ARE 231 Homework 1

Tristan Hanon

November 13, 2018

Examining the Data

The data at hand include household expenditure on various categories, prices of goods in those categories, and some demographic characteristics. Table 1 presents summary statistics. In examining these values, the first notable aspect of the data is that all of the expenditure categories have observations with zero expenditure. While a zero observations in some categories could make sense, like tobacco for households without any smokers, there are other categories for which this seems less likely, such as food at home.

There are also some strange values for net income for both the head of the household and the spouse. We can see that the minimum value for `hnety` is $-5,532$ and the minimum value for `snety` is $-15,000$. Without knowing more about the survey procedure it is hard to know whether negative values may represent something for income, but it seems safe to assume that these may have been miscoded somehow.

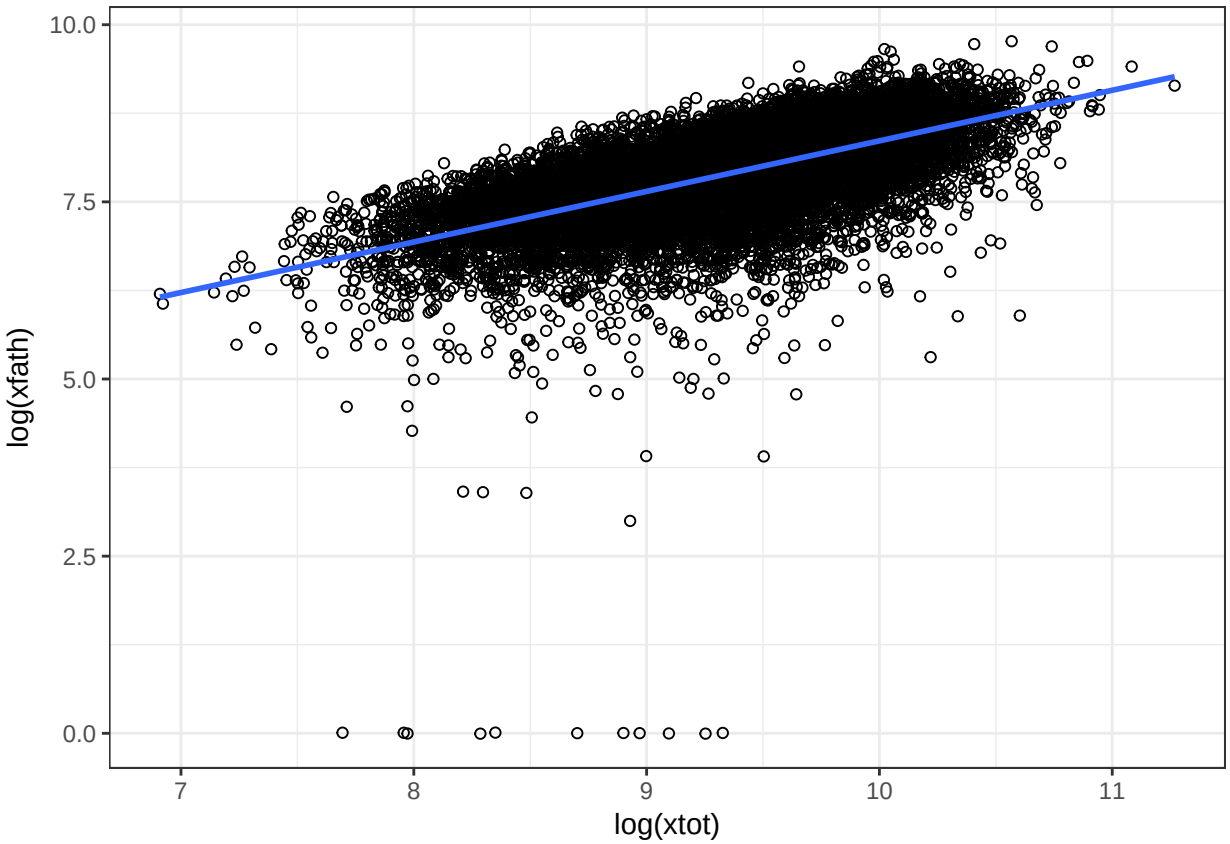
Outside of these two aspects of the data that is difficult to accept, the rest of the summary statistics seem to be unsurprising. Also worth noting is that out of the 14,996 total observations, only 9,838 records had values for `snety` and `sage`, suggesting that there are 5,158 records with only one adult in the household.

