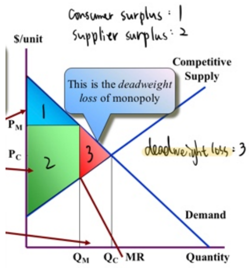
Econ Cheat Sheet

Lecture 2

- Max profit of any type: **MR(P if price taker) = MC**
- Demand: P = A – BQ; **MR = A- 2BQ**
- Aggregated MC has smaller slope
- **Deadweight loss** of monopoly: when monopolist set MR=MC at their own



Lecture 3- EndoIV

- **Endogeneity**: $Y = \alpha + \beta X + \epsilon$; (1) X is correlated with ϵ; (2) X is **endogenous explanatory variable**
- Common types of endogeneity: (1) omitted variables; (2) measurement error; (3) self-selection; (4) simultaneity e.g. demand and supply
- **Instrumental Variables**: $Y = \alpha + \beta X + \epsilon$, where $Cov(X, \epsilon) = 0$; (1) if there exists z uncorrelated with ϵ but correlated with X
- **Structural equation**: $Y = \alpha + \beta X + \epsilon$
- **Reduced form**: $X = \gamma + \delta z + v$; (write an endo variable in terms of exo variables)
- Weak instruments: $H_0 = the\ IV\ is\ weak\ (reject)$
- Hausman: $H_0 = y\ is\ not\ endogenous\ (reject)$
- Sargan: $H_0 = all\ IVs\ are\ exo\ (do\ not\ reject)$
- **2SLS**: (1) identify structural equation and reduced form; (2) regress the reduced form; (3) insert the result of 3 to structural equation; (4) regress the structural equation

Lecture 5- Price Discrimination and Monopoly

- Feasibility of price discrimination: (1) **identification**: able to identify demands of diff types of consumer or markets; (2) **arbitrage**: no arbitrage chances
- Type of price discrimination: (1) **first-degree**: fully customized; (2) **second-degree**: menu pricing, use menu price to identify customers; (3) **third-degree**: group pricing, when can identify by observable characteristics
- **Third degree (uniform price + constant MC):** (1) reverse the demand functions; (2) computed aggregated demand (! Mind the bound); (3) Reverse back to demand function and identify the aggregated MR (slope times by 2); (4) compute MC = MR(agg) and find P from aggregated demand
- **Third degree (price discrimination + constant MC):** (1) equate MC = MR separately in each market; (2) get price
- **Third degree (price discrimination + non constant MC):** (1) calculate aggregated demand; (2) identify the aggregated MR; (3) equate MR(agg) = MC; (4) use this MR and refer to demand functions in each market and find each P, Q
- 2 main rules: (1) **MR must be equalized in each market**; (2) **MR must equal aggregate MC**
- Elasticity: P lower in high elasticity market
- Discrimination by locations with identical demand function: $c_j = c_i + t$; $P_j - P_i = t/2$, difference in prices is not the same as difference in costs
- **Pd and welfare**: (1) serve weak and strong markets, $\Delta welfare = (P_u - MC)(\Delta Q_1 + \Delta Q_2)$, total **Q needs to go up to increase welfare**; (2) **open up new market**: when uniform pricing drop one market but price discrimination open up a new market → **welfare increased**

Lecture 6- Price Discrimination and Monopoly: Nonlinear Pricing

- First degree **(two-part pricing)**: (1) same unit price + different entry fees; (2) unit price (best) = MC; (3) entry price = all consumer surplus at P = MC
- First degree **(Block pricing)**: (1) charge one whole package prices (diff) to diff segments; (2) price = same as two-part pricing
- **Second degree**: (1) identify individual lower customer's max (Q,P), package price for L; (2) At the same Q, compute higher customer's surplus; (3) identify individual higher customer's max (Q,P), package price for H: H(Q,P)- surplus from buying package L
- **Second degree**: (1) extract all surplus from low custo; (2) some consumer surplus left on table to intro high customers
- **Welfare**: (1) first-degree price discrimination always increases social welfare; (2) second degree increased social welfare only when total output increased

Lecture 7- 2nd_degree

- N high-income consumers; demand: $P_H = A - Q$
- n low-income consumers; demand: $P_L = a - Q$
- cost function: $C = cQ$; c (marginal costs)
- Willingness to pay: $W_H(Q) = \int_o^Q P_H(x)dx = AQ - \frac{Q^2}{2}$; $W_L(Q) = \int_o^Q P_L(x)dx = aQ - \frac{Q^2}{2}$

