Tutorial 3: demand estimation

Yiran Hao

Oct.22.2018

1. Reading the dataset

use "F:\TA ECO310H1 2018FALL\Tutorial3\verboven_cars.dta", clear

This directory depends on where you save the dataset. By typing “clear”, it specifies that it is okay to replace the data in memory, even though the current data have not been saved to disk

2. Summary statistics

sort ma co ye

	ye	ma	co
1	83	Belgium	1
2	84	Belgium	1
3	85	Belgium	1
4	86	Belgium	1
5	87	Belgium	1
6	88	Belgium	1
7	89	Belgium	1
8	90	Belgium	1
9	91	Belgium	1
10	92	Belgium	1
11	93	Belgium	1
12	94	Belgium	1
13	86	Belgium	2
14	87	Belgium	2
15	88	Belgium	2

Sort command arrange the observations of the current data into ascending order based on the values of the variables. Data can be sorted by more than one variable, and in such cases, the sort order is lexicographic. If we sort the data by two variables, for instance, the data are placed in ascending order of the first variable, and then observations that share the same value of the first variable are placed in ascending order of the second variable. Here, we sort market, model, year. The data are in ascending order of market and within each market category, the data are in ascending order of model and within each model code, the data are in ascending order of year. Therefore, the oldest year of model No.1 in Belgium is 1983.

list ma co ye qu in 1/20

list displays the values of variables. Here we list first 20 observations’ market, model code, year, number of sales.

. list ma co ye qu in 1/20

	ma	co	ye	qu
1.	Belgium	1	83	729
2.	Belgium	1	84	1860
3.	Belgium	1	85	1771
4.	Belgium	1	86	2047
5.	Belgium	1	87	2147
6.	Belgium	1	88	2087
7.	Belgium	1	89	1803
8.	Belgium	1	90	2689
9.	Belgium	1	91	2880
10.	Belgium	1	92	2849
11.	Belgium	1	93	1615
12.	Belgium	1	94	1095
13.	Belgium	2	86	775
14.	Belgium	2	87	997
15.	Belgium	2	88	1263
16.	Belgium	2	89	1200
17.	Belgium	2	90	1064
18.	Belgium	2	91	801
19.	Belgium	2	92	230
20.	Belgium	3	89	920

tab ma

tab ye

tabulate produces a one-way table of frequency counts. Here we tabulate market and find there are five categories of market: Belgium France Germany Italy and UK. There are 2,673 cars in Belgium, which accounts for 23.14%. By tabulating year, we can see the oldest year is 1970 and the most recent year is 1999 in our dataset.

. tab ma

market (=second dimension of panel)	Freq.	Percent	Cum.
Belgium	2,673	23.14	23.14
France	2,265	19.61	42.76
Germany	2,283	19.77	62.52
Italy	2,027	17.55	80.08
UK	2,301	19.92	100.00
Total	11,549	100.00	

. tab ye

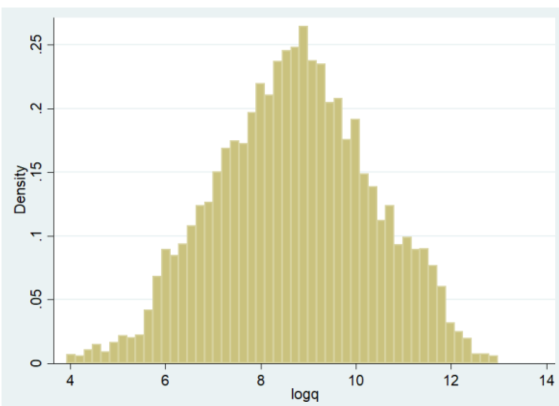
year (=first dimension of panel)	Freq.	Percent	Cum.
70	272	2.36	2.36
71	309	2.68	5.03
72	341	2.95	7.98
73	322	2.79	10.77
74	335	2.90	13.67
75	322	2.79	16.46
76	339	2.94	19.40
77	348	3.01	22.41
78	369	3.20	25.60
79	363	3.14	28.75
80	379	3.28	32.03
81	385	3.33	35.36
82	383	3.32	38.68
83	423	3.66	42.34
84	411	3.56	45.90
85	406	3.52	49.42
86	396	3.43	52.84
87	386	3.34	56.19
88	400	3.46	59.65
89	392	3.39	63.04
90	398	3.45	66.49
91	415	3.59	70.08
92	401	3.47	73.56
93	420	3.64	77.19
94	417	3.61	80.80
95	427	3.70	84.50
96	440	3.81	88.31
97	437	3.78	92.09
98	450	3.90	95.99
99	463	4.01	100.00
Total	11,549	100.00	