Exploring Expenditure on Food at Home

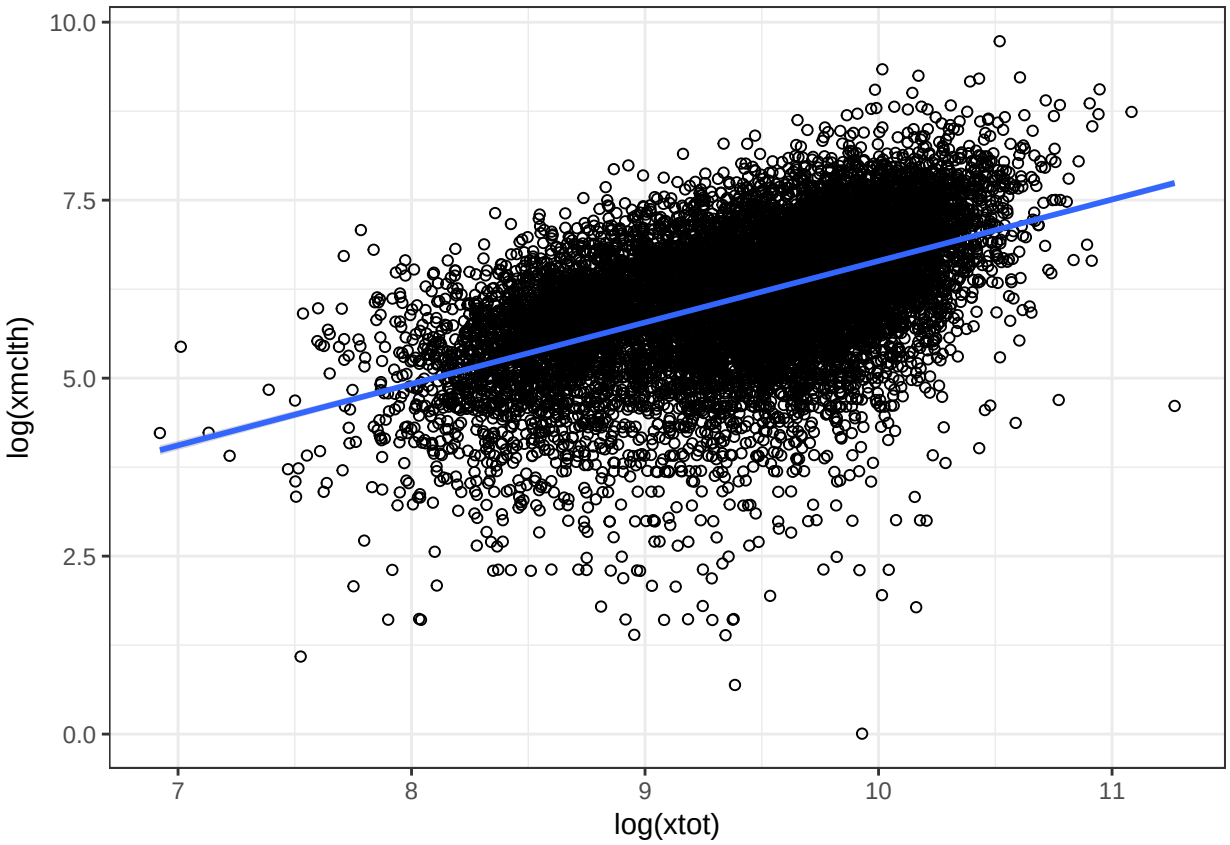
Plotting logged expenditure on food at home against logged total expenditure indicates an upward sloping, relatively linear, relationship. In fact, including a linear regression line in the plot seems to support the claim that the relationship is linear. Since the relationship is upward sloping, this suggests food at home is a normal good, which is of course consistent with expectations. Additionally, since the data is represented linearly, this suggests that food at home is neither a necessity nor a luxury.

