additional lack of clarity.

Due to model choice, differences between observed and predicted and expected and predicted values for variables of interest were not uniform across the data. This means that I was more likely to falsely perceive an observation as insignificant and reject its validity and explanatory power. Finally, I wish to point out the possibility that any variable not considered stands a chance of being correlated with variables I did consider (e.g. place of residence might be linked with public transport utilization, commute satisfaction and satisfaction with living and working in Cardiff). By omitting any such variable, I introduced the likelihood of overstating the effects considered variables have on overall satisfaction.

Word Count: 1488

Works Cited

1. *Auh, S., & Cook, C. (2009). Quality of Community Life Among Rural Residents: An Integrated Model. Social Indicators Research, 94(3), 377-389.*

2. *Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. Journal of Public Economics, 88, 1359-1386. doi:10.1016/S0047-2727(02)00168-8*

3. *Fetai, B., Abduli, S., & Qirici, S. (2015). An Ordered Probit Model of Job Satisfaction in the Former Yugoslav Republic of Macedonia. Procedia Economics and Finance, 33, 350-357. doi:10.1016/S2212-5671(15)01719-0*

4. *Glass, D. C., Singer, J., & Pennebaker, J. Behavioral and physiological effects of uncontrollable environmental events. In D. Stokols (Ed.), Perspectives on environment and behavior. New York: Plenum, 1977.*

5. *Gross Domestic Product Preliminary Estimate: Q4 2014 (pp. 1-20, Rep.). (2015). Office for National Statistics.*

6. *Ilies, R., Wilson, K. S., & Wagner, D. T. (2009). The Spillover of Daily Job Satisfaction Onto Employees' Family Lives: The Facilitating Role of Work-Family Integration. Academy of Management Journal, 52(1), 87-102. doi:10.5465/AMJ.2009.36461938*

7. *Jaensirisak, S., Wardman, M., May, A.D., 2005. Explaining variations in public acceptability of road pricing schemes. Journal of Transport Economics and Policy 39, 127–154.*

8. *Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004). A survey method for characterizing daily life experience: The day reconstruction method (DRM). Science, 306, 1776–1780.*

9. *Kottenhoff, K., & Freij, K. B. (2009). The role of public transport for feasibility and acceptability of congestion charging – The case of Stockholm. Transportation Research Part A, 43, 297-305. doi:10.1016/j.tra.2008.09.004*

10. *Lipsetz, D. A. (2000). Residential Satisfaction: Identifying the differences between suburban-ites and urbanites. Columbus: Ohio State University.*

11. *Novaco, R. W., & Gonzales, O. I. (2009). Commuting and well-being. In Y. Amichai-Hamburger (Ed.), Technology and well-being (pp. 174–205). New York: Cambridge University Press.*

12. *Novaco, R. W., Stokols, D., & Milanesi, L. (1990). Objective and subjective dimensions of travel impedance as determinants of commuting stress. American Journal of Community Psychology, 18, 231–257.*

13. *Olsson, L. E., Garling, T., Ettema, D., Friman, M., & Fujii, S. (2013). Happiness and Satisfaction with Work Commute. Soc Indic Res, 111, 255-263. doi:10.1007/s11205-012-0003-2*

14. *Rayton, B. A. (2006). Examining the interconnection of job satisfaction and organizational commitment: An application of the bivariate probit model. Int. J. of Human Resource Management, 17(1), 139-154. doi:10.1080/09585190500366649*

15. *Richey, S. (2012). Determinants of Community Satisfaction and its Relative Importance for Life Satisfaction.*

16. *Schlag, B., Teubel, U., 1997. Public acceptability of transport pricing. IATSS Research, 21(2).*

17. *Sherrod, D. R. Crowding, perceived control, and behavioral aftereffects. Journal of Applied Social Psychology, 1974, 4, 171-186.*

18. *Shimon L. Dolan & Eric Gosselin, 2000. "Job satisfaction and life satisfaction: Analysis of a reciprocal model with social demographic moderators," Economics Working Papers 484, Department of Economics and Business, Universitat Pompeu Fabra.*

19. *Stokols, D., Novaco, R. W., Stokols, J., & Campbell, J. (1978). Traffic Congestion, Type A Behavior, and Stress. Journal of Applied Psychology, 63(4), 467-480. doi:10.1037/0021-9010.63.4.467*

20. *Transport for London (TfL), 2005. Central London Congestion Charging. Impacts monitoring, Third Annual Report, April 2005*

Figure 1 – General summary

	Average	Standard Deviation	Minimum	Maximum	Explanation
Quality of public transportation	3.26	1.09	1	5	1 is very bad and 5 is very good
Being able to get from place to place with little traffic (i.e. congestion)	2.72	1.04	1	5	1 is very bad and 5 is very good
Work proximity	2.98	1.28	1	5	1 you like the most and 5 you dislike
Travel within the city	3.86	0.96	1	5	1 is not important and 5 is very important
Ease of getting to work	3.65	1.02	1	5	1 is not important and 5 is very important
Overall satisfaction with life	2.95	0.98	1	5	1 being highly satisfied and 5 being unsatisfied
# of years employed	9.02	5.46	0.5	15	How long have you worked in your present employment
Income	25,748.18	10,318.80	5,720	55,241	How much do you get paid, before taxes
Age	41.50	11.05	21	74	How old are you
Sex	0.42	0.49	0	1	1 for male, 0 for female
Relationship	0.69	0.46	0	1	1 for in a relationship, 0 for single
Children	0.56	0.49	0	1	1 for having any number of children, 0 for none
Overall satisfaction with place of residence	0.66	0.47	0	1	1 for satisfied, 0 for unsatisfied
Train station proximity	2.72	0.50	1	3	Do you have a train station within 2 miles of your residence? 1 - Don't Know, 2 - No, 3 - Yes
Education	2.53	1.10	1	6	1 - High School, 2 - Associates degree, 3 – Bachelor's Degree BA, BSc, 4 – Master's Degree, 5 - Professional degree, 6 - PhD
Primary mode of transportation	1.36	0.83	0	2	0 - Hippie, 1 - Public, 2 - Drive
Job satisfaction	0.52	0.50	0	1	0 - Unsatisfied, 1 - Satisfied