- Package: (Q,V); price V at quantity Q
- $\boxed{\text{Option 1}}$: serving only high end consumers: (1) $\boldsymbol{V = W_H(Q)}$; (2) $\boldsymbol{\pi_1 = N[W_H(Q) - cQ] = N(AQ - \frac{Q^2}{2} - cQ)}$; (3) solving $\frac{d\pi_1}{dQ} = 0$; (4) $\boldsymbol{Q_1{}^* = A - c}$; (5) $\boldsymbol{V_1{}^* = W_H(A - c) = \frac{A^2 - C^2}{2}}$; (6) $\boldsymbol{\pi_1{}^* = N\frac{(A-c)^2}{2}}$
- $\boxed{\text{Option 2}}$: serving high and low end consumers with union price (never optimal): (1) $\boldsymbol{V = W_L(Q)}$**;** (2) $\boldsymbol{\pi_2 = (N + n)[W_L(Q) - cQ] = (N + n)(aQ - \frac{Q^2}{2} - cQ)}$; (3) solving $\frac{d\pi_2}{dQ} = 0$; (4) $\boldsymbol{Q_2{}^* = a - c}$; (5) $\boldsymbol{V_2{}^* = W_L(a - c) = \frac{a^2 - C^2}{2}}$; (6) $\boldsymbol{\pi_2{}^* = (N + n)\frac{(a-c)^2}{2}}$
- $\boxed{\text{Option 3}}$: serving high and low end consumers with diff price: (1) $\boldsymbol{V_L = W_L(Q_L);\ V_H = W_H(Q_H) - W_H(Q_L) + W_L(Q_L) = AQ_H - \frac{Q_H{}^2}{2} - (A - a)Q_L}$; (2) (3) solving $\frac{d\pi_3}{dQ_H} = 0$; $\frac{d\pi_3}{dQ_L} = 0$ (4) $\boldsymbol{Q_H{}^* = A - c;\ Q_L{}^* = a - c - \frac{N}{n}(A - a)}$; (6) $\boldsymbol{\pi_3{}^* = \frac{(N+n)(a-c)^2}{2} + \frac{(N+n)N(A-a)^2}{2n}}$
- Option 3 valid only when $Q_L{}^* > 0$, namely, $n(a - c) > N(A - a)$

Lecture 8- Causal inferences

- **A/B Testing**: (1) **control**- existing system; (2) **treatment**- existing system with feature X
- Ways to identify causal relationship: (1) Randomized experiment; (2) Regression; (3) DiD; (4) instrumental variables

Lecture 8- RCT and DiD

- $y_i(w = 1)$: outcome if treated [shorthand y_{i1}]
- $y_i(w = 0)$: outcome if treated [shorthand y_{i0}]
- Issue: missing outcomes
- $y_i{}^{obs} = w_i y_i(1) + (1 - w_i)y_i(0)$; $y_i{}^{mis} = w_i y_i(0) + (1 - w_i)y_i(1)$
- Treatment effect: how much change if treated? $\tau_i = y_{i1} - y_{i0}$
- $\widehat{ATE} = \frac{1}{N}\sum_i^N \tau_i$; $\widehat{ATT} = \frac{1}{N_t}\sum_{i:wi=1} \tau_i$; $\widehat{ATC} = \frac{1}{N_c}\sum_{i:wi=0} \tau_i$
- $\hat{\tau}_{naive} = \overline{y_1{}^{obs}} - \overline{y_0{}^{obs}} = ATT + Baseline\ bias = Outcome\ bias + ATC$
- **Baseline bias**: diff between treated and controls if neither were treated
- **Outcome bias**: diff between if both were treated
- ATT-ATC= Treatment heterogeneity
- **If randomized→** Baseline bias = 0, Outcome bias = 0, ATE=ATT=ATC→regression works
- Check if data is randomized: **Fisher's sharp null:** $H_o = \tau_i = o$
- Block randomized Trials: randomize within each block

Lecture 9- EstVsPred

- R function: **rbinom(t,c,p)**; t: num of trials, c: num of coin tossed per trial, p: prob of getting 1
- R function: **runif(100,0,1)**; 100 times of trial, min = 0, max = 1
- R function: rnorm(t, mean=m, sd=s); t: num of trials
- R function: pnorm(z, mean=m, sd=s); give under probability pnorm(0,0,1)= 0.5
- Lasso regression:(1)$\sum_{i=1}^n (y_i - \sum_j x_{ji}\beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j|$; (2)Least square + penalty λ times slope; (3) tune model with a bit more bias but way less variance

Lecture 10- bandits

- Exploration: get more info by playing diff arms (1)every play improves confidence interval (2)
- allows unlucky arms to catch up; Exploitation: making money on current info
- **A/B testing**: (1) Explore all then exploit all; (2) play all arms; (3) choose best arm; (4) exploit, play that arm for the rest
- **ϵ-greedy**: (1) decide $\epsilon = \frac{\alpha(rounds\ you\ want\ to\ explore)}{t(explore+exploit\ rounds)}$; (2) each round, to explore or to exploit, decided by random select with prob(explore)= ϵ, prob(exploit)=1- ϵ; (3) regret= $\sum_{i=0}^t \epsilon_t \approx \log(t)$
- **Softmax or Boltzmann exploration**: (1) τ: temperature; (2) estimate of are i at time t is $\widehat{\mu_i}(t)$; (3) at time $t + 1$, play arm i with prob :$p_i(t + 1) = \frac{e^{\frac{\widehat{\mu_i}(t)}{\tau}}}{\sum_{j=1}^K e^{\frac{\widehat{\mu_j}(t)}{\tau}}}$; (4) $\tau = 0$, exploitation; (5) $\tau = \infty$, pure exploration
- **UCB**: (1) $n_j(t)$ denote the number of times arm j has been played; (2) $\widehat{\mu_j}(t)$ be the estimate of arm j at round t; (3) $\hat{\mu}_j{}^+(t) = \hat{\mu}_j(t) + \sqrt{\frac{2\log(t)}{n_j(t)}}$; (4) update $\hat{\mu}_j{}^+(t + 1) = \frac{\hat{\mu}_j{}^*(t)*n_j{}^*(t)+x_j{}^*}{n_j{}^*(t)+1}$; (5) increment $n_j{}^*(t + 1) = n_j{}^*(t) + 1$
- Beta distribution: (1) $\frac{\alpha}{\alpha+\beta}$; (2)begin with Beta(1,1) (3) α increase by 1 if success; (4) β increase by 1 if else