

Transport Data Analysis and Modeling Methodologies

Lab Session #3 (Seemingly Unrelated Regression Estimation)

A survey of 206 people was conducted on the campus of Purdue University in the Fall of 2007. One of the key questions of the survey was to find out how fast people drove on interstate highways with speed limits of 55 mph, 65 mph and 70 mph.

Your task is to estimate a seemingly unrelated regression model to determine the normal driving speed of individuals in this data sample. The equation system:

$$\begin{aligned}Speed_{70} &= \beta_{70}Z + \alpha_{70}X + \varepsilon_{70} \\Speed_{65} &= \beta_{65}Z + \alpha_{65}X + \varepsilon_{65} \\Speed_{55} &= \beta_{55}Z + \alpha_{55}X + \varepsilon_{55}\end{aligned}$$

In these equations, $Speed_{70}$, $Speed_{65}$ and $Speed_{55}$ are the number of miles per hour respondents normally drive above the speed limit (with little traffic) for 70, 65, and 55 mph speed limits, respectively. These variables can take on positive values if respondents normally drive above the speed limit and negative values if they normally drive below it. Also in these equations, Z is a vector of driver and driver-household characteristics, X is a vector of vector of driver preferences and opinions, β 's, α 's, are vectors of estimable parameters, and ε 's are disturbance terms. Include:

1. The results of your best model specification.
2. A discussion of the logical process that led you to the selection of your final specification (the theory behind the inclusion of your selected variables). Include t -statistics and justify the signs of your variables.

PURDUE

UNIVERSITY

A graduate class at Purdue University is undertaking a survey in an attempt to gain a better understanding of public opinions on a wide variety of subjects. We appreciate your time in completing this survey -- your responses will be strictly confidential.

Section A: Your Opinions and Preferences

- ① 1. On an interstate with a 70mph speed limit and little traffic, about how fast do you normally drive? _____ mph
- ② 2. On an interstate with a 65mph speed limit and little traffic, about how fast do you normally drive? _____ mph
- ③ 3. On an interstate with a 55mph speed limit and little traffic, about how fast do you normally drive? _____ mph
- ④ 4. About how many miles per year do you drive? _____ miles/yr
- ⑤ 5. How many vehicle accidents have you had (while driving) in the last 5 years? _____ years
- ⑥ 6. When braking quickly to avoid a collision do you usually?
1 Brake and steer 2 Brake only 3 Not sure/Depends on situation
- ⑦ 7. Do you have:
1 An engineering background 2 No engineering background
- ⑧ 8. When putting on socks and shoes do you normally:
1 Put both socks on then both shoes 2 One sock, one shoe 3 Not sure/Varies
- ⑨ 9. When waking up in the morning, what leg reaches the floor first from bed:
1 Right 2 Left 3 Both legs at the same time 4 Not sure/Varies
- ⑩ 10. Do you prefer to buy online or in a store? 1 Online 2 Store 3 No preference
- ⑪ 11. Do you think Britney Spears should get full custody of her kids? 1 Yes 2 No 3 Don't care
- ⑫ 12. Do you prefer an automatic or manual transmission in your vehicle?
1 Automatic 2 Manual 3 No preference
- ⑬ 13. Do you sometimes smoke cigarettes? 1 Yes 2 No
- ⑭ 14. Do you prefer to text or talk on a cell phone? 1 Text 2 Talk 3 No preference
- ⑮ 15. At Purdue (West Lafayette), what is your favorite season?
1 Spring 2 Summer 3 Fall 4 Winter
- ⑯ 16. During severe winter weather (snow) do you use your car or bus? 1 Car 2 Bus 3 Not sure/Varies
- ⑰ 17. In money matters (investments, gambling, etc) do you consider yourself?
1 Conservative 2 Moderate 3 Risky 4 Don't know
- ⑱ 18. What is the first thing you drink in the morning?
1 Water 2 Juice 3 Milk 4 Soft drink 5 Coffee 6 Energy drink 7 Other
- ⑲ 19. Do you normally skip breakfast? 1 Yes 2 No

(Please Turn Over)