Figure 2 – Model 1 marginal effects

		Marginal Effect	Standard error
Quality of public transportation		**0.04941 ***	[0.010541]
Being able to get from place to place with little traffic (i.e. congestion)		**0.0411897 ***	[0.0108126]
Work proximity		**0.0229768 ***	[0.0082522]
Travel within the city		0.00686	[0.0137853]
Ease of getting to work		0.01526	[0.0129639]
Overall satisfaction with life		**-0.02716 **	[0.0107514]
# of years employed		-0.00119	[0.0022301]
Income		**0.00001 ***	[1.10e-06]
Age		-0.00768	[0.0065343]
Sex		**-0.07248 ***	[0.021709]
Relationship		-0.01465	[0.0245402]
Children		0.03934	[0.0256719]
Overall satisfaction with place of residence		**0.09391 ***	[0.0226908]
Train station proximity	**Yes**	**-0.20385 ***	[0.0584247]
	No	**-0.21915 ***	[0.0609073]
Education		**0.02444 **	[0.009904]
Primary mode of transportation	**Public**	0.01019	[0.0332334]
	Drive	**-0.07591 ***	[0.0270008]

* - 90 % significance / ** - 95 % significance /
*** - 99 % significance

Figure 3 – Model 2 marginal effects

		Marginal Effect	Standard error
Quality of public transportation		**0.0616198 *****	[0.0097441]
Being able to get from place to place with little traffic (i.e. congestion)		**0.0522773 *****	[0.0100888]
Work proximity		0.00252	[0.0078019]
Travel within the city		0.01641	[0.0130512]
Ease of getting to work		**-0.02084 ***	[0.0124397]
Overall satisfaction with workplace		**0.0861031 *****	[0.0202487]
# of years employed		0.00159	[0.0021118]
Income		**0.00000286 *****	[0.00000109]
Age		-0.00624	[0.0061355]
Sex		**-0.03487**	[0.0204136]
Relationship		**0.1058187 *****	[0.0227303]
Children		0.03392	[0.0240695]
Overall satisfaction with life		0.00244	[0.0099781]
Train station proximity	**Yes**	**0.2123409 *****	[0.0749149]
	No	0.12099	[0.0769091]
Education		**0.0193598 ****	[0.0093465]
Primary mode of transportation	**Public**	**-0.05518 ***	[0.0318249]
	Drive	-0.03030	[0.0254]

