ARE 231: Part 1: HW1
Due in class on the 14th

Timothy Beatty
2151 SSH
tbeatty@ucdavis.edu

Fall 2018

1

After lecture I post a .dta file on the Canvas. Unzip it to get the Stata data set. Note that these are "real-ish" data, but they have been created to have approximately the covariance structure of some real data. Here is a description of the variables

Item	EXPENDITURES:	PRICES:
food at home	xfath	pfath
food in restaurants	xrest	prest
household operations	xhhop	phhop
car operations	xcaruse	pcaruse
public transportation	xtran	ptran
personal care	xcare	pcare
recreation	xrecr	prcer
tobbaco	xtob	ptob
alcohol	xalc	palc
women's clothing	xwclth	pwclth
men's clothing	xmclth	pmcloth
children's clothing	xkclth	pkclth

OTHER VARIABLES:	
Household net income	nety
Gender of head	hsex
Age of head	hage
Age of spouse (if present)	sage
Number of kids	kids
Number of Adults	ad
Region	region
Time	time

2

Load the data into STATA. Examine the data for outliers, oddities, and possible mis-codes. Check means, maxima, and mininima of continuous variables and frequencies of categorical data against the summary statistics below. Check the demographic characteristics of your sample.

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
region	14994	3.02581	1.246699	1	5
hnety	14994	20652.37	11643.64	-5532	71000
hage	14994	35.20108	8.266935	20	55
hsex	14994	1.328932	.4698407	1	2
snety	9838	11913.58	11407.85	-15000	70000
sage	9838	33.50508	7.63836	20	55
nety	14994	28502.31	15939.8	5050.627	71455.74
xfath	14994	3242.105	1905.458	0	17448.54
xrest	14994	1283.318	1319.746	0	16117.34
xhhop	14994	1662.291	1613.549	0	23814.82
xwclth	14994	846.2386	1035.544	0	24391.79
xmclth	14994	567.9294	717.4686	0	16857.31
xkclth	14994	53.68904	153.0824	0	2506.821
xcaruse	14994	2110.811	1805.092	0	17682.57
xtran	14994	386.751	659.7086	0	10232.39
xcare	14994	649.4383	494.8456	0	4532.063
xrecr	14994	1634.237	1710.11	0	23337.76
xtob	14994	501.8776	811.7445	0	8684.721
xalc	14994	580.4715	830.7408	0	15208.5
pfath	14994	.8366482	.2448178	.3265014	1.14347
prest	14994	.7594858	.2576474	.2619048	1.152542
phhop	14994	.7911728	.248994	.2635426	1.081685
pcaruse	14994	.7115704	.2735375	.1962521	1.31414
ptran	14994	.7560718	.3083149	.1912682	1.244134
pcare	14994	.825366	.213055	.3776119	1.039801
precr	14994	.7760642	.2356293	.3125541	1.045887
ptob	14994	.5463716	.3018981	.1021921	1.051254
palc	14994	.7670915	.2809425	.2382838	1.178878
pwclth	14994	.8597817	.1974238	.4707775	1.108847
pmclth	14994	.8823521	.2028664	.5084102	1.164948
pkclth	14994	.7977848	.19513	.406382	1.026748
kids	14994	.918434	1.022303	0	3
ad	14994	1.656129	.4750145	1	2
time	14994	3.561558	1.732526	1	6

3

Construct total (nondurable) expenditure x_h as the sum of expenditure on individual commodities. Plot the log of expenditure on 'food at home' against the log of total expenditure. Is 'food at home' a normal good? Is it a luxury or a necessity? Could you investigate the income elasticity of men's clothing this way?

4

Suppose you wanted estimate parameters of a Cobb-Douglas utility function from this data. How would you do so? Based on your preliminary inspection of the data, are the data consistent with an assumption of Cobb-Douglas preferences? Why or why not?

5

Construct budget shares for each good i and household h: $w_{ih} = \frac{x_{ih}}{x_h}$.

Also construct a 'Stone' price index for each household:

$$P_h = \exp\left(\sum_{i=1}^{12} w_{ih} \ln p_{ih}\right)$$

Check that in 1992 the values of this index are about 1. Then construct real total expenditure for each household as,

$$rx_h = \frac{x_h}{P_h}$$

6

Estimate the parameters of a share equation for food at home ($i = f$) of the Working-Leser form:

$$w_{fh} = \alpha_f + \sum_{i=1}^{12} \gamma_i \ln p_i h + \beta_i \ln rx_h + e_{fh}$$

7

Try a RESET test of this parametric functional form. Hint: This can be done with a post-estimation command in Stata: estat ovtest.

8

Create the following aggregate commodities:

- $xtran = xtran + xcaruse$ (transportation = public transportation + car use)
- $xserv = xhhop + xcare$ (services = household operations + personal care)
- $xvices = xalc + xtob$ (vices = alcohol + tobacco)
- $xcloth = xmcltoh + xwcloth + xkcloth$ (clothing)

This gives seven composite commodities: xfath, xfrest, xtran, xserv, xrecr, xvices, xcloth.

Create prices for the four new commodities. To do this, first find the mean budget shares of the 12 original goods in each period (or period and region). Then construct new price for each household for the new commodities by taking weighted geometric means of the household prices. For example:

$$p_{vices,h} = \exp\left(\frac{\overline{w}_{alc,t}}{\overline{w}_{alc,t} + \overline{w}_{tobc,t}} \ln p_{alc,t} + \frac{\overline{w}_{tob,t}}{\overline{w}_{alc,t} + \overline{w}_{tobc,t}} \ln p_{tob,t}\right)$$

where the t subscript refers to the year in which household h is sampled. Check that you have done the construction correctly by confirming that the prices are about one in the base period/region.

9

Estimate by OLS the parameters of the seven equations:

$$w_{i,h} = \alpha_{i0} + \sum_{k=1}^{m} \alpha_{i,k} d_{k,h} + \sum_{j=1}^{7} \gamma_{i,j} \ln p_{j,h} + \beta_i \ln \frac{x_h}{P_h} + u_{i,h}$$

Where the price index P_h is the Stone price index created previously; The $d_{k,h}$ are m household demographic variables that affect demands. Examples are the number of adults or children in the household, the age of the person 1, and so on. You should choose your own set of demographics (keeping in mind the demographic heterogeneity in your sample.) Be careful to include the same demographics in each equation. (Why?) You may need to construct new variables.

10

Check that adding up automatically holds. That is, the columns of estimates sum to one or zero:

$$\sum_{i=1}^{7} \alpha_{i0} = 1; \quad \sum_{i=1}^{7} \alpha_{i,k} = 0 \forall k; \quad \sum_{i=1}^{7} \gamma_{i,j} = \forall j; \quad \sum_{i=1}^{7} \beta_i = 0$$

11

From your estimates, determine whether 'food at home' and 'clothing' are luxuries or necessities. Which good is the most price elastic?

12

Estimate six equations together by Seemingly Unrelated Regression (SURE). To do this, exclude the vices equation. (First try to run all seven equations together as a SURE system and see what happens.) Check that the parameter estimates are exactly the same as the OLS estimates. (Why is this the case?) Check that the dropping some other good gives exactly the same parameter estimates.

13

Look at the error covariance matrix for the SURE estimates. Can you make any sense of the covariances? Which errors are correlated and why?

14

Test for homogeneity in each of the seven demand equations estimated by OLS above. That is, use F-tests on the individual equations to test for:

$$\sum_{j=1}^{n} \gamma_{ij} = 0 \text{ for } i = 1, 2, ...7$$