Section A: Continued

20. What drives your decision to accept a job offer?
1 Compensation 2 Career progression opportunities 3 Job location 4 Not sure
21. When showering, do you normally use: 1 Body wash 2 Bar soap 3 Unsure
22. While driving and using your cell phone, do you:
1 Use a headset 2 Use one hand on the phone - one hand on wheel 3 Other
4 Do not use phone while driving
23. When your car is the first in line at a traffic signal (dry road) and the light turns green do you normally:
1 accelerate briskly 2 accelerate moderately 3 accelerate slowly
24. What is the fastest that you (as a driver) have ever driven on an interstate, rural or urban road?
1 less than 70mph 2 70-79mph 3 80-89mph 4 90-99mph 5 100-109mph
6 110-119 mph 7 120-129mph 8 130-139mph 9 more than 140mph
-

Section B: Additional Questions About Yourself

25. Are you? 1 Female 2 Male
26. Are you? 1 Married 2 Single 3 Separated 4 Divorced 5 Other
27. What is your age? _____
28. Are you currently?
1 Not affiliated with Purdue 2 Purdue undergraduate 3 Purdue graduate
4 Purdue faculty 5 Purdue staff (other than RA/TA/faculty)
29. What is your highest completed level of education?
1 Some high school 2 High school diploma 3 Technical college degree (A.A.)
4 College degree 5 Post graduate degree
30. Please indicate your Race/Ethnicity
1 African American 2 American Indian 3 Asian 4 Caucasian
5 Hispanic/White 6 Hispanic/Non-white 7 Other 8 I would rather not answer
31. What is the approximate annual household income of the household you consider home?
1 no income 2 under \$10,000 3 \$10,000-\$19,999 4 \$20,000-\$29,999
5 \$30,000-\$39,999 6 \$40,000-\$49,999 7 \$50,000-\$74,999 8 \$75,000-\$100,000 9 Over \$100,000
32. Including yourself, how many people live in the household you consider home? _____
33. How many children, in the household you consider home, are under age 6? _____
34. How many children, in the household you consider home, are aged 6 to 16? _____
35. How many people living, in the household you consider home, work outside the home? _____
36. How many licensed and operable motor vehicles does your "home" household have? _____
37. Are you a licensed driver? 1 Yes 2 No
- 37a. If you are licensed to drive, how many years have you had a license? _____ years
38. Did you lie about your response to any of the previous questions on this survey? 1 Yes 2 No
-

40 STUDENT ID

(Thank you)

```

--> read;nvar=40;nobs=206;file=D:class-07.txt$
--> skip
--> create;ageL=x27-x38$
--> create;if(ageL>17)late=1$
--> create;if(x25=2)male=1$
--> create;if(x26=1)married=1$
--> create;if(x23=1)brisk=1$
--> create;mo70=x1-70$
--> create;mo65=x2-65$
--> create;mo55=x3-55$
--> dstat;rhs=mo70,mo65,mo55$

```

Descriptive Statistics
All results based on nonmissing observations.

```

=====
Variable      Mean      Std.Dev.   Minimum   Maximum   Cases Missing
=====
All observations in current sample
-----
MO70          5.45366   7.08727   -30.0000   20.0000   205          1
MO65          5.69024   7.15568   -25.0000   25.0000   205          1
MO55          6.24757   7.66951   -40.0000   25.0000   206          0

```

```

--> reject;x1=-999$
--> reject;x2=-999$
--> reject;x3=-999$
--> reject;x27=-999$
--> reject;x38=-999$
--> reject;x25=-999$
--> reject;x37=2$
--> Sure;lhs=mo70,mo65,mo55
      ;eq1=one,x27,x32,late,brisk,x24
      ;eq2=one,male,x27,x32,late,brisk,x24
      ;eq3=one,x33,late,brisk,x24$

```

```

*****
* NOTE: Deleted      3 observations with missing data. N is now      191 *
*****