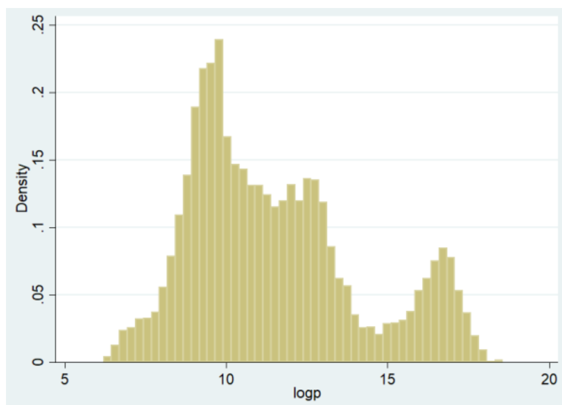
gen logq = ln(qu)
gen logp = ln(pr)
gen logpop = ln(pop)
gen loggdp = ln(ngdp)

Then we generate new variables: logq logp logpop loggdp, which are log forms of variable: number of sales, prices, population and nominal GDP.

hist logq, bin(50)
hist logp, bin(50)

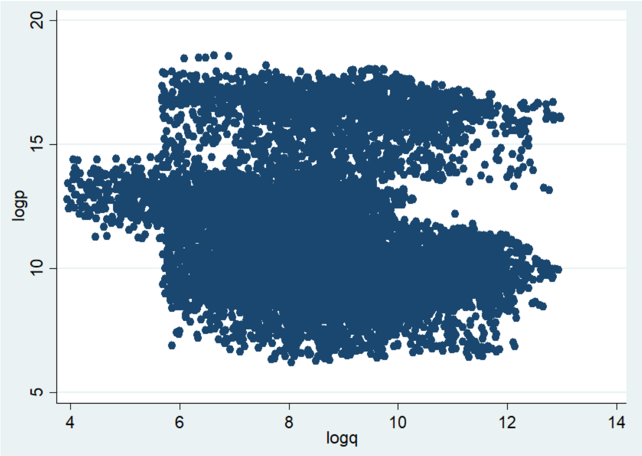
A histogram is a plot that lets you discover, and show, the underlying frequency distribution (shape) of a set of continuous data. This allows the inspection of the data for its underlying distribution (e.g., normal distribution), outliers, skewness, etc. "histogram" command assumes that the variable is continuous, so you need to type only histogram followed by the variable name. If you add up the area of the bars, you would get 1.





scatter logp logq

Scatter plots are important in statistics because they can show the extent of correlation, if any, between the values of observed quantities or phenomena (called variables). If no correlation exists between the variables, the points appear randomly scattered on the coordinate plane. If a large correlation exists, the points concentrate near a straight line.



3. Simple regressions: isoelastic demands (CES)

1) OLS

reg logq logp, robust

```
Linear regression                               Number of obs     =     11,549
                                                F(1, 11547)       =     217.88
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0185
                                                Root MSE          =     1.6118

------------------------------------------------------------------------------
             |               Robust
        logq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.0834329   .0056524   -14.76   0.000    -.0945125   -.0723533
       _cons |   9.669009   .0650023   148.75   0.000     9.541594    9.796425
------------------------------------------------------------------------------
```

Robust here means robust(unclustered) variance estimator. This simple OLS shows the negative relationship between price and quantity demanded.

Notes: The difference between three types of variance estimators: OLS, robust, and robust cluster is following:

a) OLS variance estimator: VOLS = s2 * (X'X)-1
   Where s2 = (1/(N - k)) ΣNi=1 ei2

b) Robust (unclustered) variance estimator: Vrob = (X'X)-1 * [ΣNi=1 (ei*xi)' * (ei*xi)] * (X'X)-1

c) Robust cluster variance estimator: Vcluster = (X'X)-1 * Σnj=1 uj'*uj * (X'X)-1
   Where uj = Σj_cluster ei*xi and n is the total number of clusters.