* - 90 % significance / ** - 95 % significance /
*** - 99 % significance

Appendix A – Stata log file

```
--------------------------------------------------------------------------------------------
----------------------
      name:  <unnamed>
       log:  D:\UMaine\Spring (2018)\ECO 485\Stata\Empirical_Project_3.log
  log type:  text
 opened on:   6 May 2018, 17:53:05

.
. use CBP_survey.dta //choose data

.
. // following commands represent my data clean-up process
. ////////////////////////////////////////////////////////////////
>
. encode Q3, generate(yrs_employed)

. replace yrs_employed = 0.5 if Q3 == "Less than 1 year"
variable yrs_employed was long now double
(171 real changes made)

. replace yrs_employed = 2 if Q3 == "1 to 3 years"
(221 real changes made)

. replace yrs_employed = 4 if Q3 == "3 to 5 years"
(261 real changes made)

. replace yrs_employed = 7.5 if Q3 == "5 to 10 years"
(554 real changes made)

. replace yrs_employed = 15 if Q3 == "More than 10 years"
(873 real changes made)

.
. encode Q4, generate(income)

. replace income = 5720 if Q4 == "£0 - £11,440 per year"
(85 real changes made)

. replace income = 12480.5 if Q4 == "£11,441 - £13,520 per year"
variable income was long now double
(54 real changes made)

. replace income = 14820.5 if Q4 == "£13,521 - £16,120 per year"
(162 real changes made)

. replace income = 17420.5 if Q4 == "£16,121 - £18,720 per year"
(194 real changes made)

. replace income = 20540.5 if Q4 == "£18,721 - £22,360 per year"
(372 real changes made)

. replace income = 25220.5 if Q4 == "£22,361 - £28,080 per year"
(462 real changes made)

. replace income = 31720.5 if Q4 == "£28,081 - £35,360 per year"
(420 real changes made)

. replace income = 40300.5 if Q4 == "£35,361 - £45,240 per year"
(191 real changes made)

. replace income = 55241 if Q4 == "£45,241 or more per year"
(94 real changes made)

.
. drop if Q5 == "25- 40" | Q5 == "25- 41" | Q5 == "25- 42" | Q5 == "25- 43"
(4 observations deleted)
```

```
. drop if Q5 == "25- 44" | Q5 == "25- 45" | Q5 == "25- 46" | Q5 == "25- 47" | Q5 == "25- 48" | Q5
== "25- 49"
(6 observations deleted)

. encode Q5, generate(age)

. replace age = 21 if Q5 == "18- 24"
(89 real changes made)

. replace age = 32 if Q5 == "25- 39"
(865 real changes made)

. replace age = 47 if Q5 == "40- 54"
(827 real changes made)

. replace age = 59.5 if Q5 == "55- 64"
variable age was long now double
(252 real changes made)

. replace age = 74 if Q5 == "65"
(31 real changes made)

.
. generate age_sq = age * age
(20 missing values generated)

.
. label define sex 1 "Male" 0 "Female"

. encode Q6, generate(sex)

.
. generate relationshipy = Q7
(38 missing values generated)

. replace relationshipy = "Yes" if Q7 == "Co-habiting"
(411 real changes made)

. replace relationshipy = "Yes" if Q7 == "Married"
(1,008 real changes made)

. replace relationshipy = "No" if Q7 == "Single"
(627 real changes made)

. label define relationship 1 "Yes" 0 "No"

. encode relationshipy, generate(relationship)

. drop relationshipy Q7

.
. label define children 1 "Yes" 0 "No"

. encode Q8, generate(children)

.
. rename Q14h pub_trans_quality

. rename Q14i congestion

.
. generate satisfactiony = Q15
(26 missing values generated)

. label define satisfaction 1 "Satisfied" 0 "Unsatisfied"

. encode satisfactiony, generate(satisfaction)

. drop satisfactiony Q15

.
```

```
. encode Q19, generate(train)

. label define education 1 "High School" 2 "Associates degree" 3 "Bachelors Degree BA, BSc" 4
"Masters Degree" 5 "Profess
> ional degree" 6 "PhD"

. encode Q20, generate(education)

.
. rename Q30c travel_importance

. rename Q30f work_travel_ease

. rename Q31 life_satisfaction

.
. summarize pub_trans_quality

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
pub_trans_~y |      2,038    3.257605    1.092609         1          5

. replace pub_trans_quality = r(mean) if pub_trans_quality == .
variable pub_trans_quality was byte now float
(46 real changes made)

. summarize congestion

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  congestion |      2,041    2.721705    1.048325         1          5

. replace congestion = r(mean) if congestion == .
variable congestion was byte now float
(43 real changes made)

. summarize yrs_employed

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
yrs_employed |      2,070    9.020048    5.475126        .5         15

. replace yrs_employed = r(mean) if yrs_employed == .
(14 real changes made)

. summarize income

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      income |      2,024    25713.06    10468.28      5720      55241

. replace income = r(mean) if income == .
(60 real changes made)

.
. summarize age

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         age |      2,064    41.52422    11.11513        21         74

. replace age = r(mean) if age == .
(20 real changes made)

. summarize sex

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         sex |      2,050     .422439    .4940682         0          1

. replace sex = r(mean) if sex == .
```

```
variable sex was long now double
(34 real changes made)

. summarize relationship

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
relationship |      2,046    .6935484    .4611323         0          1

. replace relationship = r(mean) if relationship == .
variable relationship was long now double
(38 real changes made)

. summarize satisfaction

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
satisfaction |      2,058    .6588921    .4741967         0          1

. replace satisfaction = r(mean) if satisfaction == .
variable satisfaction was long now double
(26 real changes made)

. summarize train

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       train |      2,052    2.718324    .4972822         1          3

. drop if train == .
(32 observations deleted)

. summarize education

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   education |      1,913    2.534239    1.141186         1          6

. replace education = r(mean) if education == .
variable education was long now double
(139 real changes made)

. summarize life_satisfaction

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
life_satis~n |      1,877    2.951518    1.020208         1          5

. replace life_satisfaction = r(mean) if life_satisfaction == .
variable life_satisfaction was byte now float
(175 real changes made)

. summarize children

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    children |      2,024    .5607708    .4964158         0          1

. replace children = r(mean) if children == .
variable children was long now double
(28 real changes made)

. summarize work_travel_ease

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
work_trave~e |      1,860    3.648387    1.072429         1          5

