1 Applied Econometrics: Demand Analysis

Demand analysis is one of the first topics come to in economics. Very important especially in the Keynesian paradigm.

Traditional consumer theory is based on the Neoclassical model of consumer choice.

Demand function

$$
\begin{aligned}
q_i &= q_i(p_1, p_2, ..., p_n, x) : i = 1, ..., n \\
x &= \sum_i p_i q_i
\end{aligned}
$$

Theory actual gives very little information on functional form, other relevant variables, form of variables. But it does imply restriction which can be useful in reducing the degrees of freedom taken up.

Expect homogeneous of degree zero in prices and total expenditure. This means that equal increases in prices and income should leave demand unchanged.

The Slutsky equation suggests own price substitution effects are negative.

$$\frac{\delta q_i}{\delta p_i} + q_i \left(\frac{\delta q_i}{\delta x} \right) < 0$$

This basic theory is used in a relatively 'ad hoc' way in applied work. The functional form is chosen for ease of exposition. There are a limited number of explanatory variables: own price, prices of substitutes and complements and possibly the general price level, with a time trend to capture changing tastes.

Popular specification is log linear, which has an added advantage that the coefficients are elasticities.

$$
\begin{aligned}
q_i &= A(p_i^{\beta_1}, p_j^{\beta_2}, \overline{p}^{\beta_3} x^{\beta_4} \exp(\beta_5 t) \exp(\varepsilon) \\
\log q_i &= \beta_0 + \beta_1 p_i + \beta_2 p_j + \beta_3 \overline{p} + \beta_4 \overline{x} + \beta_5 t + \varepsilon
\end{aligned}
$$

Can impose homogeneity prior to estimation, relative to $\overline{p}$. This implies $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 0$ or $\beta_3 = -\beta_1 - \beta_2 - \beta_4$

$$\log q_i = \beta_0 + \beta_1 \log \left(\frac{p_i}{\overline{p}} \right) + \beta_2 \log \left(\frac{p_j}{\overline{p}} \right) + \beta_4 \log \left(\frac{\overline{x}}{\overline{p}} \right) + ...$$

This restriction could be tested in the normal way by comparing the unrestricted estimates with:

$$\log q_i = \beta_0 + \beta_1 \log p_1 + \beta_2 \log p_j + (-\beta_1 - \beta_2 - \beta_4) \log \overline{p} + ...$$

In some case, where the researcher has needed to save dof this restriction is simply imposed. Using relative prices and real income can also have the advantage of reducing multicollineartiy, something that is often present in time series.

The negativity restriction is an inequality and difficult to impose.

In general studies of demand for single goods or policy orientated studies are often more concerned with estimation of elasticities rather tan testing the theory. They just impose the theory.

There are a number of problems, which are discussed in more detail in Thomas.:

1.0.1 Aggregation Problems

Consumer theory has considerable problems at a conceptual level -family versus individual, but at practical level there are further problems. The data is often presented in broad categories, large groups of individuals, in both time series and cross section.

Just having a theory of the individual doesn't mean it will hold at the aggregate level.

- The implications are that it should.
- Part of N-C/New Keynesian methodology of individual as focus of analysis
- Often cop out -use representative individual

Can aggregate over commodities as long as the groupings mean something, but with individuals require some restrictive assumptions -see Thomas

Problem is worse in the case of non-inear relations. For a linear model

$$\begin{aligned} y_i &= \alpha + \beta x_i \\ \overline{y} &= \alpha + \beta \overline{x} \end{aligned}$$

But if in logs then shouldnt use the arithmetic mean. For the logit

$$\begin{aligned} P_i &= \frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)} \\ \overline{P} &\neq \frac{\exp(\overline{x}'\beta)}{1+\exp(\overline{x}'\beta)} \end{aligned}$$

1.0.2 Identification Problem

Simultaneity is possible:

Early studies focused on agricultural products, where the data was available and identification was not a problem. Later manufacturing studies hit the classical identification problems. No reason why supply conditions should be more variable than demand

Simultaneity implies biased and inconsistent estimators of the demand equations.