Above, ei is the residual for the ith observation and xi is a row vector of predictors including the constant. For simplicity, I omitted the multipliers (which are close to 1) from the formulas for Vrob and Vclusters. The formula for the clustered estimator is simply that of the robust (unclustered) estimator with the individual ei*xi's replaced by their sums over each cluster.

However, simple OLS does not consider simultaneity issue between price and unobservable attributes. Therefore we would like to control Fixed effects to see the difference:

2) OLS controlling model Fixed Effect

areg logq logp, robust a(co)

```
Linear regression, absorbing indicators         Number of obs     =     11,549
                                                F(   1,  11192)   =     336.30
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4565
                                                Adj R-squared     =     0.4392
                                                Root MSE          =     1.2183

------------------------------------------------------------------------------
             |               Robust
        logq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.0873493   .0047632   -18.34   0.000     -.096686   -.0780126
       _cons |   9.714052   .0544106   178.53   0.000     9.607397    9.820706
-------------+----------------------------------------------------------------
          co |   absorbed                                     (356 categories)
```

Typical use of areg: "areg depvar indvar1 indvar2, absorb(groupvar)"
areg and absorb() allows you to dummy for qualitative variables and obtain OLS regression results, except it does not create new variables or include coefficients for these dummies in regression results.

Notes: Regression results for our variable of interest and other quantitative covariates remain identical whether you: 1) manually generate dummies and include them in the regression (and use "reg") or 2) use the areg method

3) OLS controlling model & time Fixed Effect

areg logq logp i.ye, robust a(co)

Here we can see time dummies are manually generated and added into results table instead of being absorbed.

Notes: areg fits a linear regression absorbing one categorical factor. areg is designed for datasets with many groups, but not a number of groups that increases with the sample size. xtreg, fe command for an estimator that handles the case in which the number of groups increases with the sample size.

```
Linear regression, absorbing indicators          Number of obs     =     11,549
                                                 F(  30,  11163)   =      13.54
                                                 Prob > F          =     0.0000
                                                 R-squared         =     0.4600
                                                 Adj R-squared     =     0.4414
                                                 Root MSE          =     1.2159
```

logq	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
logp	-.0840172	.0048286	-17.40	0.000	-.093482	-.0745524
ye						
71	-.1727481	.1210149	-1.43	0.153	-.4099587	.0644625
72	-.1397776	.1186462	-1.18	0.239	-.3723452	.0927899
73	-.2341855	.1269346	-1.84	0.065	-.4829997	.0146288
74	-.4747776	.1232544	-3.85	0.000	-.7163781	-.2331771
75	-.4320747	.1250837	-3.45	0.001	-.6772608	-.1868886
76	-.429757	.1284696	-3.35	0.001	-.6815801	-.1779338
77	-.4272939	.1271746	-3.36	0.001	-.6765785	-.1780092
78	-.4016003	.1274453	-3.15	0.002	-.6514156	-.151785
79	-.3554762	.127734	-2.78	0.005	-.6058574	-.1050951
80	-.3345294	.1270051	-2.63	0.008	-.5834817	-.0855771
81	-.34567	.1251581	-2.76	0.006	-.5910019	-.1003381
82	-.3581478	.1246306	-2.87	0.004	-.6024457	-.1138498
83	-.4817167	.1248784	-3.86	0.000	-.7265004	-.2369331
84	-.520714	.1262474	-4.12	0.000	-.7681812	-.2732469
85	-.5705927	.1266984	-4.50	0.000	-.8189439	-.3222415
86	-.4603905	.1275012	-3.61	0.000	-.7103153	-.2104656
87	-.3783774	.1275767	-2.97	0.003	-.6284503	-.1283046
88	-.3902727	.1285992	-3.03	0.002	-.6423499	-.1381955
89	-.3415429	.1282957	-2.66	0.008	-.5930251	-.0900606
90	-.3950173	.1299559	-3.04	0.002	-.6497539	-.1402808
91	-.428876	.1285952	-3.34	0.001	-.6809452	-.1768068
92	-.3698482	.1281457	-2.89	0.004	-.6210363	-.1186601
93	-.5713861	.1285448	-4.45	0.000	-.8233567	-.3194155
94	-.5195882	.128253	-4.05	0.000	-.7709867	-.2681896
95	-.6096484	.1301979	-4.68	0.000	-.8648594	-.3544375
96	-.5894632	.1293554	-4.56	0.000	-.8430226	-.3359038
97	-.5540222	.1293488	-4.28	0.000	-.8075688	-.3004757
98	-.5514813	.1301695	-4.24	0.000	-.8066365	-.296326
99	-.6114863	.131037	-4.67	0.000	-.8683418	-.3546307
_cons	10.10385	.1172574	86.17	0.000	9.874001	10.33369
co	absorbed				(356 categories)	

