Advanced Industrial Organization I
Identification of Demand Functions

Måns Söderbom, University of Gothenburg

January 25, 2011

1 Introduction

This is primarily an empirical lecture in which I will discuss practical issues that you will encounter if you want to estimate a demand equation based on real data. Before we get going, please note the following:

- The core reading for the lecture consists **only** of the present notes including the empirical examples. That is, as far as the exam is concerned, if you know & understand everything in these notes you will be fine. Of course, the notes are based on previous research and existing teaching material, and I indicate the underlying references in case you want to consult these.

- I am assuming that you have a good understanding of basic microeconomics and econometric methods including OLS, panel data methods and 2-stage least squares. Please ask if something is not clear.

- My discussion of demand function estimation revolves around the **interplay** between demand and supply side mechanisms.

- Assuming that the market is in **equilibrium**, so that demand equals supply, I write down a simple system of equations modelling supply and demand. These equations, no doubt, will be very familiar to you.

- I then discuss important issues that arise when we want to **estimate** the model, e.g. what types of data will be needed, what is meant by identification and how can it be achieved, and what estimator should we choose. Throughout I assume we are dealing with a single homogeneous good.

2 Why do we care?

- The role played by consumer demand in IO cannot be over-emphasized.

- **Firms** care deeply about the demand for their products because it determines profits. Firms therefore have to understand demand very well.

- The **regulator**, whose task it is to monitor competition in the market, needs to understand the effect of price changes on demand in order to gauge the welfare loss due to market power (e.g. low price elasticity + monopolistic power = high welfare loss).

3 Consumer preferences and the demand curve

- I focus in this lecture primarily on how we can estimate, as reliably as possible, the demand curve in a certain market using econometric techniques. Before we get to that, though, I think it's useful to clarify in our minds where the demand curve comes from in the first place - what are its theoretical origins?

- Answer: the consumers' **preferences**, which we model by means of the **utility function**.

- Here's a very simple example (it's not real world at all but it delivers some nice intuitive results). Suppose you're looking at a market populated by z individuals (i.e. z measures the size of the market). Each individual

i in this market buys either 0 or 1 unit of some good produced by one or several firms (we have nothing to say about the producer side at this point). Consumer utility from consuming the good is defined simplistically as

$$U_i = \theta_i - P,$$

where θ_i is a parameter that varies across consumers reflecting the perceived quality of the produced good (alternatively, think of θ_i as measuring the 'joy' associated with consuming the good).

- Suppose that θ_i is **distributed uniformly** on the [0,1] interval. What does the demand curve look like?

- Individual i will buy the good if and only if $U_i > 0$ i.e. if

$$\theta_i > P.$$

Based on this insight, how many units will be bought in this market? The answer is pretty straightforward:

$$\begin{aligned}
Q^D &= \text{proportion of people for which } \theta_i > P \textit{ times} \text{ market size} \\
Q^D &= \Pr(\theta > P)\, z_i \\
Q^D &= [1 - \Pr(\theta \leq P)]\, z_i \\
Q^D &= \left[1 - \frac{P-0}{1-0}\right] z_i \\
Q^D &= [1 - P]\, z_i
\end{aligned}$$

(it follows from the assumption that θ is uniformly distributed that $\Pr(\theta \leq P) = \left(\frac{P-0}{1-0}\right)$ - check a basic statistics book if you are not convinced).

- Hence we've derived a **linear demand function**,

$$Q^D = [1 - P]\, z_i$$

directly from the assumptions we have made about consumer preferences and behaviour. It's telling us that quantity demanded depends positively on market size and negatively on price.

- Checkpoint:
  - What have we assumed about consumer behaviour above?
  - Suppose utility is less sensitive to price changes than in the model above - how would that affect the demand function?

4 Definitions

- We have just illustrated the basic underpinnings of the demand curve. From now on, we will take as given the derivation of an aggregate consumer demand curve. That is, we assume that quantity demanded is a decreasing function of the price, and don't worry too much about the details involved in actually deriving this demand curve from consumers' utility maximization problem. Next, consider some definitions.

- The two most common functional forms for demand are

  - **Linear** demand function, e.g.

    $$Q^D = a - b \times P,$$

    where $a > 0$ and $b > 0$ are demand parameters; Q^D denotes quantity demanded; and P is the price.

– Demand function with **constant price elasticity of demand**, e.g.

$$Q^D = X \times P^{-\eta},$$

where $-\eta < -1$ is the price elasticity of demand, and X is a demand shift parameter. Usually, you would express this in logarithmic form:

$$q^D = x - \eta \times p,$$

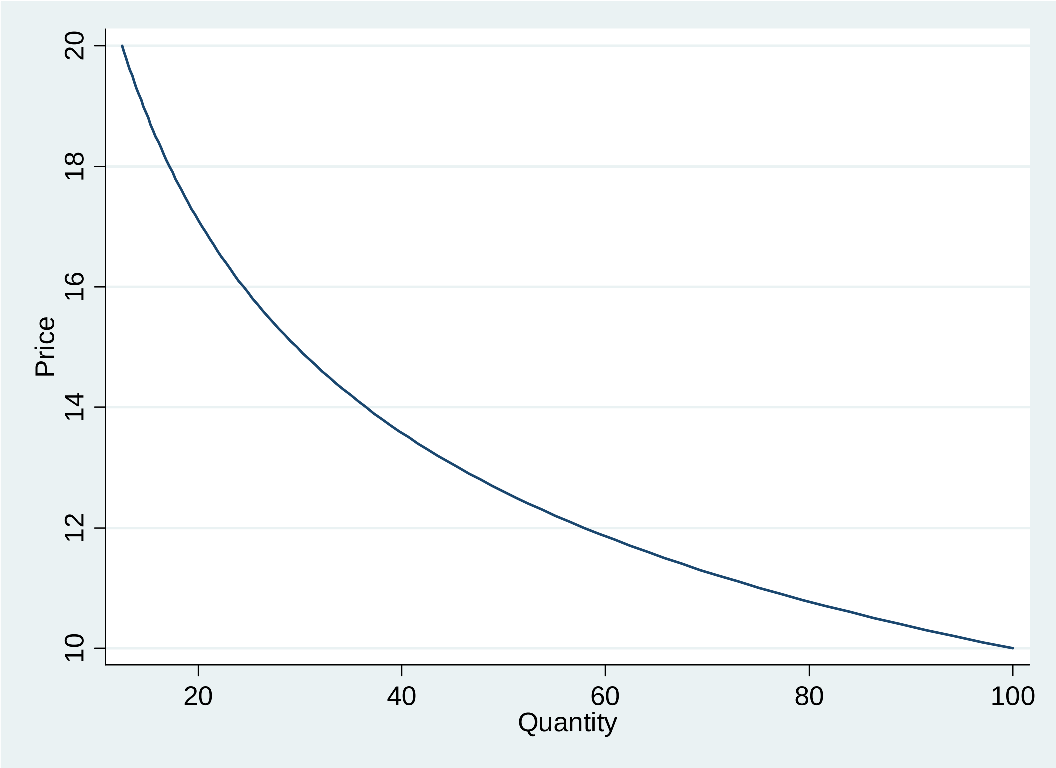
where $q^D = \ln Q^D$; $x = \ln X$; $p = \ln P$.

- These are illustrated in Figure 1 (notice that the price appears on the vertical axes in these graphs, i.e. these are inverse demand functions).

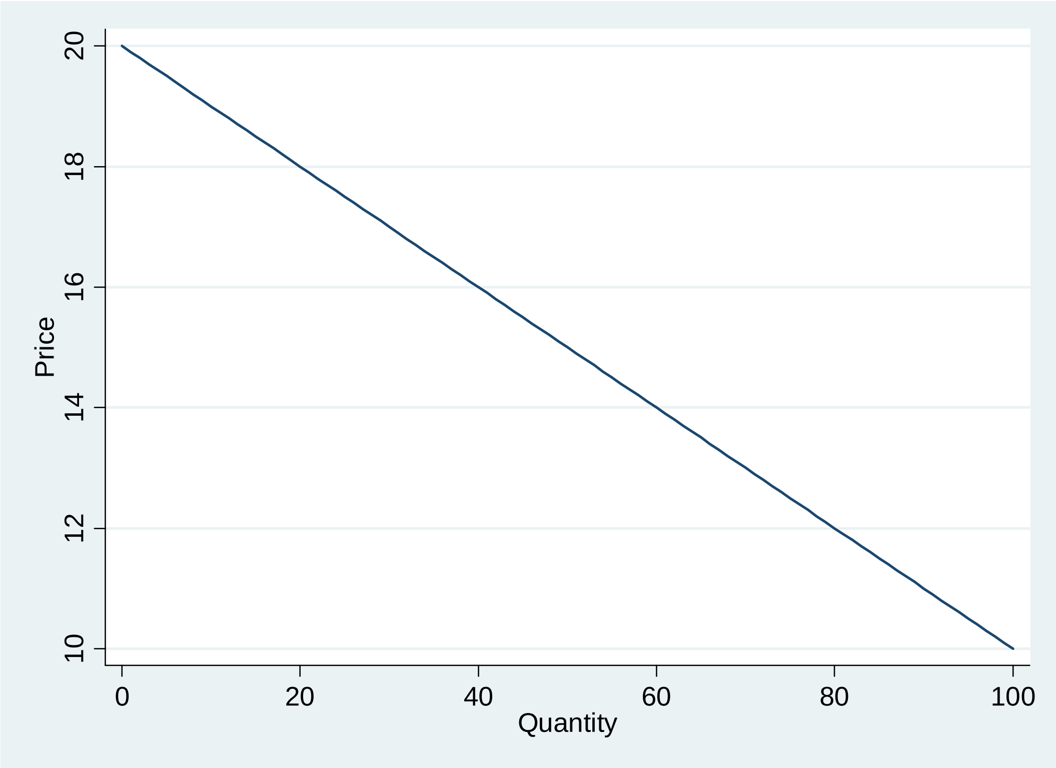
- [Figure 1 here]

Figure 1: Two common demand functions

A. Demand Function: Constant Price Elasticity of Demand
Quantity Demanded = 100,000*Price^{-3}



B. Demand Function: Linear
Quantity Demanded = 200 – 10*Price



- Of course we can easily re-arrange the demand curves above, so as to put the price on the left-hand side. This is known as the **inverse demand function**. For our two demand models:
  - Linear inverse demand function:

    $$P = A - B \times Q^D,$$

    where $A = a/b, B = 1/b$.(What's the interpretation of A?)