1.0.3 Multicollinearity

As noted we might expect multicollinearity between expenditure and prices as both are trended. This will increase standard errors and reduce precision. It can be reduced by imposing homogeneity, but if their is insufficient variance in the explanatory variables, then any remaining multicollinearity, when homogeneity is imposed could make matters worse.

In the past used extraneous estimates of realexpenditure elasticity from cross section studies -assuming absense of price variation, but wide variation in real expenditures. If the extraneous estimates are unbiased then so are the estimates.

$$q_t = \alpha + \beta p_t + \gamma x_t$$

get estimate from cross section $\widehat{\gamma}$

$$q_t - \widehat{\gamma} x_t = \alpha + \beta p_t$$

but problem:

- standard error attached is not considered
- interpretation: cross section and time series are not the same thing, the former can be considered to represent the represents the long run effect and the latter the short run
- So care must be taken

1.1 Early Studies

Engel curves

$$p_i q_i = \alpha + \beta y_i$$

Cross section studies provided a test for Engel's law, that the income elasticity of demand for food was always less than one

- In cross section prices are pretty much fixed, expenditures vary
- But problem that other factors may be important which could give ommited variable bias
  - household characteristics: in particular houselhold size. Early investigators used equivalent adult scales
  - social effects. Can use dummies for social status grouping etc...
  - functional form. Major problem as goods can change from luxuries at low income to necessities at high income. Obvious functional form is sigmoid but is non linear and complex to estimate in practice. Could estimate income ranges separately: lower log linear, upper semi log, middle linear

1.2 Recent Developments

Duality: Use concept of duality to reformulate the consumer problem as choosing quantities so as to minimise the total expenditure necessary to achieve a given utility level U

$$\max(U) \textit{ subject to } x$$
$$\min(x) \textit{ subject to } U^*$$

$$\begin{aligned} \min x &= p_1 q_1 + p_2 q_2 \\ stU^* &= U^*(q_1, q_2) \end{aligned}$$

which gives the Hicksian compensated demand functions

$$\begin{aligned} q_1 &= f_1(p_1, p_2, U^*) \\ q_2 &= f_2(p_1, p_2, U^*) \end{aligned}$$

substituting back into x gives

$$x^* = x(p_1, p_2, U^*)$$

which is homogeneous of degree one in prices.

$$\frac{\delta x^*}{\delta p_i}$$

gives the demand functions

Substitute for U^* in

$$q_i = f_i(p_1, p_2, U^*)$$

$$U^* = U^*\left[q_1(p_1, p_2, x), q_2(p_1, p_2, x)\right]$$

Thus any fucntion which is homogeneous of degree unity can be used to generate a system of demand equations that satisfy the theoretical restrictions.

1.2.1 Demand systems

Rather than focus on individual commodities much recent work has been concerned with complete systems of demand equations.

Advantages are:

- reduce degrees of freedom problem
- can test restrictions to limit number of parameters ratehr than impose ad hoc

Restrictions come from the budget constraint (see Deaton and Muellbauer, Ch1):

$$p_1 q_1 + p_2 q_2 + ... + p_n q_n = x$$

and aggregation

$$\begin{aligned} p_1 \frac{\delta q_1}{\delta x} + ... + p_n \frac{\delta q_n}{\delta x} &= 1 \\ p_1 \frac{\delta q_1}{\delta p_j} + ... + p_n \frac{\delta q_n}{\delta p_j} &= -q_j \end{aligned}$$

Homogeneity implies that the sum of the price elasticities is equal to zero

$$\begin{aligned} \frac{p_1}{q_i} \frac{\delta q_i}{\delta p_1} + ... + \frac{p_n}{q_i} \frac{\delta q_i}{\delta p_n} + \frac{x}{q_i} \frac{\delta q_i}{\delta x} &= 0 \\ p_1 \frac{\delta q_i}{\delta p_1} + ... + p_n \frac{\delta q_i}{\delta p_n} - x \frac{\delta q_i}{\delta x} &= 0 \end{aligned}$$

Slutsky equation implies negativity

$$\frac{\delta q_i}{\delta p_j} + q_i \frac{\delta q_i}{\delta x} < 0$$

and symmetry implies

$$\frac{\delta q_i}{\delta p_j} + q_j \frac{\delta q_i}{\delta x} = \frac{\delta q_j}{\delta pi} + q_i \frac{\delta q_j}{\delta x}$$