4) OLS controlling model & time & market Fixed Effect

areg logq logp i.ye i.ma, robust a(co)

```
Linear regression, absorbing indicators         Number of obs     =     11,549
                                                F(  34,  11159)   =     130.95
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5698
                                                Adj R-squared     =     0.5548
                                                Root MSE          =     1.0854

------------------------------------------------------------------------------
             |               Robust
        logq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.3278852   .0536588    -6.11   0.000    -.4330659   -.2227044
             |
          ye |
         71  |  -.1519985   .1127467    -1.35   0.178     -.373002     .069005
         72  |  -.1224336   .1109064    -1.10   0.270    -.3398297    .0949624
         73  |  -.1782098   .1180508    -1.51   0.131    -.4096102    .0531906
         74  |  -.3651415   .1167065    -3.13   0.002    -.5939069   -.1363761
         75  |  -.2554398   .1180773    -2.16   0.031    -.4868922   -.0239875
         76  |  -.2197915   .1233863    -1.78   0.075    -.4616504    .0220674
         77  |  -.2046685   .1244622    -1.64   0.100    -.4486364    .0392993
         78  |  -.1751628   .1247158    -1.40   0.160    -.4196279    .0693023
         79  |  -.0653323   .1255881    -0.52   0.603    -.3115073    .1808426
         80  |  -.0201461   .1251806    -0.16   0.872    -.2655222    .2252301
         81  |  -.0151788   .1247302    -0.12   0.903     -.259672    .2293145
         82  |  -.0016677   .1269493    -0.01   0.990    -.2505107    .2471754
         83  |  -.1105638   .1290756    -0.86   0.392    -.3635747     .142447
         84  |  -.1227381   .1321203    -0.93   0.353    -.3817173    .1362411
         85  |  -.1616334   .1340867    -1.21   0.228     -.424467    .1012002
         86  |  -.0310021   .1368161    -0.23   0.821    -.2991858    .2371817
         87  |   .0574673   .1390321     0.41   0.679    -.2150601    .3299947
         88  |   .0566239   .1421265     0.40   0.690    -.2219691    .3352168
         89  |   .1104069   .1429447     0.77   0.440     -.16979     .3906038
         90  |   .0888126   .1462462     0.61   0.544    -.1978558     .375481
         91  |   .0464162    .146854     0.32   0.752    -.2414435     .334276
         92  |   .1128623   .1474214     0.77   0.444    -.1761096    .4018342
         93  |  -.0724912   .1497373    -0.48   0.628    -.3660028    .2210203
         94  |  -.0347442   .1506816    -0.23   0.818    -.3301068    .2606184
         95  |  -.1100498   .1527152    -0.72   0.471    -.4093985     .189299
         96  |   -.085951   .1527998    -0.56   0.574    -.3854655    .2135636
         97  |   -.050661     .15286    -0.33   0.740    -.3502937    .2489716
         98  |  -.0505303   .1547951    -0.33   0.744    -.3539561    .2528955
         99  |  -.0962571    .156096    -0.62   0.537    -.4022328    .2097186
             |
          ma |
     France  |   .5672602   .1018319     5.57   0.000     .3676518    .7668686
    Germany  |    .693624    .157776     4.40   0.000      .384355    1.002893
      Italy  |    2.34081   .1946571    12.03   0.000     1.959248    2.722372
         UK  |    -.06162      .2163    -0.28   0.776    -.4856063    .3623662
             |
       _cons |   11.90324   .6235748    19.09   0.000     10.68092    13.12555
-------------+----------------------------------------------------------------
          co |   absorbed                                     (356 categories)
```

