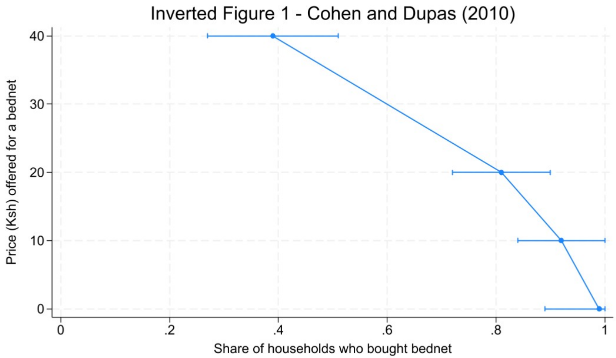
MIT 14.01: Principles of Microeconomics
Sp 2025, Lecture 2: Demand (Intro)

Dave Donaldson

Plan for Today

1. Randomized controlled trials and why correlation is not (necessarily) causation
2. How “good” is a product? Consumer surplus
3. How “sensitive” is demand? Elasticity of demand
4. What can move demand around? Comparative statics

Recall: Demand for Bednets (in rural Kenya)



- How was this obtained?
- Not an auction
- A lottery in a sense: groups of households (villages) were randomly chosen to be offered the chance to buy at different prices
- Also known as a *randomized controlled trial* (RCT)

Why do scientists love RCTs?

- Suppose I read this morning in the NYT: "Chocolate causes cancer, study finds"
- Details of study: researchers did a survey of randomly chosen American adults and found that, on average, those who eat more chocolate tend to have higher risk of developing cancer
- Does this evidence support the NYT claim?

Why do scientists love RCTs?

- Suppose I read this morning in the NYT: "Chocolate causes cancer, study finds"
- Details of study: researchers did a survey of randomly chosen American adults and found that, on average, those who eat more chocolate tend to have higher risk of developing cancer
- Does this evidence support the NYT claim?
- No! It tells us that there is a *correlation* (or "association") between chocolate and cancer, but not necessarily that chocolate has a *causal effect* (or "treatment effect") on cancer

Treatment Effects (in the Context of Demand)

- Recall our demand model from Lecture #1:
  - Given any price p_B, will consumer i choose to buy a bednet or not: $D_i(p_B)$
  - (B_i was binary. But that is not important to any of what follows today.)
- Suppose we examine two possibilities for the price: it can be either "high" ($p_B = H$) or "low" ($p_B = L$)

Treatment Effects (in the Context of Demand)

- Recall our demand model from Lecture #1:
  - Given any price p_B, will consumer i choose to buy a bednet or not: $D_i(p_B)$
  - (B_i was binary. But that is not important to any of what follows today.)
- Suppose we examine two possibilities for the price: it can be either "high" ($p_B = H$) or "low" ($p_B = L$)
- We can define the *treatment effect* (or "causal effect") for consumer i of changing from low to high: $D_i(H) - D_i(L)$
  - (This is borrowing notation from the canonical "Neyman/Rubin Causal – or Potential Outcomes – Model" in the field of statistics.)

Treatment Effects (in the Context of Demand)

- Recall our demand model from Lecture #1:
  - Given any price p_B, will consumer i choose to buy a bednet or not: $D_i(p_B)$
  - (B_i was binary. But that is not important to any of what follows today.)
- Suppose we examine two possibilities for the price: it can be either "high" ($p_B = H$) or "low" ($p_B = L$)
- We can define the *treatment effect* (or "causal effect") for consumer i of changing from low to high: $D_i(H) - D_i(L)$
  - (This is borrowing notation from the canonical "Neyman/Rubin Causal – or Potential Outcomes – Model" in the field of statistics.)
- Individual-level (and aggregate) demand functions, like all functions in math, are causal relationships
  - If $y = f(x)$, then we know how changing x will "cause" y to change, *ceteris paribus* (i.e. holding all else constant, which in this case means that $f(\cdot)$ itself is held constant)
  - That is, $D_i(H) - D_i(L)$ is not changing the consumer i, or anything about that consumer (i.e. $D_i(\cdot)$ is held fixed). All that is changing is the price offered.