. replace work_travel_ease = r(mean) if work_travel_ease == .
variable work_travel_ease was byte now float
(192 real changes made)
```

```
.  summarize travel_importance

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
travel_imp~e |     1,867    3.859668    1.005673         1          5

.  replace travel_importance = r(mean) if travel_importance == .
variable travel_importance was byte now float
(185 real changes made)


.
.  generate transport1 = Q16

.  replace transport1 = "Hippie" if Q16 ==  "Walk" | Q16 == "Cycle"
(466 real changes made)

.  replace transport1 = "Public" if Q16 == "Train" | Q16 == "Bus"
(383 real changes made)

.  replace transport1 = "Drive" if Q16 == "Car /Motorcycle"
(1,203 real changes made)

.  label define trans1 0 "Hippie" 1 "Public" 2 "Drive"

.  encode transport1, generate(trans1)

.  //generate transport2 = Q18
.  //replace transport2 = "Hippie" if Q18 ==  "Walk" | Q18 == "Cycle"
.  //replace transport2 = "Public" if Q18 == "Train" | Q18 == "Bus"
.  //replace transport2 = "Drive" if Q18 == "Car /Motorcycle"
.  //encode transport2, generate(trans2)
.
.  drop if trans1 == .
(0 observations deleted)


.
.
.  //generate transporty = Q16
.  //replace transporty = "Drive" if Q16 == "Car /Motorcycle"
.  //generate transports = Q18
.  //replace transports = "Drive" if Q18 == "Car /Motorcycle"
.  //generate transport = transporty + transports
.  //replace transport = "Bus" if transport == "BusBus"
.  //replace transport = "Drive" if transport == "DriveDrive"
.  //replace transport = "Train" if transport == "TrainTrain"
.  //replace transport = "Walk" if transport == "WalkWalk"
.  //replace transport = "Bus / Cycle" if transport == "BusCycle"
.  //replace transport = "Bus / Cycle" if transport == "CycleBus"
.  //replace transport = "Bus / Drive" if transport == "BusDrive" & transport == "DriveBus"
.  //replace transport = "Bus / Drive" if transport == "BusDrive" | transport == "DriveBus"
.  //replace transport = "Bus / Train" if transport == "BusTrain" | transport == "TrainBus"
.  //replace transport = "Bus / Walk" if transport == "BusWalk" | transport == "WalkBus"
.  //replace transport = "Cycle / Drive" if transport == "CycleDrive" | transport == "DriveCycle"
.  //replace transport = "Cycle / Train" if transport == "CycleTrain" | transport == "TrainCycle"
.  //replace transport = "Cycle / Walk" if transport == "CycleWalk" | transport == "WalkCycle"
.  //replace transport = "Cycle / Train" if transport == "Cycletrain"
.  //replace transport = "Drive / Train" if transport == "DriveTrain" | transport == "TrainDrive"
.  //replace transport = "Drive / Walk" if transport == "DriveWalk" | transport == "WalkDrive"
.  //replace transport = "Train / Walk" if transport == "TrainWalk" | transport == "WalkTrain"
.
.  //encode transport, generate(transportation)
.  //drop if transportation == .
.  //drop transport transporty transports Q16 Q18
.
.  summarize Q28a

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        Q28a |     1,879    3.185205    1.118998         1          5
```

```
. replace Q28a = r(mean) if Q28a == .
variable Q28a was byte now float
(173 real changes made)

. summarize Q28b

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28b |      1,879     3.25173    1.067461         1          5

. replace Q28b = r(mean) if Q28b == .
variable Q28b was byte now float
(173 real changes made)

. summarize Q28c

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28c |      1,877    2.936601    1.102782         1          5

. replace Q28c = r(mean) if Q28c == .
variable Q28c was byte now float
(175 real changes made)

. summarize Q28d

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28d |      1,875      2.8368     1.15271         1          5

. replace Q28d = r(mean) if Q28d == .
variable Q28d was byte now float
(177 real changes made)

. summarize Q28e

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28e |      1,876    2.684435    1.119806         1          5

. replace Q28e = r(mean) if Q28e == .
variable Q28e was byte now float
(176 real changes made)

. summarize Q28f

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28f |      1,877    2.750133    1.153888         1          5

. replace Q28f = r(mean) if Q28f == .
variable Q28f was byte now float
(175 real changes made)

. summarize Q28g

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q28g |      1,887     3.45204    1.105134         1          5

. replace Q28g = r(mean) if Q28g == .
variable Q28g was byte now float
(165 real changes made)

. summarize Q29a

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29a |      1,878    3.056443    1.367874         1          5

. replace Q29a = r(mean) if Q29a == .
```

```
variable Q29a was byte now float
(174 real changes made)

. summarize Q29b

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29b |      1,877     2.94033    1.182521         1          5

. replace Q29b = r(mean) if Q29b == .
variable Q29b was byte now float
(175 real changes made)

. summarize Q29c

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29c |      1,855    2.979515    1.349662         1          5

. replace Q29c = r(mean) if Q29c == .
variable Q29c was byte now float
(197 real changes made)

. summarize Q29d

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29d |      1,872    2.995192     1.18176         1          5

. replace Q29d = r(mean) if Q29d == .
variable Q29d was byte now float
(180 real changes made)

. summarize Q29e

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29e |      1,866    3.002144    1.413074         1          5

. replace Q29e = r(mean) if Q29e == .
variable Q29e was byte now float
(186 real changes made)

. summarize Q29f

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29f |      1,868     2.88651    1.129992         1          5

. replace Q29f = r(mean) if Q29f == .
variable Q29f was byte now float
(184 real changes made)

. summarize Q29g

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29g |      1,867    2.846277    1.231265         1          5

. replace Q29g = r(mean) if Q29g == .
variable Q29g was byte now float
(185 real changes made)

. summarize Q29h

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Q29h |      1,860    2.950538    1.158139         1          5

. replace Q29h = r(mean) if Q29h == .
variable Q29h was byte now float
```

```
(192 real changes made)

. rename Q29c work_proximity

. //generate job = (Q28a + Q28b + Q28c + Q28d + Q28e + Q28f + Q28g + Q29a + Q29b + Q29c + Q29d +
Q29e + Q29f + Q29g + Q29
> h)/15
. generate job = (Q28a + Q28b + Q28c + Q28d + Q28e + Q28f + Q28g)/7

. generate job_satisfaction = job

. summarize job

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         job |     2,052    3.013849    .7719429         1          5

. replace job_satisfaction = 1 if job > 3
(1,068 real changes made)