  - Inverse demand function with constant price elasticity of demand:

    $$P = X^{\frac{1}{\eta}} \times \left(Q^D\right)^{-\frac{1}{\eta}},$$

5 Estimating the Demand Function

- Suppose you face the task of estimating the price elasticity of demand, based on the following model:

$$q^D = x - \eta \times p$$

(recall: $q^D = \ln Q^D$; $x = \ln X$; $p = \ln P$). Your parameter of interest is η. You have data on prices and quantities consumed. How are you going to proceed?

- Suppose someone proposes you run an OLS regression of the following kind:

$$q^D = \beta_0 - \eta \times p + \text{control variables} + \varepsilon, \tag{1}$$

where ε is an error term, reflecting unobserved factors shifting the demand curve conditional on price and the other determinants of demand. Of course you can run the regression - but unfortunately most people will not be convinced your OLS estimate of η is unbiased. They will say that your analysis is basically not satisfactory.

- [Application: Demand for coffee in the Netherlands]

Advanced Industrial Organization I
University of Gothenburg
Måns Söderbom
mans.soderbom@economics.gu.se

Application: Estimating the Coffee Demand Function

Here I use a dataset on coffee consumption and production in the Dutch market to illustrate estimation of the demand function. The data contain monthly information about Dutch coffee market during the time period 1990-1996 (see Bettendorf and Verboven, 1998, for more details[1]). The following variables are included:

```
. use "C:\......\data\dutch_coffee_exe3.dta"

. describe

Contains data from C:\....\data\dutch_coffee_exe3.dta

obs:        84                    Written by R.
 vars:      14
 size:   7,728 (99.3% of memory free)
-------------------------------------------------------------------------------
              storage  display
variable name type     format     variable label
-------------------------------------------------------------------------------
maand         long     %9.0g      year and month of observation
year          double   %9.0g      year of observation
month         long     %9.0g      month of observation
qu            double   %9.0g      per capita consumption of roasted coffee in kg
cprice        double   %9.0g      price of roasted coffee per kg in current guilders
tprice        double   %9.0g      price of per kg tea in current guilders
oprice        double   %9.0g      price index for other goods
incom         double   %9.0g      Income per capita in current guilders
q1            long     %9.0g      season dummy 1
q2            long     %9.0g      season dummy 2
q3            long     %9.0g      season dummy 3
q4            long     %9.0g      season dummy 4
bprice        double   %9.0g      price of coffee beans per kg in current guilders
wprice        double   %9.0g      price of labor per man hours (work 160 hours per month)
-------------------------------------------------------------------------------
Sorted by:  year  month
```

Construction of variables in logs:

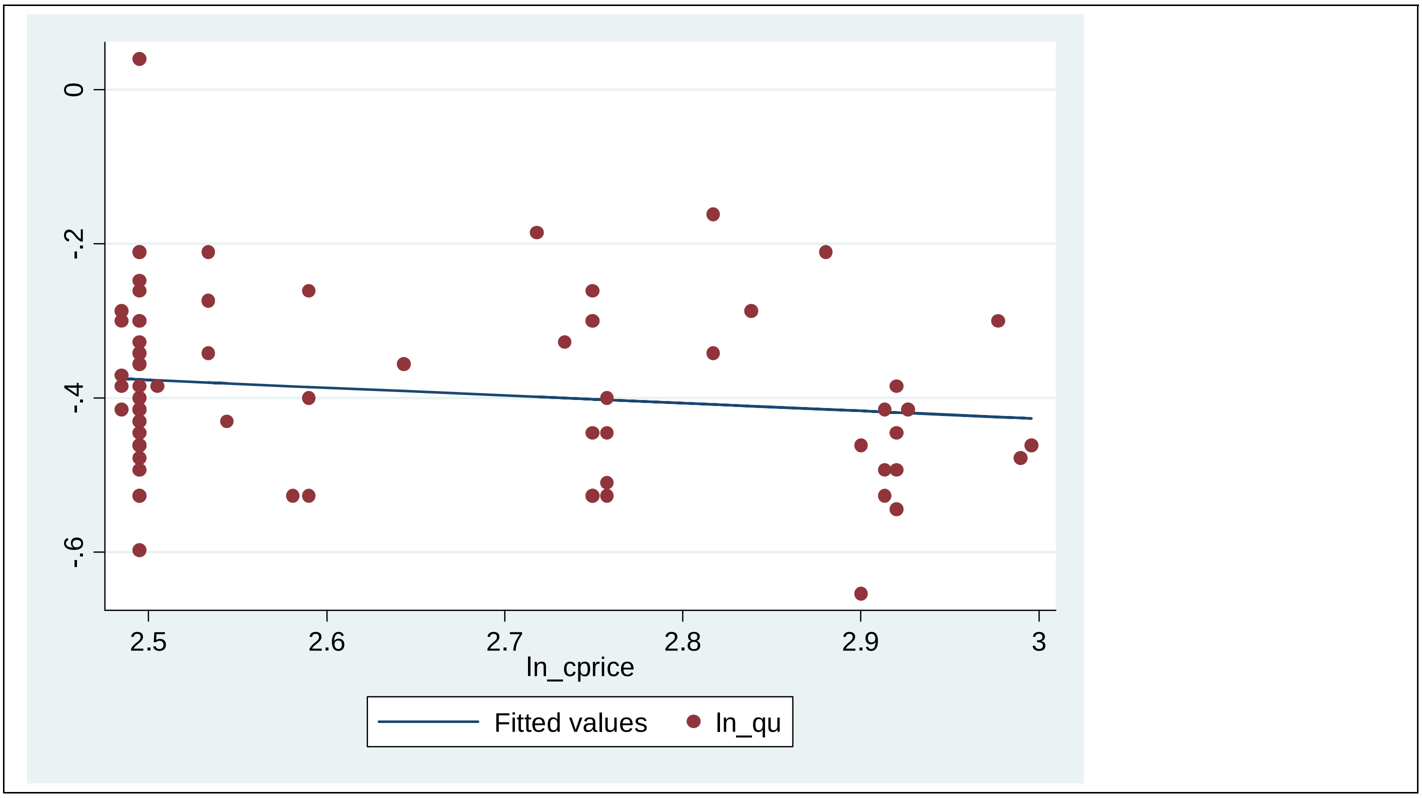
```
ge ln_qu=ln(qu)
ge ln_cprice=ln(cprice)
ge ln_tprice=ln(tprice)
ge ln_wprice=ln(wprice)
ge ln_bprice=ln(bprice)
ge ln_incom=ln(incom)
```