Average Treatment Effects

- Now look at the aggregate demand change $D(H) - D(L) = \sum_i [D_i(H) - D_i(L)]$
- Define the population-level *average treatment effect* as

$$ATE \equiv \frac{1}{N_C} \sum_i D_i(H) - D_i(L) = \mathbb{E}[D_i(H) - D_i(L)]$$

  where I am using the notation that $\mathbb{E}[x_i] \equiv \frac{1}{N} \sum_i x_i$ for any variable "x"

- Then we have $D(H) - D(L) = N_C \cdot ATE$
  - That is, our quest to learn aggregate demand curves is the same as a quest to learn (up to N_C) the average treatment effect of changing the price that consumers are offered

Suppose We Executed the Following Observational Study

- We run a survey in rural Kenya and observe some consumers i who happen to pay $p_{Bi} = H$ and some who happen to pay $p_{Bi} = L$
- And we collect data on B_i among all (or just a random subset of) individuals i in these villages, and use that data to compute (estimates of) the conditional averages:
  - $\mathbb{E}[B_i | p_{Bi} = H]$: average bednet purchase rate among those who pay $p_{Bi} = H$
  - $\mathbb{E}[B_i | p_{Bi} = L]$: analog among those who pay $p_{Bi} = L$
- And we can calculate the difference: $\mathbb{E}[B_i | p_{Bi} = H] - \mathbb{E}[B_i | p_{Bi} = L]$
  - This is the average difference in B_i *across* the H and L people
  - (This object is closely related to $\mathbb{C}ov(B_i, p_{Bi})$, and hence also $\mathbb{C}orr(B_i, p_{Bi})$, amongst the individuals we surveyed.)

Why is Correlation Usually Not Causation?

- Our observational study would correctly use correlation to estimate (average) causation iff:

$$\underbrace{\mathbb{E}[B_i|p_{Bi} = H] - \mathbb{E}[B_i|p_{Bi} = L]}_{\text{The comparison we observe (i.e. correlation)}} \quad = \quad \underbrace{\mathbb{E}[D_i(H) - D_i(L)] \equiv ATE}_{\text{What we want to know (i.e. avg. causation)}}$$

- But this is usually not true!
- The problem:
  - LHS (correlation): average *across* H and L people
  - RHS (avg. causation): average "*within*-person" effect on each i—hypothetical exercise in which observe i's choice $D_i(\cdot)$ when (holding everything constant) we offer them the H price and then the L price

When *Does* Correlation = Causation?

- Can manipulate things to show that the following is always true:

$$\underbrace{\mathbb{E}[B_i|p_{Bi}=H]-\mathbb{E}[B_i|p_{Bi}=L]}_{\text{The comparison we observe (i.e. correlation)}}$$

$$=\mathbb{E}[B_i|p_{Bi}=H]\underbrace{-\mathbb{E}[D_i(L)|p_{Bi}=H]+\mathbb{E}[D_i(L)|p_{Bi}=H]}_{\text{subtract and add a term that }=0}-\mathbb{E}[B_i|p_{Bi}=L]$$

$$=\underbrace{\mathbb{E}[B_i-D_i(L)|p_{Bi}=H]}+\underbrace{\mathbb{E}[D_i(L)|p_{Bi}=H]-\mathbb{E}[B_i|p_{Bi}=L]}$$

$$=\underbrace{\mathbb{E}[D_i(H)-D_i(L)|p_{Bi}=H]}_{\text{Avg. causal effect on those with }p_{Bi}=H}+\underbrace{\mathbb{E}[D_i(L)|p_{Bi}=H]-\mathbb{E}[D_i(L)|p_{Bi}=L]}_{\text{“Selection (into treatment) effect”}}$$

- NB: the last step uses the fact that (e.g.) $\mathbb{E}[B_i|p_{Bi}=H]=\mathbb{E}[D_i(H)|p_{Bi}=H]$
  - That is, among those who pay $p_{Bi}=H$, their B_i choice is (by definition) $D_i(H)$
  - And an analogous argument applies to L
- So we have learned that correlation = causation + the selection effect

How Would Randomization of Treatment Assignment Help?

- First, it makes the "selection effect" go away:
  - That is, if p_{Bi} is random then $\mathbb{E}[D_i(L)|p_{Bi} = H] - \mathbb{E}[D_i(L)|p_{Bi} = L] = 0$
  - This is because if some "x" is random then conditioning on it doesn't matter (i.e. always have $\mathbb{E}[w|x] = \mathbb{E}[w]$ for any variable w)
  - Intuitively, if treatment p_{Bi} is randomly given out then people can't select their own "treatments"; so no selection effect.