. replace job_satisfaction = 0 if job <= 3
(984 real changes made)


.
. //alpha Q28a-Q29h
. alpha Q28a-Q28g

Test scale = mean(unstandardized items)

Average interitem covariance:      .5046498
Number of items in the scale:             7
Scale reliability coefficient:      0.8469


.
. drop Q3 Q4 Q5 Q6 Q8 Q19 Q20

. drop Q1 Q2 Q11 Q13 Q12 Q14a Q14b Q14c Q14d Q14e Q14f Q14g Q14j Q14k Q14l Q14m Q14n Q14o

. drop Q17 Q21 Q22 Q23a Q23b Q24 Q25 Q26 Q27 Q28a Q28b Q28c Q28d Q28e Q28f Q28g

. drop Q29a Q29b Q29d Q29e Q29f Q29g Q29h Q30a Q30b Q30d Q30e Q30g Q30h Q30i Q30j Q30k

. drop Q32 Q33

. drop job Q16 Q18 transport1

. ////////////////////////////////////////////////////////////////
> // end of data clean-up process
.
. summarize

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
pub_trans_~y |     2,052    3.255213    1.086029         1          5
  congestion |     2,052    2.718616    1.039988         1          5
work_proxi~y |     2,052    2.979515    1.283208         1          5
travel_imp~e |     2,052    3.859668    .9592459         1          5
work_trave~e |     2,052    3.648387       1.021         1          5
-------------+--------------------------------------------------------
life_satis~n |     2,052    2.951518    .9757139         1          5
yrs_employed |     2,052    9.021482    5.464283        .5         15
      income |     2,052    25748.18     10318.8      5720      55241
         age |     2,052    41.50008    11.05147        21         74
      age_sq |     2,045    1844.743    978.7284       441       5476
-------------+--------------------------------------------------------
         sex |     2,052    .4234265    .4916883         0          1
relationship |     2,052    .6931475    .4584834         0          1
    children |     2,052    .5607708    .4930157         0          1
satisfaction |     2,052    .6583593    .4736818         0          1
       train |     2,052    2.718324    .4972822         1          3
-------------+--------------------------------------------------------
```

```
    education |      2,052    2.534239    1.101837           1          6
       trans1 |      2,052    1.359162     .8274595          0          2
job_satisf~n |      2,052     .5204678    .4997027          0          1

.
. probit job_satisfaction satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_import
> ance work_travel_ease work_proximity income age age_sq sex relationship i.train education
i.trans1,robust

Iteration 0:   log pseudolikelihood = -1415.9597
Iteration 1:   log pseudolikelihood = -1296.4749
Iteration 2:   log pseudolikelihood =  -1296.344
Iteration 3:   log pseudolikelihood =  -1296.344

Probit regression                                Number of obs    =      2,045
                                                 Wald chi2(19)    =     217.66
                                                 Prob > chi2      =     0.0000
Log pseudolikelihood =  -1296.344                Pseudo R2        =     0.0845

------------------------------------------------------------------------------------
                  |              Robust
 job_satisfaction |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+-----------------------------------------------------------------
     satisfaction |   .2591297   .0634221     4.09   0.000     .1348246    .3834348
pub_trans_quality |   .1363437   .0295617     4.61   0.000     .0784039    .1942835
         children |    .108547   .0709842     1.53   0.126    -.0305795    .2476735
       congestion |   .1136604   .0301621     3.77   0.000     .0545438     .172777
life_satisfaction |  -.0749545   .0297842    -2.52   0.012    -.1333304   -.0165786
     yrs_employed |  -.0032802   .0061555    -0.53   0.594    -.0153447    .0087844
 travel_importance |   .0189328   .0380542     0.50   0.619    -.0556521    .0935176
 work_travel_ease |   .0420978   .0358022     1.18   0.240    -.0280732    .1122688
   work_proximity |   .0634031   .0228823     2.77   0.006     .0185545    .1082516
           income |   .0000288   3.22e-06     8.95   0.000     .0000225    .0000351
              age |  -.0212006   .0180474    -1.17   0.240    -.0565729    .0141717
           age_sq |      .0002   .0001983     1.01   0.313    -.0001886    .0005887
              sex |  -.2000173   .0604519    -3.31   0.001    -.3185008   -.0815338
     relationship |  -.0404363    .067735    -0.60   0.551    -.1731946    .0923219
                  |
            train |
               No |  -.6380189   .1955144    -3.26   0.001     -1.02122   -.2548177
              Yes |  -.5958631   .1901787    -3.13   0.002    -.9686066   -.2231196
                  |
        education |   .0674359   .0274517     2.46   0.014     .0136315    .1212403
                  |
           trans1 |
           Public |   .0283593   .0925539     0.31   0.759    -.1530431    .2097616
            Drive |  -.2091751   .0748236    -2.80   0.005    -.3558267   -.0625236
                  |
            _cons |  -.6705619   .4533203    -1.48   0.139    -1.559053    .2179296
------------------------------------------------------------------------------------

. //probit job_satisfaction satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_impo
> rtance work_travel_ease work_proximity income age sex relationship train education
trans2,robust
. //probit job_satisfaction satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_impo
> rtance work_travel_ease work_proximity income age sex relationship train education trans1
trans2,robust
. //probit job_satisfaction satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_impo
> rtance work_travel_ease work_proximity income age sex relationship train education,robust
. margins, dydx(*)

Average marginal effects                         Number of obs    =      2,045
Model VCE    : Robust

Expression   : Pr(job_satisfaction), predict()
dy/dx w.r.t. : satisfaction pub_trans_quality children congestion life_satisfaction yrs_employed
travel_importance
```

```
            work_travel_ease work_proximity income age age_sq sex relationship 2.train 3.train
education 1.trans1
            2.trans1


---------------------------------------------------------------------------------
                 |            Delta-method
                 |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+---------------------------------------------------------------
    satisfaction |   .0939068   .0226908     4.14   0.000     .0494337    .1383799
pub_trans_quality |    .04941    .010541     4.69   0.000      .02875      .07007
        children |   .0393367   .0256719     1.53   0.125    -.0109794    .0896527
      congestion |   .0411897   .0108126     3.81   0.000     .0199974    .0623821
life_satisfaction |  -.027163   .0107514    -2.53   0.012    -.0482353   -.0060907
    yrs_employed |  -.0011887   .0022301    -0.53   0.594    -.0055596    .0031822