[1] Bettendorf and F. Verboven, “Competition on the Dutch Coffee Market. Research Memorandum, no.141. Central Planning Bureau”, 1998.

Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
maand	84	42.5	24.39262	1	84
year	84	1993	2.012012	1990	1996
month	84	6.5	3.472786	1	12
qu	84	.6815476	.0796301	.52	1.04
cprice	84	14.02976	2.555492	12	20
tprice	84	19.19155	.5254499	18.41	20.27
oprice	84	1.089048	.0574691	1	1.18
incom	84	1956.982	183.0193	1538.6	2297.2
q1	84	.25	.4356134	0	1
q2	84	.25	.4356134	0	1
q3	84	.25	.4356134	0	1
q4	84	.25	.4356134	0	1
bprice	84	4.030119	1.233319	2.44	7.18
wprice	84	31.79143	1.839375	28.15	34.21
ln_qu	84	-.3897157	.1116681	-.6539265	.0392207
ln_cprice	84	2.626139	.1707395	2.484907	2.995732
ln_tprice	84	2.954101	.0272999	2.912894	3.009142
ln_wprice	84	3.457502	.0589346	3.337547	3.532518
ln_bprice	84	1.352554	.2815938	.8919981	1.971299
ln_incom	84	7.574766	.0947155	7.338628	7.739446

i) The association between consumption and prices in the data



Note: The solid line shows the prediction based on a simple bivariate regression of log consumption on log prices:

```
. regress ln_qu ln_cprice

      Source |       SS       df       MS              Number of obs =      84
-------------+------------------------------           F(  1,    82) =    1.97
       Model |  .024251731     1  .024251731           Prob > F      =  0.1645
    Residual |   1.0107379    82  .012326072           R-squared     =  0.0234
-------------+------------------------------           Adj R-squared =  0.0115
       Total |  1.03498963    83  .012469755           Root MSE      =  .11102

------------------------------------------------------------------------------
       ln_qu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ln_cprice |  -.1001148   .0713739    -1.40   0.164    -.2421002    .0418706
       _cons |  -.1268003   .1878288    -0.68   0.502    -.5004516     .246851
------------------------------------------------------------------------------
```

- So, if you are using OLS to estimate the following equation:

$$q^D = \beta_0 - \eta \times p + \text{control variables} + \varepsilon,$$

the problem is that your price variable p is likely to be econometrically **endogenous** - i.e. correlated with the error term ε.

- Why is this? Let's think carefully about this. First of all, it should be obvious that in order to stand any chance of estimating a demand equation, there must be **variation** in the price data (as well as in the other variables, if you want to estimate their impact on demand).

- Why might prices vary? Being economists we believe in the idea that prices are the result of the interplay between supply and demand mechanisms.

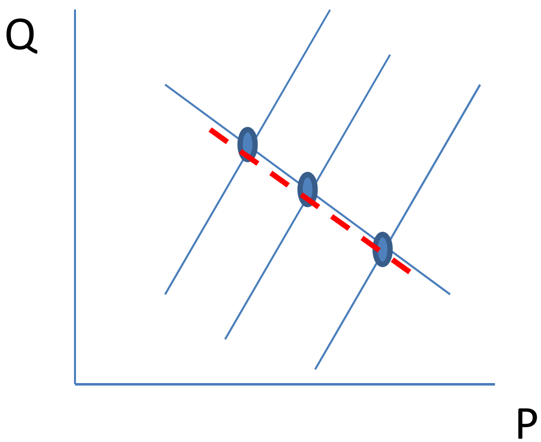
If we also assume that markets always clear, so that prices are equilibrium outcomes, it then becomes clear that in order to have any chance of estimating the **demand** function, we need **supply curve shifts**. This is important.