These are not independent aggregation and symmetry imply homogeneity

cross equation restrictions mean symmetry can only be tested when using simultaneous equation methods MLE

n equations implies n^2 price parameters and n total expenditure parameters. With restrictions

$$\begin{aligned} & \left(n^2+n\right)-1/2(n+1)-1 \\ = \; & 1/2(n+1)-1 \end{aligned}$$

- May still be too few degrees of freedom.
- Impose further restrictions derived from making assumptions about the underlying utility function
- Most popular is additivity, which implies utility functions are additive, that the marginal utility of one good is independent of the quantity consumed of another.
- Needs goods defined in broad categories for validity. Cant be inferior goods or complements in Hicksian sense. If valid there is a massive reduction in parameters to n. But too restrictive?

  Two possibilities:

1. Specify form of utility function and then derive demand curves that satisfy the theoretical restrictions. Advantage is dof saved, but disadvantage is that cant test restrictions and there is a loss of generality.
2. Begin with demand system capable of satisfying restrictions, but that doesn't necessarily do so, and test if they hold. Advantage is can test, disadvantage is dof problem.

Consider forms of demand system

- Linear expenditure system -Stone: explicitly specified utility function
- Rotterdam model -test restrictions, popular until recently
- Indirect addilog and double log -from indirect utility function
- Direct and indirect translog -providing flexible functional form
- Almost Ideal demand system -from duality
- Variants

1.2.2 Linear Expenditure System

First used by Stone (1954) this system has an explicitly specified utility function:

$$\begin{aligned} U &= \beta_1 \log(q_1 - \gamma_1) + ... + \beta_n \log(q_n - \gamma_n) \\ \textit{subject to} \sum p_i q_i &= x \end{aligned}$$

$$p_i q_i = p_i \gamma_i + \beta_i \left[x - \sum_j p_j \gamma_j \right]$$

with $p_i \gamma_i$ representing subsistence expenditure and $x - \sum_j p_j \gamma_j$ supernumerary expenditure.

Advantages:

- expresses q_i as a linear function of real total expenditure x/p_i and of relative prices p_j/p_i

- is the only demand system that satisfies all the theoretical restrictions

But suffers from the fact that the underlying utility function is is additive and hence not general.

1.2.3 Rotterdam System

Until recently the most popular way to 'test' restrictions. Developed by Theil (1965) and Barton (1966) the dependent varaible is not the share but takes a more complex form:

$$\begin{aligned} w_i \left(\frac{dq_i}{q_i} \right) &= w_i d \log q_i \\ w_i &= \frac{p_i q_i}{x} \end{aligned}$$

The proportion of total expenditure allocated to good i.

We get demand equation as total differential of $q_i = f(...)$ and multiply by w_i/q_i.

$$w_i d \log q_i = \sum \Pi_{ij} d \log p_j + \mu_i \sum w_j dloq_j$$

where

$$\mu_i = p_i \left(\frac{\delta q_i}{\delta x} \right)$$

and Π_{ij} is the product of w_i and 'income compensated' elasticity of i wrt j

- Note equation is completely general -a differential demand equation
- Rotterdam model 'reparameterises' μ_i and Π_{ij} and treats them as constants, ignoring their dependence on income and prices. This is drastic, but necessary for estimation.

- To estimate use changes for total differential, meaning $\sum w_i \Delta \log q_i$ or

  $w_{it} (\log q_{it} - \log q_{it-1}) = w_{it} \Delta \log q_i$for a single commodity
- Allows theoretical restrictions to be written as equations which are unchanged for all values of total expenditure and prices.
- So aggregation implies $\mu_1 + ... + \mu_n = 1$ and $\Pi_{1j} + ... + \Pi_{nj} = 0$

Homogeneity implies $\Pi_{i1} + ... + \Pi_{in} = 0$
Negativity implies $\Pi_{ij} < 0$
Symmetry implies $\Pi_{ij} = \Pi_{ji}$

- Note that these are simple linear restrictions for all observations and are easy to impose.
- First order approximation to any arbitrary demand system