 travel_importance |  .0068611   .0137853     0.50   0.619    -.0201575    .0338797
 work_travel_ease |    .015256   .0129639     1.18   0.239    -.0101528    .0406647
   work_proximity |   .0229768   .0082522     2.78   0.005     .0068028    .0391508
          income |   .0000104   1.10e-06     9.50   0.000     8.29e-06    .0000126
             age |   -.007683   .0065343    -1.18   0.240    -.0204899     .005124
          age_sq |   .0000725   .0000718     1.01   0.313    -.0000683    .0002132
             sex |  -.0724849    .021709    -3.34   0.001    -.1150337   -.0299361
    relationship |  -.0146538   .0245402    -0.60   0.550    -.0627517     .033444
                 |
           train |
              No |  -.2191524   .0609073    -3.60   0.000    -.3385286   -.0997763
             Yes |  -.2038484   .0584247    -3.49   0.000    -.3183586   -.0893381
                 |
       education |   .0244383    .009904     2.47   0.014     .0050269    .0438497
                 |
          trans1 |
          Public |   .0101872   .0332334     0.31   0.759    -.0549491    .0753235
           Drive |  -.0759055   .0270008    -2.81   0.005    -.1288261   -.0229848
---------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

.
. probit satisfaction job_satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_import
> ance work_travel_ease work_proximity income age age_sq sex relationship i.train education
i.trans1,robust

Iteration 0:   log pseudolikelihood = -1312.7005
Iteration 1:   log pseudolikelihood = -1189.1064
Iteration 2:   log pseudolikelihood = -1188.7229
Iteration 3:   log pseudolikelihood = -1188.7229

Probit regression                               Number of obs   =      2,045
                                                Wald chi2(19)   =     227.58
                                                Prob > chi2     =     0.0000
Log pseudolikelihood = -1188.7229               Pseudo R2       =     0.0944


---------------------------------------------------------------------------------
                 |             Robust
    satisfaction |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+---------------------------------------------------------------
 job_satisfaction |   .2611661   .0621132     4.20   0.000     .1394263    .3829058
pub_trans_quality |   .1869038   .0302912     6.17   0.000     .1275341    .2462735
        children |   .1028822   .0731293     1.41   0.159    -.0404485    .2462129
      congestion |   .1585664   .0311072     5.10   0.000     .0975974    .2195353
life_satisfaction |  .0074104   .0302692     0.24   0.807    -.0519162     .066737
    yrs_employed |   .0048189   .0064074     0.75   0.452    -.0077395    .0173772
 travel_importance |  .0497645   .0396253     1.26   0.209    -.0278997    .1274286
 work_travel_ease |  -.0632043   .0378078    -1.67   0.095    -.1373062    .0108976
   work_proximity |   .0076496   .0236666     0.32   0.747    -.0387361    .0540354
          income |   8.69e-06   3.32e-06     2.62   0.009     2.19e-06    .0000152
             age |  -.0189366   .0186143    -1.02   0.309    -.0554198    .0175467
          age_sq |    .000267   .0002037     1.31   0.190    -.0001323    .0006663
             sex |  -.1057789   .0621114    -1.70   0.089     -.227515    .0159572
    relationship |   .3209668   .0699578     4.59   0.000      .183852    .4580816
                 |
           train |
```

```
          No  |   .3341056    .2115034     1.58   0.114    -.0804334    .7486446
         Yes  |   .6023461    .2059518     2.92   0.003     .1986879    1.006004
              |
   education  |   .0587216    .0284313     2.07   0.039     .0029973    .1144458
              |
       trans1 |
      Public  |  -.1673986     .096405    -1.74   0.082    -.3563489    .0215517
       Drive  |  -.0932303    .0788408    -1.18   0.237    -.2477554    .0612947
              |
       _cons  |  -1.524988      .47128    -3.24   0.001     -2.44868   -.6012961
------------------------------------------------------------------------------

. margins, dydx(*)

Average marginal effects                        Number of obs     =      2,045
Model VCE    : Robust

Expression   : Pr(satisfaction), predict()
dy/dx w.r.t. : job_satisfaction pub_trans_quality children congestion life_satisfaction
yrs_employed travel_importance
               work_travel_ease work_proximity income age age_sq sex relationship 2.train 3.train
education 1.trans1
               2.trans1


------------------------------------------------------------------------------
                   |            Delta-method
                   |     dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------
 job_satisfaction  |   .0861031   .0202487     4.25   0.000     .0464165    .1257898
pub_trans_quality  |   .0616198   .0097441     6.32   0.000     .0425217     .080718
         children  |    .033919   .0240695     1.41   0.159    -.0132564    .0810943
       congestion  |   .0522773   .0100888     5.18   0.000     .0325037    .0720509
life_satisfaction  |   .0024431   .0099781     0.24   0.807    -.0171136    .0219999
     yrs_employed  |   .0015887   .0021118     0.75   0.452    -.0025503    .0057278
travel_importance  |   .0164067   .0130512     1.26   0.209    -.0091731    .0419866
 work_travel_ease  |  -.0208377   .0124397    -1.68   0.094     -.045219    .0035436
   work_proximity  |    .002522   .0078019     0.32   0.747    -.0127695    .0178135
           income  |   2.86e-06   1.09e-06     2.63   0.008     7.32e-07    5.00e-06
              age  |  -.0062432   .0061355    -1.02   0.309    -.0182684    .0057821
           age_sq  |    .000088   .0000671     1.31   0.190    -.0000436    .0002196
              sex  |   -.034874   .0204136    -1.71   0.088    -.0748838    .0051359
     relationship  |   .1058187   .0227303     4.66   0.000     .0612681    .1503693
                   |
            train  |
              No   |   .1209914   .0769091     1.57   0.116    -.0297476    .2717305
             Yes   |   .2123409   .0749149     2.83   0.005     .0655103    .3591715
                   |
        education  |   .0193598   .0093465     2.07   0.038      .001041    .0376785
                   |
           trans1  |
          Public   |  -.0551791   .0318249    -1.73   0.083    -.1175549    .0071966
           Drive   |  -.0303022      .0254    -1.19   0.233    -.0800854    .0194809
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

.
. save CBP_survey_clean.dta, replace
file CBP_survey_clean.dta saved

. log close
      name:  <unnamed>
       log:  D:\UMaine\Spring (2018)\ECO 485\Stata\Empirical_Project_3.log
  log type:  text
 closed on:   6 May 2018, 17:53:07
-----------------------------------------------------------------------------------------------
-----------------------
```