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
region	14,994	3.026	1.247	1	5
hnety	14,994	20,652.370	11,643.640	−5,532	71,000
hage	14,994	35.201	8.267	20	55
hsex	14,994	1.329	0.470	1	2
snety	9,838	11,913.580	11,407.850	−15,000	70,000
sage	9,838	33.505	7.638	20	55
nety	14,994	28,502.310	15,939.800	5,050.627	71,455.740
xfath	14,994	3,242.105	1,905.458	0.000	17,448.540
xrest	14,994	1,283.318	1,319.746	0.000	16,117.340
xhhop	14,994	1,662.291	1,613.549	0.000	23,814.820
xwclth	14,994	846.239	1,035.544	0.000	24,391.790
xmclth	14,994	567.929	717.469	0.000	16,857.310
xkclth	14,994	53.689	153.082	0.000	2,506.821
xcaruse	14,994	2,110.811	1,805.092	0.000	17,682.570
xtran	14,994	386.751	659.709	0.000	10,232.390
xcare	14,994	649.438	494.846	0.000	4,532.063
xrecr	14,994	1,634.237	1,710.110	0.000	23,337.760
xtob	14,994	501.878	811.744	0.000	8,684.721
xalc	14,994	580.472	830.741	0.000	15,208.500
pfath	14,994	0.837	0.245	0.327	1.143
prest	14,994	0.759	0.258	0.262	1.153
phhop	14,994	0.791	0.249	0.264	1.082
pcaruse	14,994	0.712	0.274	0.196	1.314
ptran	14,994	0.756	0.308	0.191	1.244
pcare	14,994	0.825	0.213	0.378	1.040
precr	14,994	0.776	0.236	0.313	1.046
ptob	14,994	0.546	0.302	0.102	1.051
palc	14,994	0.767	0.281	0.238	1.179
pwclth	14,994	0.860	0.197	0.471	1.109
pmclth	14,994	0.882	0.203	0.508	1.165
pkclth	14,994	0.798	0.195	0.406	1.027
kids	14,994	0.918	1.022	0	3
ad	14,994	1.656	0.475	1	2
time	14,994	3.562	1.733	1	6



We could also attempt to estimate the income elasticity of men's clothing in a similar way. If we used a similar procedure to the above,



Estimating a Cobb-Douglas Function

Budget Shares and a Stone Price Index

Working-Leser Share Equation

[1] “pfath” “prest” “phhop” “pcaruse” “ptran” “pcare” “precr”
[8] “ptob” “palc” “pwclth” “pmclth” “pkclth”

RESET test

data: wl_reg RESET = 10.863, df1 = 2, df2 = 14978, p-value = 1.93e-05

Composite Commodities and Prices

```
## # A tibble: 30 x 6
## # Groups:   time [?]
##     time region ptransport pserv pvices pcloth
##    <int>  <int>      <dbl> <dbl>  <dbl>  <dbl>
##  1     1      1      0.226 0.302  0.196  0.324
##  2     1      2      0.257 0.313  0.166  0.323
##  3     1      3      0.229 0.305  0.181  0.317
##  4     1      4      0.196 0.294  0.174  0.325
##  5     1      5      0.202 0.324  0.178  0.360
##  6     2      1      0.575 0.821  0.477  0.681
```

Table 2:

	time	region	Mean	SD	Min	Max
1	1	1	1	0	1	1
2	1	2	1	0	1	1
3	1	3	1	0	1	1
4	1	4	1	0	1	1
5	1	5	1	0	1	1
6	2	1	1	0	1	1
7	2	2	1	0	1	1
8	2	3	1	0	1	1
9	2	4	1	0	1	1
10	2	5	1	0	1	1
11	3	1	1	0	1	1
12	3	2	1	0	1	1
13	3	3	1	0	1	1
14	3	4	1	0	1	1
15	3	5	1	0	1	1
16	4	1	1	0	1	1
17	4	2	1	0	1	1
18	4	3	1	0	1	1
19	4	4	1	0	1	1
20	4	5	1	0	1	1
21	5	1	1	0	1	1
22	5	2	1	0	1	1
23	5	3	1	0	1	1
24	5	4	1	0	1	1
25	5	5	1	0	1	1
26	6	1	1	0	1	1
27	6	2	1	0	1	1
28	6	3	1	0	1	1
29	6	4	1	0	1	1
30	6	5	1	0	1	1