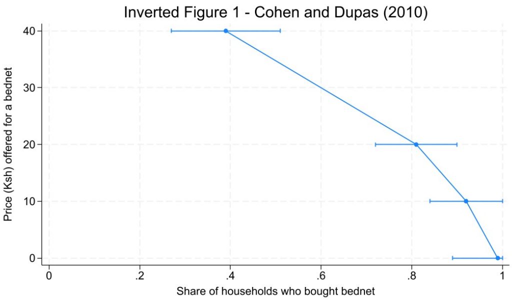
How Would Randomization of Treatment Assignment Help?

- First, it makes the "selection effect" go away:
  - That is, if p_{Bi} is random then $\mathbb{E}[D_i(L)|p_{Bi} = H] - \mathbb{E}[D_i(L)|p_{Bi} = L] = 0$
  - This is because if some "x" is random then conditioning on it doesn't matter (i.e. always have $\mathbb{E}[w|x] = \mathbb{E}[w]$ for any variable w)
  - Intuitively, if treatment p_{Bi} is randomly given out then people can't select their own "treatments"; so no selection effect.
- Second, it makes the "average causal effect on those with $p_{Bi} = H$" = the ATE
  - That is, $\mathbb{E}[D_i(H) - D_i(L)|p_{Bi} = H] = \mathbb{E}[D_i(H) - D_i(L)] \equiv ATE$
  - This is again because if some variable (like p_{Bi}) is random then the average obtained when conditioning on it is the same as the average obtained when not conditioning on it (which is the overall average, ie the ATE)
- Put these together and we have that, when p_{Bi} is random:

$$\underbrace{\mathbb{E}[B_i|p_{Bi} = H] - \mathbb{E}[B_i|p_{Bi} = L]}_{\text{The comparison we observe (i.e. correlation)}} \quad = ATE$$

- Bottom line: need treatments to be assigned randomly (or in a way that we believe is as good as randomly assigned) to learn about the causal effect of treatments.

Demand for Bednets (in rural Kenya) Again

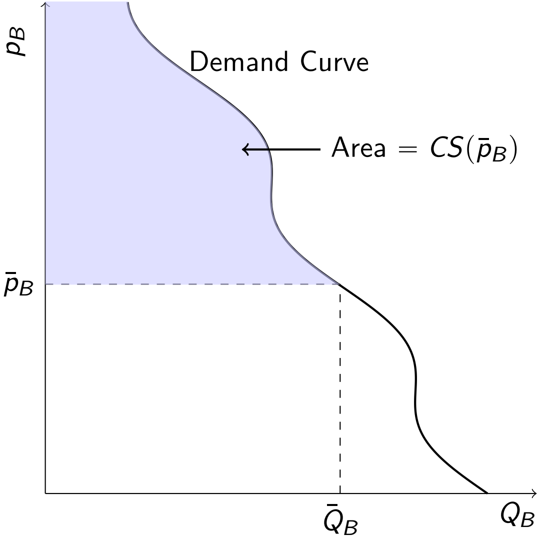


- Authors randomized p_{Bi} across i. So selection effect should be zero (in expectation at least)
- E.g. difference between dot at $p_B = 0$ and dot at $p_B = 40$ is an unbiased estimate of $\mathbb{E}[D_i(40) - D_i(0)]$
- Hence if we multiply the x-axis by N_C we have an estimated demand curve (well, 4 points on it)
- (You can ignore the bars, but they are 95% confidence intervals on these estimates.)

How “good” are bednets?