Appendix B – Stata do-file

```stata
. /*@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
> @ Author: Antonio Jurlina           @
> @ Date: 4/23/2018                    @
> @ Filename: EMPIRICAL_PROJECT_3.do @
> @@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@*/

cls
clear //clear previous data

cd "D:\UMaine\Spring (2018)\ECO 485\Stata" //setting my directory

// Store results in a log file (diary)
log using "D:\UMaine\Spring (2018)\ECO 485\Stata\Empirical_Project_3", replace text

use CBP_survey.dta //choose data

// following commands represent my data clean-up process
//////////////////////////////////////////////////////////////////

encode Q3, generate(yrs_employed)
replace yrs_employed = 0.5 if Q3 == "Less than 1 year"
replace yrs_employed = 2 if Q3 == "1 to 3 years"
replace yrs_employed = 4 if Q3 == "3 to 5 years"
replace yrs_employed = 7.5 if Q3 == "5 to 10 years"
replace yrs_employed = 15 if Q3 == "More than 10 years"

encode Q4, generate(income)
replace income = 5720 if Q4 == "£0 - £11,440 per year"
replace income = 12480.5 if Q4 == "£11,441 - £13,520 per year"
replace income = 14820.5 if Q4 == "£13,521 - £16,120 per year"
replace income = 17420.5 if Q4 == "£16,121 - £18,720 per year"
replace income = 20540.5 if Q4 == "£18,721 - £22,360 per year"
replace income = 25220.5 if Q4 == "£22,361 - £28,080 per year"
replace income = 31720.5 if Q4 == "£28,081 - £35,360 per year"
replace income = 40300.5 if Q4 == "£35,361 - £45,240 per year"
replace income = 55241 if Q4 == "£45,241 or more per year"

drop if Q5 == "25- 40" | Q5 == "25- 41" | Q5 == "25- 42" | Q5 == "25- 43"
drop if Q5 == "25- 44" | Q5 == "25- 45" | Q5 == "25- 46" | Q5 == "25- 47" | Q5 == "25-
48" | Q5 == "25- 49"
encode Q5, generate(age)
replace age = 21 if Q5 == "18- 24"
replace age = 32 if Q5 == "25- 39"
replace age = 47 if Q5 == "40- 54"
replace age = 59.5 if Q5 == "55- 64"
replace age = 74 if Q5 == "65"

generate age_sq = age * age

label define sex 1 "Male" 0 "Female"
encode Q6, generate(sex)

generate relationshipy = Q7
replace relationshipy = "Yes" if Q7 == "Co-habiting"
replace relationshipy = "Yes" if Q7 == "Married"
replace relationshipy = "No" if Q7 == "Single"
label define relationship 1 "Yes" 0 "No"
encode relationshipy, generate(relationship)
drop relationshipy Q7
```

```
label define children 1 "Yes" 0 "No"
encode Q8, generate(children)

rename Q14h pub_trans_quality
rename Q14i congestion

generate satisfactiony = Q15
label define satisfaction 1 "Satisfied" 0 "Unsatisfied"
encode satisfactiony, generate(satisfaction)
drop satisfactiony Q15

encode Q19, generate(train)
label define education 1 "High School" 2 "Associates degree" 3 "Bachelors Degree BA,
BSc" 4 "Masters Degree" 5 "Professional degree" 6 "PhD"
encode Q20, generate(education)

rename Q30c travel_importance
rename Q30f work_travel_ease
rename Q31 life_satisfaction

summarize pub_trans_quality
replace pub_trans_quality = r(mean) if pub_trans_quality == .
summarize congestion
replace congestion = r(mean) if congestion == .
summarize yrs_employed
replace yrs_employed = r(mean) if yrs_employed == .
summarize income
replace income = r(mean) if income == .

summarize age
replace age = r(mean) if age == .
summarize sex
replace sex = r(mean) if sex == .
summarize relationship
replace relationship = r(mean) if relationship == .
summarize satisfaction
replace satisfaction = r(mean) if satisfaction == .
summarize train
drop if train == .
summarize education
replace education = r(mean) if education == .
summarize life_satisfaction
replace life_satisfaction = r(mean) if life_satisfaction == .
summarize children
replace children = r(mean) if children == .
summarize work_travel_ease
replace work_travel_ease = r(mean) if work_travel_ease == .
summarize travel_importance
replace travel_importance = r(mean) if travel_importance == .

generate transport1 = Q16
replace transport1 = "Hippie" if Q16 ==  "Walk" | Q16 == "Cycle"
replace transport1 = "Public" if Q16 == "Train" | Q16 == "Bus"
replace transport1 = "Drive" if Q16 == "Car /Motorcycle"
label define trans1 0 "Hippie" 1 "Public" 2 "Drive"
encode transport1, generate(trans1)
//generate transport2 = Q18
//replace transport2 = "Hippie" if Q18 ==  "Walk" | Q18 == "Cycle"
//replace transport2 = "Public" if Q18 == "Train" | Q18 == "Bus"
//replace transport2 = "Drive" if Q18 == "Car /Motorcycle"
//encode transport2, generate(trans2)
```

```
drop if trans1 == .