Another command reghdfe performs exactly same as above:

reghdfe logq logp, vce(robust) a(co ma ye)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 3 HDFE groups                           F(   1,  11159) =      37.39
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.5691
                                                  Adj R-squared   =     0.5547
                                                  Within R-sq.    =     0.0044
                                                  Root MSE        =     1.0854

------------------------------------------------------------------------------
             |               Robust
        logq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.3278852    .053624    -6.11   0.000    -.4329976   -.2227727
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -  Redundant  |
-------------+---------------------------------------|
          co |         341           341              0    |
          ma |           4             5              1    |
          ye |          29            30              1 ?  |
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

reghdfe is a Stata package that runs linear and instrumental-variable regressions with many levels of fixed effects. Within Stata, it can be viewed as a generalization of areg/xtreg, with several additional features:

a) Supports two or more levels of fixed effects.
b) It can estimate not only ols regressions but two-stage least squares, instrumental-variable regressions, and linear gmm (via the ivreg2 and ivregress commands).

c) Careful estimation of degrees of freedom, taking into account nesting of fixed effects within clusters, as well as many possible sources of collinearity within the fixed effects.

d) Even with only one level of fixed effects, it is faster than areg/xtreg

From the result, we can see the endogeneity problem due to the correlation of price with time-invariant country heterogeneity seems much more important than the endogeneity problem due to the correlation of price with time-invariant model heterogeneity

5) OLS adding population, GDP into explanatory variables, controlling model & time & market Fixed Effect

reghdfe logq logp logpop loggdp, vce(robust) a(co ma ye)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 3 HDFE groups                           F(   3,  11157) =     134.08
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.5836
                                                  Adj R-squared   =     0.5695
                                                  Within R-sq.    =     0.0377
                                                  Root MSE        =     1.0672

------------------------------------------------------------------------------
             |               Robust
        logq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -1.977126    .103842   -19.04   0.000    -2.180675   -1.773578
      logpop |   .2430475   .2128435     1.14   0.254    -.1741634    .6602583
      loggdp |   1.965084   .1089951    18.03   0.000     1.751435    2.178734
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -   Redundant |
-------------+---------------------------------------|
          co |         341           341              0   |
          ma |           4             5              1   |
          ye |          29            30              1 ? |
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

By controlling GDP and population, the coefficient of logp starts to look like a reasonable elasticity of demand.

4. Construction of market shares

We use population as measure of market size, consider the demand is at the household level and assume an average family size of 4 members. Therefore, market size H = pop/4 . This will not be very important for the empirical results because eventually we are going to control for market*year fixed effects.

gen msize = pop/4

Then we generate market share, which equals to number of sales divided by market size:

gen share = qu/msize

egen sum_share = sum(share), by(ma ye)

One of Stata's most powerful and useful commands is egen. Like generate, it is used to create new variables, but it is much more than that. Using egen difficult and tedious variables can be created easily. Some examples are variables whose values are the mean of another variable for each group such as sociability for males and females. You can also use egen to create other variables that count the number of observations that fit a certain criteria, or even simply number observations. Here we are producing sum of market shares, separately for groups defined by one or more variables specified as arguments to by(), i.e, by market and year.

	ye	ma	sum_share
1	70	Belgium	.1079466
2	70	Belgium	.1079466
3	70	Belgium	.1079466
4	70	Belgium	.1079466
5	70	Belgium	.1079466
6	70	Belgium	.1079466
7	70	Belgium	.1079466
8	70	Belgium	.1079466
9	70	Belgium	.1079466
10	70	Belgium	.1079466
11	70	Belgium	.1079466
12	70	Belgium	.1079466
13	70	Belgium	.1079466
14	70	Belgium	.1079466
15	70	Belgium	.1079466