- What can we use demand curves for?
- One use: How much “value” is created by some product?
  - What do we mean by “value”?
  - And surely it depends on not just whether the product exists, but whether the product is actually affordable, right?
- Recall from Lecture #1 that:
  - Each consumer’s demand was $D_i(p_B) = 1$ if $v_i \geqslant p_B$ (and $= 0$ otherwise)
  - We therefore called v_i the WTP_i for each i: their (max) willingness to pay for a bednet
- Since v_i is what they are willing to pay and p_B is what they actually pay, we define $CS_i(p_B) \equiv \max\{0, v_i - p_B\}$ as the *consumer surplus* for consumer i when the price is p_B
  - When you buy a cold beverage on a hot day, you’re probably getting a lot of consumer surplus
  - You lose p_B, but gain v_i, and the difference – the surplus – is your net gain from the exchange

Aggregate consumer surplus



- Define aggregate CS (at some $\bar{p}_B$) as:

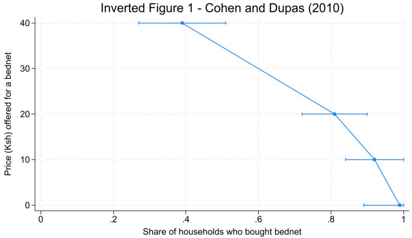
$$CS(\bar{p}_B) \equiv \sum_i CS_i(\bar{p}_B) = \int_{\bar{p}_B}^{\infty} D(p)dp$$

- But if you like to think of integrals as "area below the curve" then use the inverse demand function instead (and you'll get the same answer):

$$CS(\bar{Q}_B) \equiv \int_0^{\bar{Q}_B} [D^{-1}(Q_B) - \bar{p}_B]dQ_B$$

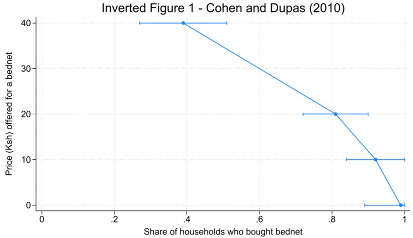
- Note how CS for a good doesn't have much to do with how "big" that good is (i.e. $\bar{p}_B$ or $\bar{Q}_B$ or $\bar{p}_B\bar{Q}_B$)

How Much Value Would be Created by Free Bednets (in rural Kenya)?



- To simplify, let's assume that $D(\cdot)$ is linear
- Then $CS(\bar{p}_B = 0)$ is equal to N_C times the area of the triangle from the $p_B = 0$ point to the $Q_B = 0$ point
- And let's use the slope from $p_B = 0$ to $p_B = 40$ in order to estimate the slope of $D(\cdot)$ (and to extrapolate linearly above $p_B = 40$)
- Then triangle has area of around $\frac{1}{2} \times 1 \times 67 \approx 34$. So total $CS(\bar{p}_B) = 34N_C$

How Much Value Would be Created by Free Bednets (in rural Kenya)?



- To simplify, let's assume that $D(\cdot)$ is linear
- Then $CS(\bar{p}_B = 0)$ is equal to N_C times the area of the triangle from the $p_B = 0$ point to the $Q_B = 0$ point
- And let's use the slope from $p_B = 0$ to $p_B = 40$ in order to estimate the slope of $D(\cdot)$ (and to extrapolate linearly above $p_B = 40$)
- Then triangle has area of around $\frac{1}{2} \times 1 \times 67 \approx 34$. So total $CS(\bar{p}_B) = 34N_C$
- Does that answer surprise you?

Elasticity of Demand

- *Elasticity of demand*: a summary of demand sensitivity to price
- Given $D(p)$ we could of course evaluate its derivative at any point: $\frac{dD(p)}{dp}$
- But economists prefer to use the related concept of *elasticity* rather than derivative:

$$\text{“Elasticity of demand at } p\text{”} \equiv \varepsilon_{D,p}(p) = \frac{p}{D}\frac{dD(p)}{dp} = \frac{d\ln D(p)}{d\ln p} \approx \frac{\text{\% change in D}}{\text{\% change in p}}$$

- Economists prefer elasticity to slope because elasticities are unit-free
  - And we are often talking about things whose units will differ a lot across applications...e.g. here we have $D(\cdot)$ in units of “number of bednets” and p in units of “Kenyan shillings”)

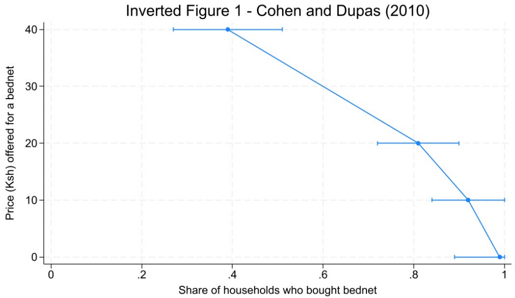
Elasticity of Demand: Comments

- Will often use synonymous language when referring to $\varepsilon_{D,p}(p)$:
  - "Elasticity of demand"
  - "Demand elasticity"
  - "*Price* elasticity of demand" – because later we will also talk about the *income* elasticity of demand
  - "*Own* price elasticity of demand" – because later we will talk about how the demand for some good "B" depends on the price of some good "A"