Table 3:

	Dependent variable:
	wfath
log(pfath)	-0.125^{**}
	(0.054)
log(prest)	-0.058^{*}
	(0.031)
log(phhop)	0.104^{*}
	(0.053)
log(pcare)	-0.072
	(0.053)
log(ptran)	-0.019
	(0.024)
log(pcaruse)	0.006
	(0.012)
log(precr)	-0.006
	(0.058)
log(palc)	0.147^{***}
	(0.033)
log(ptob)	0.003
	(0.010)
log(pwclth)	-0.158^{*}
	(0.096)
log(pmclth)	0.069
	(0.055)
log(pkclth)	0.060
	(0.072)
log(realx)	-0.085^{***}
	(0.002)
Constant	1.072^{***}
	(0.018)
Observations	14,994
R^2	0.165
Adjusted R^2	0.164
Residual Std. Error	0.115 (df = 14980)
F Statistic	228.030^{***} (df = 13; 14980)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
##  7     2      2      0.722 0.772  0.426  0.603
##  8     2      3      0.652 0.800  0.478  0.619
##  9     2      4      0.582 0.726  0.452  0.701
## 10     2      5      0.691 0.825  0.482  0.699
## # ... with 20 more rows
```

Table 4:

	Dependent variable:					
	wfath	wrest	wserv	wtransport	wrecr	wvices
	(1)	(2)	(3)	(4)	(5)	(6)
ad	0.077***	−0.040***	−0.015***	0.025***	−0.027***	−0.014***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
kids	0.054***	−0.021***	0.019***	−0.018***	−0.004***	−0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(pfath)	−0.071***	0.062***	0.078***	−0.118***	0.032	−0.032
	(0.026)	(0.019)	(0.021)	(0.026)	(0.021)	(0.023)
log(prest)	0.025	−0.019	−0.078***	0.142***	0.066***	−0.076***
	(0.023)	(0.017)	(0.019)	(0.023)	(0.019)	(0.021)
log(pserv)	0.053**	−0.008	−0.029	0.014	−0.096***	0.086***
	(0.026)	(0.019)	(0.021)	(0.026)	(0.021)	(0.023)
log(ptransport)	0.036***	−0.003	−0.012	0.004	−0.015**	−0.006
	(0.009)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)
log(precr)	0.119***	−0.063***	0.067**	−0.020	−0.013	−0.083***
	(0.033)	(0.024)	(0.026)	(0.033)	(0.027)	(0.029)
log(pvices)	0.001	0.002	−0.017**	−0.030***	−0.037***	0.069***
	(0.010)	(0.007)	(0.008)	(0.010)	(0.008)	(0.009)
log(pcloth)	−0.076***	−0.014	0.066***	0.007	0.058***	0.001
	(0.018)	(0.013)	(0.015)	(0.018)	(0.015)	(0.016)
log(xtot/index)	−0.157***	0.049***	−0.009***	0.018***	0.048***	0.008***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Constant	1.595***	−0.302***	0.265***	−0.002	−0.301***	0.043***
	(0.016)	(0.012)	(0.013)	(0.016)	(0.013)	(0.014)
Observations	14,994	14,994	14,994	14,994	14,994	14,994
R^2	0.453	0.173	0.106	0.056	0.093	0.046
Adjusted R^2	0.453	0.172	0.106	0.056	0.093	0.046
Residual Std. Error (df = 14983)	0.093	0.068	0.074	0.093	0.075	0.082
F Statistic (df = 10; 14983)	1,242.218***	313.196***	177.945***	89.186***	154.156***	72.488***

Note: *p<0.1; **p<0.05