The outside good's market share in a given country, given year, is defined as follows:

gen share0 = 1 - sum_share

sum share share0

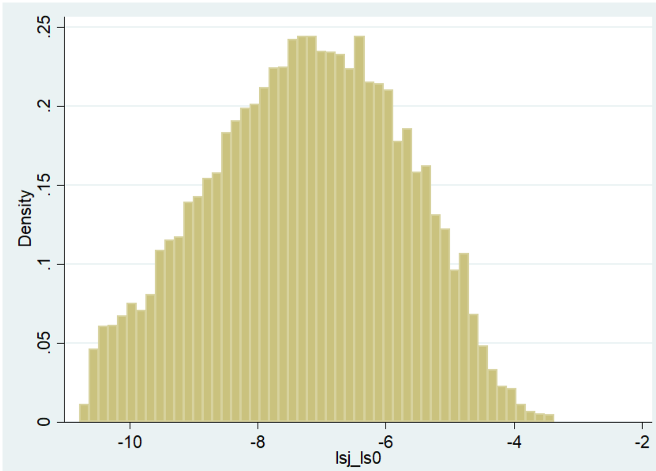
```
sum share share0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
share	11,549	.0016367	.0025921	.0000176	.0303018
share0	11,549	.8717519	.0234954	.8181894	.9356163

Let's generate log of odds-ratio and make the histogram:

gen lsj_ls0 = ln(share/share0)

hist lsj_ls0, bin(50)



5. Logit demand regressions (OLS and FE)

Now we are ready for logit demand estimation:

1) OLS

reg lsj_ls0 logp, robust

```
. reg lsj_ls0 logp, robust

Linear regression                               Number of obs     =     11,549
                                                F(1, 11547)       =       5.36
                                                Prob > F          =     0.0206
                                                R-squared         =     0.0005
                                                Root MSE          =     1.5014

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.0122702   .0052977    -2.32   0.021    -.0226545   -.0018859
       _cons |  -7.106636   .0621008  -114.44   0.000    -7.228364   -6.984908
------------------------------------------------------------------------------
```

Remember that we need to solve the price endogeneity problem by including FEs:

2) OLS with model FE

reghdfe lsj_ls0 logp, vce(robust) a(co)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 1 HDFE group                            F(   1,  11192) =       0.69
Statistics robust to heteroskedasticity           Prob > F        =     0.4065
                                                  R-squared       =     0.4613
                                                  Adj R-squared   =     0.4449
                                                  Within R-sq.    =     0.0001
                                                  Root MSE        =     1.1189

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.0038331   .0046181    -0.83   0.407    -.0128853    .0052192
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -   Redundant     |
-------------+---------------------------------------|
          co |         341          341              0        |
-----------------------------------------------------+
```

3) OLS with model & year FE

reghdfe lsj_ls0 logp, vce(robust) a(co ye)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 2 HDFE groups                           F(   1,  11163) =       0.34
Statistics robust to heteroskedasticity           Prob > F        =     0.5624
                                                  R-squared       =     0.4693
                                                  Adj R-squared   =     0.4517
                                                  Within R-sq.    =     0.0000
                                                  Root MSE        =     1.1120

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |   .0027145   .0046861     0.58   0.562    -.0064711    .0119002
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -   Redundant  |
-------------+---------------------------------------|
          co |       341             341              0    |
          ye |        29              30              1    |
-----------------------------------------------------+
```

4) OLS with model & year & market FE

reghdfe lsj_ls0 logp, vce(robust) a(co ye ma)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 3 HDFE groups                           F(   1,  11159) =      23.72
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.4932
                                                  Adj R-squared   =     0.4762
                                                  Within R-sq.    =     0.0028
                                                  Root MSE        =     1.0868

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -.2613373   .0536553    -4.87   0.000    -.3665112   -.1561635
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -   Redundant  |
-------------+---------------------------------------|
          co |       341             341              0    |
          ye |        29              30              1    |
          ma |         4               5              1 ?  |
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

5) OLS controlling population & GDP with model & year & market FE

reghdfe lsj_ls0 logp logpop loggdp, vce(robust) a(co ma ye)