- In what follows I will assume (unless I say otherwise) that the outcomes we observe (prices, quantities etc.) are **equilibrium** outcomes. (Think of 'equilibrium' as a situation in which no consumer and no firm has an incentive to change their decision about how much to buy or produce - i.e. nothing changes, the market is at rest).

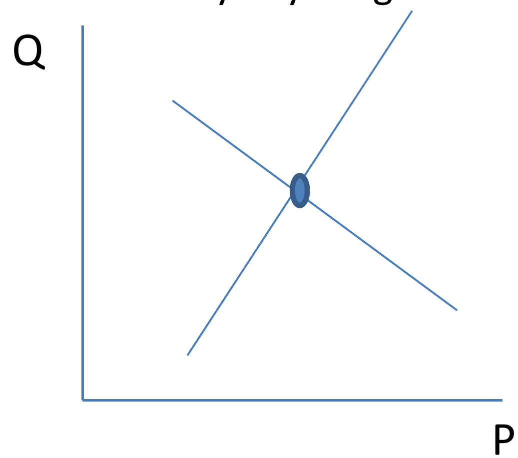
- [Illustration]

The first insight: You need supply shifts to identify the demand curve

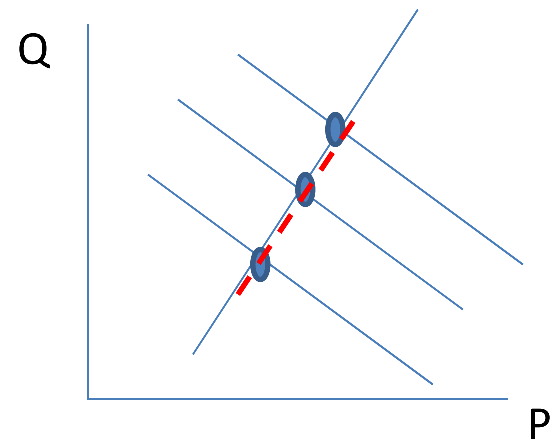
(a) You need supply shifts in order to trace the demand curve…



(b) otherwise you might not be able to identify anything…



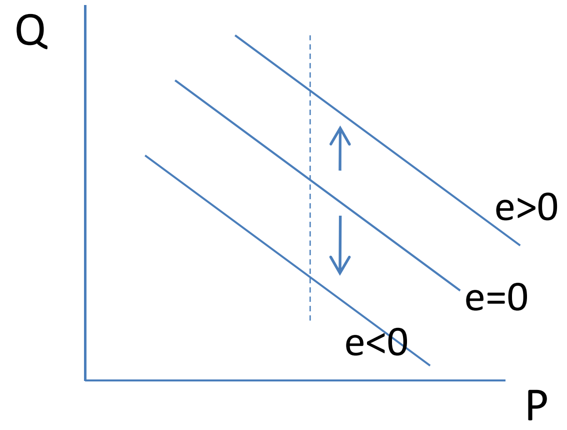
(c) alternatively you might end up identifying the supply curve(!)…



(the points indicated by represent observed data; the dashed red line illustrates what you would get if you run a regression of the type Q = a – b*P)

Now think about the residual in the demand function: $Q = a - b*P + e$

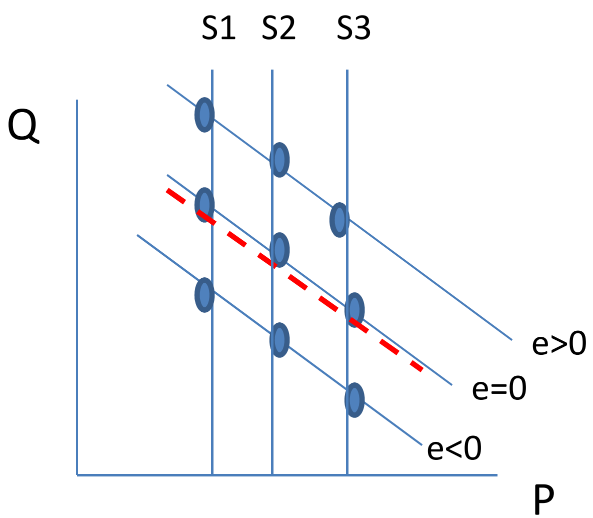
Different values of e are interpretable as shifting the demand curve up or down:



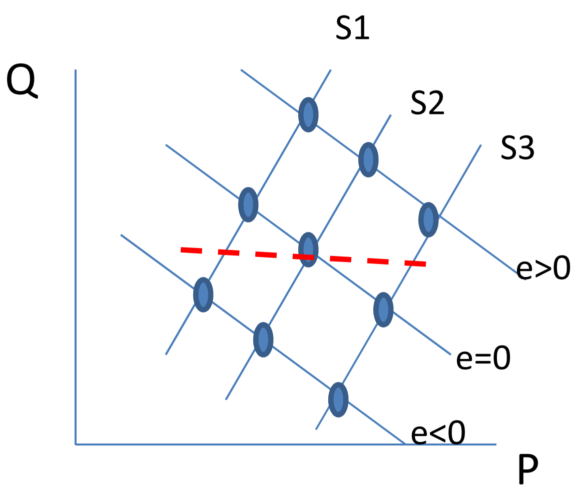
The second insight:

Depending on the slope of the supply curves, the price may respond endogenously to demand shocks

(a) No endogeneity problem if the supply curve is not price sensitive:



(b) But if demand shocks affect the price, you have an endogeneity problem:



Note: represent observed data;

The dashed red line illustrates what you would get if you run a regression of the type $Q = a - b*P + e$.

- As you know, one of the assumptions that needs to hold for the OLS estimator to be unbiased is that the error term (the residual) is **uncorrelated** with the explanatory variables. In panel (b) on the previous slide, high values of the residual ε (or e, same thing) are associated with high values of the price variable P and vice versa. Hence P and ε are positively correlated, which leads to **upward bias** in the OLS estimate of the price elasticity of demand η (yes?).

- Now let's formalize these thoughts a little. Consider a simple supply-demand model, in which the market price P and the quantity q are jointly determined by demand and supply, and discuss how we can use **data** to estimate parameters of interest. Homogeneous good.

- We continue to take **demand** as a given - i.e. we don't derive explicitly it from consumer preferences (but, of course, we understand that demand is determined by preferences and other factors such as income). We write demand in period t in constant elasticity form**:**

$$q_t = X_t P_t^{-\eta},$$

where $-\eta < -\mathbf{1}$ is the price elasticity of demand, and X_t is a demand shift parameter. Thus high values of X_t (could be income) will be associated with high demand, and vice versa; i.e. changes in X_t will **shift** the demand curve, and thus **influence** the equilibrium price.

- To motivate the **supply** curve, we turn to the firms' technology and costs. Let's write down the quantity (optimally) supplied by firms as a function of output prices and input prices (how might you derive such a supply function?). Hence:

- **Supply** in logarithmic form:

$$q_t^s = \alpha_0 + \alpha_1 p_t + \alpha_2 w_t + u_t \tag{2}$$

- **Demand** in logarithmic form:

$$q_t^d = \beta_0 + \beta_1 p_t + \beta_2 y_t + v_t, \tag{3}$$

  where I have added income to the right-hand side of the demand function (makes intuitive sense).

- Parameters: $\alpha_1 \geq 0$; $\alpha_2 < 0$; $\beta_1 \leq 0; \beta_2 > 0$.

- Equations (2) and (3) form a system of equations in **structural form,** in the sense that each equation specifies **causal, theoretical** relationships.

The parameter β_1 is interpretable as the price elasticity of demand, which is a key parameter in our theoretical model.

- Suppose our goal is to estimate the parameters of the model. What type of data do we need?
  - Quantity supplied & demanded
  - Output price
  - Demand shifters - e.g. income y
  - Supply shifters - e.g. input prices w
- Our empirical equations:

$$q_t^s = \alpha_0 + \alpha_1 p_t + \alpha_2 w_t + u_t \qquad \text{(Supply)}$$

$$q_t^d = \beta_0 + \beta_1 p_t + \beta_2 y_t + v_t \qquad \text{(Demand)}$$

- Equilibrium:

$$q_t^s = q_t^d$$

- Econometrics: Again, price is an endogenous variable. To see this, combine the supply and demand equations and solve for price and quantity in reduced form. You will obtain equations of the following form:

$$\begin{aligned} q_t &= \pi_1 w_t + \pi_2 y_t + \upsilon_1 (u_t, v_t) \\ p_t &= \omega_1 w_t + \omega_2 y_t + \upsilon_2 (u_t, v_t) . \end{aligned}$$

A shock to demand (v_t) may impact on the price - hence price is endogenous. [Exercise: Figure out how the reduced form parameters π_1, π_2, ω_1, ω_2 depend on the parameters in the supply and demand equations]

6 Identification by instrumental variables

- Remember our goal is to estimate the parameter β_1, which measures the causal effect of a change in the price on quantity demanded. That is, this parameter measures the slope of the demand function.

- In the language of simultaneous equation econometrics, we cannot identify β_1 unless the rank and order conditions are fulfilled (see an econometrics book if you are interested).

- More intuitively, we cannot infer β_1 from the observed relationship in the data between quantity and price, because we can't be sure about whether this relationship in the data reflects movement along the demand curve, the supply curve, or a combination. Please refer to the graphs above.

- Unless you insist that supply is completely price insensitive, you cannot use OLS to identify β_1.

- Now, remember you always need variation in prices in order to be able to identify β_1. One way of looking at our current problem is that the price variable P varies for **two** reasons:

  - i) because of shifts to the supply curve;

  - ii) because of unobserved shifts to the demand curve (variation in the residual in the demand equation).

- If we could somehow **"keep"** the first source of variation in prices and **"remove"** the other source, this would be great because then we are looking at the effects of shifts in supply on demand *along the same demand curve*.