  Problems:
- Parameterisation of μ_i and Π_{ij}. McFadden shows true only if expenditure is a constant proportion of total expenditure no matter what the relative price structure is. Rather implausible.
- Claims as general system disputed as underlying utility fuction is additive
- Approximation to the log differential. Need ML to estimate
- Overtaken by Almost Ideal system

1.2.4 Other Models

Havent time to go through these, but check out Thomas

- Indirect addilog and double log -from indirect utility function
- Direct and indirect translog -providing flexible functional form

1.2.5 Almost Ideal Demand System

Deaton and Muellbauer: start with a general cost function

$$\log c(u,p) = (1-u)\log[a(p)] + u\log[b(p)]$$

where $a(p)$ is interpreted as the cost of subsistence and $b(p)$ the cost of bliss.

$$\log c(u,p) = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2}\sum_k \sum_j \gamma^*_{kj} \log p_k \log p_j + u\beta_0 \Pi p_k^{\beta_k}$$

Now

$$\frac{\delta \log c(u,p)}{\delta \log p_i} = \frac{p_i q_i}{c(u,p)} = w_i$$

so

$$w_i = \alpha_i + \sum \gamma_{ij} \log p_j + \beta_i u \beta_0 \Pi p_k^{\beta_k}$$

where

$$\gamma_{ij} = \frac{1}{2}\left(\gamma^*_{ij} + \gamma^*_{ij}\right)$$

Now

$$x = c(u,p) \Longrightarrow u = f(p,x)$$

by inversion of the indirect utility function. Do this for the cost function and substitute into the share equation gives

$$\begin{aligned} w_i &= \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x}{P}\right) \\ \log P &= \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2}\sum_j \sum_k \log p_k \log p_j \end{aligned}$$

Usually

$$\log P^* = \sum_j w_i \log p_j$$

is used as an approximation to estimate by OLS, but to test symmetry need to use the proper version, which requires systems estimation as it implies cross equation restrictions

So estimate

$$\begin{aligned} w_i &= \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x}{P^*}\right) + \varepsilon_i \\ \log P^* &= \sum_j w_i \log p_j \end{aligned}$$

which means a system of equations:

$$
\begin{aligned}
w_1 &= \alpha_1 + \gamma_{11}\log p_1 + \gamma_{12}\log p_2 + ... + \gamma_{1n}\log p_n + \beta_1 \log\left(\frac{x}{P^*}\right) + \varepsilon_1 \\
w_2 &= \alpha_2 + \gamma_{21}\log p_1 + \gamma_{22}\log p_2 + ... + \gamma_{2n}\log p_n + \beta_2 \log\left(\frac{x}{P^*}\right) + \varepsilon_2 \\
& .... \\
w_n &= \alpha_n + \gamma_{n1}\log p_1 + \gamma_{n2}\log p_2 + ... + \gamma_{nn}\log p_n + \beta n \log\left(\frac{x}{P^*}\right) + \varepsilon n
\end{aligned}
$$

This is a singular system as the dependent variables are shares and so add up to one across all the commodities. This means that adding up is automatically satisfied

$$
\begin{aligned}
\sum_i \alpha_i &= 1 \\
\sum_i \gamma_{ij} &= 0 \\
\sum_i \beta_i &= 0
\end{aligned}
$$

Homogeneity is testable and implies

$$\sum_j \gamma_{ij} = 0$$

This can be tested equation by equation using OLS, but symmetry

$$\gamma_{ij} = \gamma_{ji}$$

requires MLE

Consider a 4 commodity system

$$
\begin{aligned}
w_1 &= \alpha_1 + \gamma_{11}\log p_1 + \gamma_{12}\log p_2 + \gamma_{13}\log p_3 + \gamma_{14}\log p_4 + \beta_1(\log x - \log P^*) + \varepsilon_1 \\
w_2 &= \alpha_1 + \gamma_{21}\log p_1 + \gamma_{22}\log p_2 + \gamma_{23}\log p_3 + \gamma_{24}\log p_4 + \beta_2(\log x - \log P^*) + \varepsilon_2 \\
w_3 &= \alpha_1 + \gamma_{31}\log p_1 + \gamma_{32}\log p_2 + \gamma_{33}\log p_3 + \gamma_{34}\log p_3 + \beta_3(\log x - \log P^*) + \varepsilon_3 \\
w_4 &= \alpha_1 + \gamma_{41}\log p_1 + \gamma_{42}\log p_2 + \gamma_{43}\log p_3 + \gamma_{44}\log p_4 + \beta_4(\log x - \log P^*) + \varepsilon_4
\end{aligned}
$$