- Note that $\varepsilon_{D,p}(p)$ is a function of p, so it can vary depending on where we evaluate it
  - So, be careful and say "elasticity of demand at p"
  - Though sometimes we might work with a demand function for which the elasticity *is* constant at all p

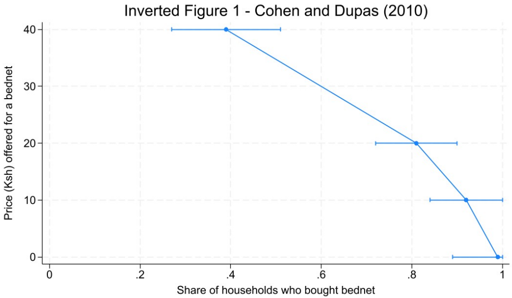
- Most demand curves in this class (and Econ in general) slope down, so most demand elasticities are negative. We therefore sometimes neglect the sign and say "the demand elasticity is 2" (because it's obvious that it's really -2).

What is the Elasticity of Demand for Bednets (in Rural Kenya)?



- Let's try it "at" $p_B = 30$...
- Fine to approximate the slope there by the estimated slope from $p_B = 40$ to $p_B = 20$

What is the Elasticity of Demand for Bednets (in Rural Kenya)?



- Let's try it "at" $p_B = 30$...
- Fine to approximate the slope there by the estimated slope from $p_B = 40$ to $p_B = 20$
- $\varepsilon_{D,p}(30) \approx \frac{30}{D(30)} \frac{D(40)-D(20)}{40-20} = \frac{30}{N_C \times 0.6} \frac{N_C \times (0.4-0.8)}{20} = -1$
- Note how the N_C canceled out – another consequence of elasticity being unit-free.

Elastic vs Inelastic

- We will use the definitions:
  - $\varepsilon_{D,p}(p) < -1$: demand is *elastic* at p
  - $\varepsilon_{D,p}(p) \geqslant -1$: demand is *inelastic* at p
  - $\varepsilon_{D,p}(p) = -\infty$: demand is *perfectly elastic* at p
  - $\varepsilon_{D,p}(p) = 0$: demand is *perfectly inelastic* at p
- Why is the cutoff of -1 important?
- One answer: imagine you are a firm facing the demand curve $D(p)$ for your product. Will raising your price a bit (starting from some p) increase your revenue (i.e. $R(p) \equiv p \cdot D(p)$)? That is, is your *marginal revenue* ($MR(p) \equiv \frac{dR(p)}{dp}$) positive or negative at p? The answer depends on whether demand is elastic at p or not because:

$$MR(p) \equiv \frac{dR(p)}{dp} = D(p) + p\frac{dD(p)}{dp} = D(p)\left[1 + \varepsilon_{D,p}(p)\right]$$

Comparative Statics

- Aka (in this context) "curve-shifting"
- For example: what would happen to the demand function $D(p)$ if...
  - We doubled the number of households N_C?
  - One new consumer arrived who had a really high v_i? (Or a really low v_i?)
  - A marketing campaign told consumers about the benefits of using bednets?
  - That same marketing campaign was effective, but only targeted at those consumers who already had high v_i?
  - The Kenyan government gave 500 Ksh to everyone in the village?
- What about this one: what would happen to our demand function $D(p)$ if some event caused p to rise?

Concluding Remarks

- **Key concepts from today's lecture:**
  - Correlation is not (necessarily) causation
  - Because of selection bias, need randomization (or something like it) to estimate causal/treatment effects (even just the average treatment/causal effect)
  - Aggregate consumer surplus: area under demand function
  - Elasticity of demand: proportional derivative of demand function
  - Elastic vs. inelastic demand
  - Comparative statics: what would move the demand function?
  - Events that cause p to move don't cause the demand *function* $D(\cdot)$ to move

- **Next lecture:**
  - Supply!
  - ...who makes these goods the consumers demand?
  - ... and why would they be willing to make them for a given price p?