- What is needed, conceptually, is a way of holding demand **constant** while varying supply.

- **Instrumental variables**: instrument the price variable using supply side variables that do not affect demand directly (only indirectly through the price variable).

- In the model above, this means we will use the **input prices** w as instruments for the price.

- Now write the equations in a form suitable for IV estimation:

$$\begin{aligned} p_t &= \omega_1 w_t + \omega_2 y_t + \upsilon_{2t} \quad \text{(stage 1)} \\ q_t^d &= \beta_0 + \beta_1 p_t + \beta_2 y_t + \upsilon_t \quad \text{(stage 2)} \end{aligned}$$

- Clearly $\alpha_2 \neq 0$, or otherwise we cannot identify β_1. Intuitively, the reason is that, while q depends on p, causation runs in the opposite direction as well. By using an IV approach, we consider how movements in the price that are only attributable to supply side shocks correlate with quantities produced and consumed. Our theory then tells us we can interpret the results as telling us what the demand curve looks like.

Correction: "Clearly alpha2 not equal zero..." should be replaced by "Clearly omega1 not equal zero..." (MS, Jan -11).

7 Detour: Causality in Applied Econometrics

- Goal of most empirical studies in economics: investigate if and how a change in an 'explanatory' variable X **causes** a change in another variable Y, the dependent variable - in our context, how a change in the price causes demand to fall.

- In order to find the causal effect, we must hold all other relevant determinants of Y **fixed** - ceteris paribus analysis. In the social sciences, we rarely have access to data generated in a laboratory (where the analyst controls the explanatory variables). We therefore need a technique that enables us to analyze the data and draw inferences about the role played by X **as if** other factors determining y are held fixed.

- Regression analysis is one such approach.

- We may achieve a lot by including control variables in our regressions. But when estimating demand-supply models, you typically suspect you don't observe all relevant determinants of demand and supply. As a result, the theory tells us we will have an endogeneity problem.

7.0.1 Instrumental Variables

Arguably the most important econometric problem for estimation of the demand-supply model is posed by the output price being likely endogenous. Suppose my goal is to estimate the price elasticity of demand. To do this, I consider the demand equation

$$q_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + v_t.$$

My problem is that the output price p_t is determined jointly with quantity demanded. In particular, a shock to demand not captured by income y (like what?) is likely to affect the price. Where in our demand equation above would such a shock enter?

It would enter the residual v_t. If, as a result, the residual v_t is correlated with the price, we clearly have an endogeneity problem.

Some of you may be very familiar with the instrumental variables approach, others may not. In this subsection, I briefly discuss the following:

- The key **assumptions** that need to hold for the IV approach to work

- **How** the IV estimator works

- Some intuition into **why** it works

My exposition is informal but hopefully sufficient given our current purposes. If you have difficulties, you need to consult a basic econometrics textbook (I'd recommend "Introductory Econometrics" by Jeffrey Wooldridge, or "A Guide to Modern Econometrics" by Marco Verbeek, but there are many others too).

7.0.2 Two key assumptions underlying the IV approach

- We suspect that the residual in our demand equation is correlated with the market price:

$$q_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + v_t.$$

$$cov\left(p_t, v_t\right) \neq 0.$$

This amounts to saying that the price is econometrically endogenous.

- Now, the price varies for many reasons. In our model, the price varies because of shocks to supply and shocks to demand:

$$p_t = \phi_1 \times w_t + \phi_2 \times y_t + e_t,$$

where ϕ_1, ϕ_2 are non-zero coefficients. Price will clearly be correlated with a determinant of supply in this case (i.e. the input price variable w_t). The key point, however, is that if w_t is uncorrelated with shocks to demand, then there is **some variation** in the price that is **not correlated with demand shocks**. That is, there is some **exogenous** variation in the price.

- The IV estimator uses **only** this source of variation in the price to identify the demand curve.

- Note the **analogy** with moving around the supply curve whilst holding demand constant.

- We say that w_t is our instrument (by which we really mean there is an exclusion restriction: w_t does not enter the structural demand equation - it is excluded).

- For the IV estimator to work, the following conditions need to hold:

$$\begin{aligned} cov\left(w_t, v_t\right) &= 0, \quad (4) \\ cov\left(w_t, p_t\right) &\neq 0. \quad (5) \end{aligned}$$

- The first of the conditions, (4), says that the instrument must be uncorrelated with the residual in the demand equation. This is sometimes referred to as **instrument validity**.

- The second condition, (5), says that the instrument must be correlated with the endogenous explanatory variable, i.e the price This is sometimes referred to as **instrument relevance**.

- If these conditions hold, then w_t can be used as an instrument for the price in the demand equation.