//generate transporty = Q16
//replace transporty = "Drive" if Q16 == "Car /Motorcycle"
//generate transports = Q18
//replace transports = "Drive" if Q18 == "Car /Motorcycle"
//generate transport = transporty + transports
//replace transport = "Bus" if transport == "BusBus"
//replace transport = "Drive" if transport == "DriveDrive"
//replace transport = "Train" if transport == "TrainTrain"
//replace transport = "Walk" if transport == "WalkWalk"
//replace transport = "Bus / Cycle" if transport == "BusCycle"
//replace transport = "Bus / Cycle" if transport == "CycleBus"
//replace transport = "Bus / Drive" if transport == "BusDrive" & transport ==
"DriveBus"
//replace transport = "Bus / Drive" if transport == "BusDrive" | transport ==
"DriveBus"
//replace transport = "Bus / Train" if transport == "BusTrain" | transport ==
"TrainBus"
//replace transport = "Bus / Walk" if transport == "BusWalk" | transport == "WalkBus"
//replace transport = "Cycle / Drive" if transport == "CycleDrive" | transport ==
"DriveCycle"
//replace transport = "Cycle / Train" if transport == "CycleTrain" | transport ==
"TrainCycle"
//replace transport = "Cycle / Walk" if transport == "CycleWalk" | transport ==
"WalkCycle"
//replace transport = "Cycle / Train" if transport == "Cycletrain"
//replace transport = "Drive / Train" if transport == "DriveTrain" | transport ==
"TrainDrive"
//replace transport = "Drive / Walk" if transport == "DriveWalk" | transport ==
"WalkDrive"
//replace transport = "Train / Walk" if transport == "TrainWalk" | transport ==
"WalkTrain"

//encode transport, generate(transportation)
//drop if transportation == .
//drop transport transporty transports Q16 Q18

summarize Q28a
replace Q28a = r(mean) if Q28a == .
summarize Q28b
replace Q28b = r(mean) if Q28b == .
summarize Q28c
replace Q28c = r(mean) if Q28c == .
summarize Q28d
replace Q28d = r(mean) if Q28d == .
summarize Q28e
replace Q28e = r(mean) if Q28e == .
summarize Q28f
replace Q28f = r(mean) if Q28f == .
summarize Q28g
replace Q28g = r(mean) if Q28g == .
summarize Q29a
replace Q29a = r(mean) if Q29a == .
summarize Q29b
replace Q29b = r(mean) if Q29b == .
summarize Q29c
replace Q29c = r(mean) if Q29c == .
summarize Q29d
replace Q29d = r(mean) if Q29d == .
summarize Q29e
replace Q29e = r(mean) if Q29e == .
```

```
summarize Q29f
replace Q29f = r(mean) if Q29f == .
summarize Q29g
replace Q29g = r(mean) if Q29g == .
summarize Q29h
replace Q29h = r(mean) if Q29h == .
rename Q29c work_proximity
//generate job = (Q28a + Q28b + Q28c + Q28d + Q28e + Q28f + Q28g + Q29a + Q29b + Q29c
+ Q29d + Q29e + Q29f + Q29g + Q29h)/15
generate job = (Q28a + Q28b + Q28c + Q28d + Q28e + Q28f + Q28g)/7
generate job_satisfaction = job
summarize job
replace job_satisfaction = 1 if job > 3
replace job_satisfaction = 0 if job <= 3

//alpha Q28a-Q29h
alpha Q28a-Q28g

drop Q3 Q4 Q5 Q6 Q8 Q19 Q20
drop Q1 Q2 Q11 Q13 Q12 Q14a Q14b Q14c Q14d Q14e Q14f Q14g Q14j Q14k Q14l Q14m Q14n
Q14o
drop Q17 Q21 Q22 Q23a Q23b Q24 Q25 Q26 Q27 Q28a Q28b Q28c Q28d Q28e Q28f Q28g
drop Q29a Q29b Q29d Q29e Q29f Q29g Q29h Q30a Q30b Q30d Q30e Q30g Q30h Q30i Q30j Q30k
drop Q32 Q33
drop job Q16 Q18 transport1
//////////////////////////////////////////////////////////////////////
// end of data clean-up process

summarize

probit job_satisfaction satisfaction pub_trans_quality children congestion
life_satisfaction yrs_employed travel_importance work_travel_ease work_proximity
income age age_sq sex relationship i.train education i.trans1,robust
//probit job_satisfaction satisfaction pub_trans_quality children congestion
life_satisfaction yrs_employed travel_importance work_travel_ease work_proximity
income age sex relationship train education trans2,robust
//probit job_satisfaction satisfaction pub_trans_quality children congestion
life_satisfaction yrs_employed travel_importance work_travel_ease work_proximity
income age sex relationship train education trans1 trans2,robust
//probit job_satisfaction satisfaction pub_trans_quality children congestion
life_satisfaction yrs_employed travel_importance work_travel_ease work_proximity
income age sex relationship train education,robust
margins, dydx(*)

probit satisfaction job_satisfaction pub_trans_quality children congestion
life_satisfaction yrs_employed travel_importance work_travel_ease work_proximity
income age age_sq sex relationship i.train education i.trans1,robust
margins, dydx(*)

save CBP_survey_clean.dta, replace
log close
```