Homogeneity is tested by imposing the restrictions on the individual equations. Taking the second equation:

$$\gamma_{21} + \gamma_{22} + \gamma_{23} + \gamma_{24} = 0$$

which implies

$$\gamma_{21} + \gamma_{22} + \gamma_{23} = -\gamma_{24}$$

so

$$w_2 = \alpha_1 + \gamma_{21}(\log p_1 - \log p_4) + \gamma_{22}(\log p_2 - \log p_4) + \gamma_{23}(\log p_3 - \log p_4) + \beta_2(\log x - \log P^*) + \varepsilon_2$$

To test the restrictions estimate each equation by OLS restricted and unrestricted and the do LLR or F test.

As the dependent variable is shares the coefficients are not elasticities. We compute the elasticities as:

Expenditure

$$e_i = 1 + \frac{\beta_i}{w_i}$$

Compensated

$$e_{ij}^* = \frac{1}{w_i}\left(\gamma_{ij} + \beta_i \beta_j \log\left(\frac{x}{P^*}\right)\right) - \delta_{ij} + w_j$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise

Uncompensated

$$e_{ij} = e_{ij}^* - e_i w_j$$

So the uncompensated is the uncompensated minus the share weighted compensated.

Deaton and Muellbauer reject homogeneity for 4 commodity groups in the UK (food, clothing, housing, transport) and also not sharp decrease in DW statistic. Suggest rejection of homogeneity may be a problem:

- Omitted Variables: dynamics or conditioning variables that may be important
- Price expectations
- Aggregation problem
- Static model assumptions are inadequate
- Argue premature to reject consumer theory as consumption involves intertemporal choices, might need to consider labour supply, failure in models to take account of dynamic factors.
- Note have ommited durables: earlier studies didnt necessarily do so.

Developments:

- General dynamic model: Anderson and Blundell ( )set up a general first order dynamic model, that allows for short run dynemics and long run solutions. Basically a vector error correction model:

$$\Delta \mathbf{w}_t = A\Delta \mathbf{x}_t - B\left(\Delta \mathbf{w}_{t-1}^n - \Pi^n(\theta)\,\mathbf{x}_{t-1}\right) + \varepsilon_t$$

  where the bold lower case are vectors and the superscript n defines an operator that deletes the nth rpow of any vector or matrix. A and B are appropriately dimensioned short run coefficients. They use the share weighted price index to estimate the system.

- Ng develops this approach to consider the time series properties.

- Dynamise theory: Dunne Pashardes and Smith (1984) :

  - make an attempt to dynamise the theoretical model when applying the model to the allocation of government spending, where unitary State

    $$w_i = w_i(E, p, D)$$

  - Introduce a dynamic adjustment process.

    $$s_{it} = q_{it} + d_i s_{it-1}$$

    the effectiveness input is less than the actual expenditure because of a 'hangover' from previous expenditures. Treasury is well aware of this asn uses effective prices in its evaluation.

  - this gives a share equation

    $$w_i^* = \alpha_i + \sum_j \gamma_{ij} \log p_j^* + \beta_i \log\left(\frac{x}{P^*}\right) + \varepsilon_i$$

    where the star superscript on p represents the effective variable and $w_i^* = (p_{it}^* s_{it})/x_t$ and $x_t = \sum p_{it}^* s_{it}$

  - various conditioning and need variables are introduced and a demand system estimated successfully, with the homogeneity and symmetry restrictions accepted.

- Banks, Blundell and Lewbel (1997) "Quadratic Engel Curves and Consumer demand", Review of Economics and Statistics, 79 (4)

  - Propose: Quadratic Almost Ideal Demand System

  - Suggest a specifatcion of the Engel curves that reflects the observed behaviur better than preceeding specifications

1.2.6 Other Models

Havent had time to go through these, but check out Thomas

- Indirect addilog and double log -from indirect utility function
- Direct and indirect translog -providing flexible functional form