```

```

+-----+
| Estimates for equation: MO70
| Generalized least squares regression
| Model was estimated Jan 23, 2008 at 11:22:26AM
| LHS=MO70      Mean          = 5.748691
|                Standard deviation = 6.822539
| WTS=none      Number of observs. = 191
| Model size    Parameters     = 6
|                Degrees of freedom = 185
| Residuals     Sum of squares = 7071.643
|                Standard error of e = 6.182645
| Fit           R-squared      = .1744634
|                Adjusted R-squared = .1521516
| Model test    F[ 5, 185] (prob) = 7.82 (.0000)
| Diagnostic    Log likelihood = -615.9227
|                Restricted(b=0) = -637.2802
|                Chi-sq [ 5] (prob) = 42.72 (.0000)
| Info criter.  LogAmemiya Prd. Crt. = 3.674423
|                Akaike Info. Criter. = 3.674402
| Not using OLS or no constant. Rsqd & F may be < 0.
| Log|W|        8.0840 Log-Likelihood = -1585.0730
| Durbin-Watson 1.549 Autocorrelation = .2255
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	1.03937387	1.57273427	.661	.5087	
X27	-.02485828	.02481934	-1.002	.3166	31.4136126
X32	.12286014	.20565573	.597	.5502	2.83246073
LATE	-1.00489400	.90224677	-1.114	.2654	.45549738
BRISK	-.06583367	1.15896419	-.057	.9547	.18848168
X24	1.29620039	.21742426	5.962	.0000	4.32984293

```

-----
Estimates for equation: MO65
Generalized least squares regression
Model was estimated Jan 23, 2008 at 11:22:26AM
LHS=MO65      Mean          =    5.850785
              Standard deviation =    7.007514
WTS=none     Number of observs. =         191
Model size   Parameters      =          7
              Degrees of freedom =         184
Residuals   Sum of squares   =    7614.406
              Standard error of e =    6.432934
Fit          R-squared        =    .1528310
              Adjusted R-squared =    .1252059
Model test   F[ 6, 184] (prob) =    5.53 (.0000)
Diagnostic   Log likelihood    =   -622.9848
              Restricted(b=0)    =   -642.3897
              Chi-sq [ 6] (prob) =   38.81 (.0000)
Info criter. LogAmemiya Prd. Crt. =    3.758855
              Akaike Info. Criter. =    3.758822
Not using OLS or no constant. Rsqd & F may be < 0.
Log|W|      8.0840 Log-Likelihood =   -1585.0730
Durbin-Watson 1.522 Autocorrelation =    .2390
-----

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	2.77846956	1.50616043	1.845	.0651	
MALE	-.08593715	.36576014	-.235	.8142	.62827225
X27	-.01630271	.02089760	-.780	.4353	31.4136126
X32	-.16336038	.17645812	-.926	.3546	2.83246073
LATE	-2.06350679	.93873401	-2.198	.0279	.45549738
BRISK	-.22663173	1.20641004	-.188	.8510	.18848168
X24	1.17412675	.22757164	5.159	.0000	4.32984293

```

+-----+
| Estimates for equation: MO55
| Generalized least squares regression
| Model was estimated Jan 23, 2008 at 11:22:26AM
| LHS=MO55      Mean          = 6.649215
|                Standard deviation = 6.953037
| WTS=none      Number of observs. = 191
| Model size    Parameters     = 5
|                Degrees of freedom = 186
| Residuals    Sum of squares  = 7682.128
|                Standard error of e = 6.426645
| Fit          R-squared       = .1411856
|                Adjusted R-squared = .1227165
| Model test    F[ 4, 186] (prob) = 7.64 (.0000)
| Diagnostic    Log likelihood  = -623.8304
|                Restricted(b=0) = -640.8991
|                Chi-sq [ 4] (prob) = 34.14 (.0000)
| Info criter. LogAmemiya Prd. Crt. = 3.746746
|                Akaike Info. Criter. = 3.746735
| Not using OLS or no constant. Rsqd & F may be < 0.
| Log|W|       8.0840 Log-Likelihood = -1585.0730
| Durbin-Watson 1.598 Autocorrelation = .2012
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	3.29385084	1.17903549	2.794	.0052	
X33	-.01507763	.41775487	-.036	.9712	.22513089
LATE	-2.57853632	.94018290	-2.743	.0061	.45549738
BRISK	.51618891	1.20400057	.429	.6681	.18848168
X24	1.02451353	.22509710	4.551	.0000	4.32984293