```
HDFE Linear regression                            Number of obs   =     11,534
Absorbing 3 HDFE groups                           F(   3,  11157) =     128.56
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.5109
                                                  Adj R-squared   =     0.4945
                                                  Within R-sq.    =     0.0377
                                                  Root MSE        =     1.0677

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |   -2.00379   .1038401   -19.30   0.000    -2.207335   -1.800245
      logpop |  -.7852034   .2129166    -3.69   0.000    -1.202558   -.3678493
      loggdp |    2.00369   .1089785    18.39   0.000     1.790073    2.217307
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -  Redundant  |
-------------+---------------------------------------|
          co |         341           341           0     |
          ma |           4             5           1     |
          ye |          29            30           1 ?   |
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

6) OLS controlling population & GDP with brand & year & market FE

reghdfe lsj_ls0 logp logpop loggdp, vce(robust) a(ma ye brd)

```
. reghdfe lsj_ls0 logp logpop loggdp, vce(robust) a(ma ye brd)
(converged in 8 iterations)

HDFE Linear regression                            Number of obs   =     11,549
Absorbing 3 HDFE groups                           F(   3,  11473) =     625.41
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.3535
                                                  Adj R-squared   =     0.3493
                                                  Within R-sq.    =     0.1574
                                                  Root MSE        =     1.2114

------------------------------------------------------------------------------
             |               Robust
     lsj_ls0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logp |  -1.632788   .0378404   -43.15   0.000    -1.706962   -1.558615
      logpop |  -.6008506   .2377051    -2.53   0.011    -1.066793    -.134908
      loggdp |   1.652463   .0698674    23.65   0.000     1.515511    1.789415
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Num. Coefs.  =   Categories  -  Redundant  |
-------------+---------------------------------------|
          ma |           5             5           0     |
          ye |          29            30           1     |
         brd |          39            40           1 ?   |
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

7) OLS controlling population & GDP with brand & year & market FE and including model attributes:

reghdfe lsj_ls0 logp sp ac li wi cy hp we pl do le he logpop loggdp, vce(robust) a(ma ye brd)

```
. reghdfe lsj_ls0 logp sp ac li wi cy hp we pl do le he logpop loggdp, vce(robust) a(ma ye brd)
(converged in 8 iterations)

HDFE Linear regression                            Number of obs   =      9,227
Absorbing 3 HDFE groups                           F(  14,    9141) =     193.03
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.4208
                                                  Adj R-squared   =     0.4154
                                                  Within R-sq.    =     0.2427
                                                  Root MSE        =     1.1577
```

lsj_ls0	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
logp	-1.339531	.1342044	-9.98	0.000	-1.602602	-1.076461
sp	.0272126	.0025619	10.62	0.000	.0221908	.0322344
ac	.0148311	.0058055	2.55	0.011	.0034511	.0262111
li	-.0520444	.018889	-2.76	0.006	-.0890711	-.0150177
wi	.0596532	.0040658	14.67	0.000	.0516833	.0676231
cy	-.0005557	.000104	-5.34	0.000	-.0007595	-.0003519
hp	-.0335244	.0026927	-12.45	0.000	-.0388026	-.0282462
we	.0004381	.0002447	1.79	0.073	-.0000416	.0009178
pl	.3583704	.0508141	7.05	0.000	.2587634	.4579774
do	-.0478343	.0197557	-2.42	0.015	-.0865599	-.0091088
le	-.0026625	.0009728	-2.74	0.006	-.0045694	-.0007556
he	-.0107419	.0040727	-2.64	0.008	-.0187252	-.0027585
logpop	-.8305754	.250483	-3.32	0.001	-1.321578	-.3395727
loggdp	1.41743	.1406121	10.08	0.000	1.141798	1.693061

Absorbed degrees of freedom:

Absorbed FE	Num. Coefs.	=	Categories	-	Redundant
ma	5		5		0
ye	29		30		1
brd	38		39		1 ?

? = number of redundant parameters may be higher

The effect of some characteristics have the expected sign (sp ac li wi), ,i.e, maximum speed, acceleration time, fuel efficiency, width. The own price-elasticity of demand seems kind of reasonable.