7.0.3 How the IV estimator works

- We can obtain an instrumental variable estimate by means of a **two-stage procedure:**

1. Run an OLS regression in which price is the dependent variable, and the instrument w_t, and other exogenous variables in the model are the explanatory variables:

$$p_t = \phi_1 \times w_t + \phi_2 \times y_t + e_t$$

   Once you've got your results, calculate the **predicted** values of the price based on the regression:

$$\hat{p}_t = \hat{\phi}_1 \times w_t + \hat{\phi}_2 \times y_t$$

You see how this "new" measure of the price will **not** be correlated with the demand residual - since the latter is assumed uncorrelated with w_t (and y_t)

2. In the demand equation, use the **predicted** values of the price (instead of the **actual** values) as the explanatory variable, and run the following regression using OLS:

$$q_t = \beta_0 + \beta_1 \hat{p}_t + \beta_2 y_t + v_t.$$

The resulting estimate of β_1 is the instrumental variable estimate, denoted b_1^{IV}.

If your sample is large and/or w_t is a very important explanatory variable for price for supply-related reasons, the IV estimate b_1^{IV} is likely to be much closer to the true value β_1 than the biased OLS estimate b_1^{OLS}. (To say what I have just said "properly" would require a lot of statistical jargon - consult an econometrics book if you are interested). This is the basic reasons for using the IV estimator in applied research.

7.0.4 Intuition

- I find it easiest to think of the IV estimator as a way of "purging" the price of endogeneity. That is, we **remove** from the price variable the part that correlates with the residual in the demand equation, but **keep** the part that is not correlated with residual in the demand equation. This is what the prediction after the first-stage regression achieves. Predicted price is then "exogenous" and there will therefore be no endogeneity bias.

[Discuss instrumental variables application:. Demand for Coffee in the Netherlands]

Application continued: Estimating the Coffee Demand Function

ii) Estimation by OLS, with controls for income and quarter dummies included:

```
. regress ln_qu ln_cprice q1 q2 q3 ln_incom

      Source |       SS       df       MS              Number of obs =      84
-------------+------------------------------           F(  5,    78) =    6.01
       Model |  .287909094     5  .057581819           Prob > F      =  0.0001
    Residual |  .747080533    78  .009577956           R-squared     =  0.2782
-------------+------------------------------           Adj R-squared =  0.2319
       Total |  1.03498963    83  .012469755           Root MSE      =  .09787

------------------------------------------------------------------------------
       ln_qu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ln_cprice |  -.2672689   .0903828    -2.96   0.004    -.4472072   -.0873305
          q1 |  -.1146401   .0310976    -3.69   0.000    -.1765508   -.0527295
          q2 |  -.0917195   .0303397    -3.02   0.003    -.1521213   -.0313177
          q3 |  -.1096042    .030499    -3.59   0.001    -.1703231   -.0488853
    ln_incom |   .3665619    .166024     2.21   0.030     .0360336    .6970903
       _cons |  -2.385461   1.104394    -2.16   0.034    -4.584141   -.1867807
------------------------------------------------------------------------------
```

- Compare these OLS results to those shown above (with no control variables added). How do the results differ and why?
- So adding control variables seems like a good thing to do. But we may still have an endogeneity problem, yes?

iii) Estimation by IV (two-stage least squares), with controls for income and quarter dummies included:

```
. ivregress 2sls ln_qu (ln_cprice=ln_bprice ln_wprice) q1 q2 q3 ln_incom, first

First-stage regressions
-----------------------

                                                Number of obs   =         84
                                                F(   6,      77) =     272.01
                                                Prob > F        =     0.0000
                                                R-squared       =     0.9549
                                                Adj R-squared   =     0.9514
                                                Root MSE        =     0.0376

------------------------------------------------------------------------------
   ln_cprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          q1 |  -.0099168   .0127581    -0.78   0.439    -.0353214    .0154879
          q2 |  -.0175034    .011685    -1.50   0.138    -.0407712    .0057643
          q3 |  -.0043489   .0122726    -0.35   0.724    -.0287867    .0200889
    ln_incom |   .0375377   .1464156     0.26   0.798    -.2540131    .3290884
   ln_bprice |   .4889238   .0178391    27.41   0.000     .4534015     .524446
   ln_wprice |    .738258   .2200838     3.35   0.001     .3000152    1.176501
       _cons |  -.8640816   .4795459    -1.80   0.075    -1.818979    .0908162
------------------------------------------------------------------------------


Instrumental variables (2SLS) regression               Number of obs =      84
                                                       Wald chi2(5)  =   32.71
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.2778
                                                       Root MSE      =  .09433

------------------------------------------------------------------------------
       ln_qu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ln_cprice |  -.2859159    .091475    -3.13   0.002    -.4652037   -.1066281
          q1 |  -.1141084   .0299852    -3.81   0.000    -.1728783   -.0553386
          q2 |  -.0920802   .0292491    -3.15   0.002    -.1494073   -.0347531
          q3 |   -.109086   .0294078    -3.71   0.000    -.1667243   -.0514477
    ln_incom |    .390966    .164139     2.38   0.017     .0692595    .7126725
       _cons |  -2.521518   1.083791    -2.33   0.020     -4.64571   -.3973267
------------------------------------------------------------------------------
Instrumented:  ln_cprice
Instruments:   q1 q2 q3 ln_incom ln_bprice ln_wprice
```

- Compare the OLS and the 2sls results – does endogeneity appear to be a big problem for the OLS estimator?
- What is the estimated price elasticity of demand?
- How should the quarter (season) dummies be interpreted?
- Does higher income raise coffee consumption?