Contracts, Health, Education

1 Table of Contents

2 CONTRACTS

- See Part IIA papers for evidence and detail!!

2.1 SET UP

2.1.1 ASSUMPTIONS

- Principal/Employer (firm) is risk neutral; Agent/employee (CEO) is risk averse
  - This creates trade-off between allocative efficiency (employers bear all risk) and productive efficiency (employee bears all risk)
- Effort unobservable and costly, increasing at an increasing rate $[c(e) < 0;\ c'^{(e)} > 0;\ c''(e) > 0]$
  - This necessitates principal to incentivize employee
- Principal uses a linear wage contract $w(z, y) = \alpha + \beta(z + \gamma y)$
  - α shifts base salary and β intensity of incentives (i.e. bonus)
  - γ relative weight of y (part of output irrelevant with e) compared to $z = x + e$ (noise measure of effort)
- One time period

2.1.2 SETTING UP PROBLEMS

Principal

- Employer $\boldsymbol{\max_{\alpha,\beta,\gamma} M = e - w(z, y) = e - [\alpha + \beta(x + e + \gamma y)]}$ s.t. $\boldsymbol{U^{CE} \geq \underline{U}}$,
  - Benefits from employee's effort e
  - Loses from paying a wage
  - Needs to ensure incentive compatibility constraint holds else agent exerts no effort

Agent

- Mean-variance utility: $U^{CE} = E[w(z, y)] - \frac{1}{2}\rho var[w(z, y)] - c(e) = E[\alpha + \beta(x + e + \gamma y)] - \frac{1}{2}\rho var\big(\alpha + \beta(x + e + \gamma y)\big) - c(e) = \alpha + \beta e - \frac{1}{2}\rho\beta^2 var(x + \gamma y) - c(e)$
  - Individuals prefer high return, low volatility, and low effort
  - Uncertainty comes from z, y

Formally deriving Utility Function

- Certainty Equivalence is $w^{CE} = E[w] - r$ such that $u(w^{CE}) = EU(w)$
  - where r is amount willing to avoid risk

- Second Order Taylor expansion: $u(w) = u(w_0) + u'(w_0)(w - w_0) + \frac{1}{2}u''(w_0)(w - w_0)^2$
  - Now sub in $w_0 = E[w]$ to get $u(w) = u(E[w]) + \frac{1}{2}u''(E[w])E[(w - E[w])^2]$
- First-order Talyor expansion of $u(w^{CE}) = u(E[w]) + u'(E[w])(w^{CE} - E[w])$
- Equating these and rearrange: $w^{CE} = E[w] - \frac{1}{2}\frac{-u''(E[w])}{u'(E[w])}E[(w - E[w])^2] = E[w] - \frac{1}{2}\rho var(w)$

2.2 SOLVING

- Use backward recursion strategy where (1) find optimal e^* given any wage scheme and (2) find optimal α, β, γ given e^*

2.2.1 SOLVING AGENT PROBLEM

- Agent solves $\max_{e} \alpha + \beta e - \frac{1}{2}\rho\beta^2 var(x + \gamma y) - c(e)$
- Thus FOC wrt e gives $\beta - c'(e) = 0$

- $\frac{d\beta}{de^*} = c''(e^*) > 0$ thus larger β gives agent more incentive to put in effort
    - To incentivize agent to exert more effort $\uparrow e^*$, employers should increase incentive intensity $\uparrow \beta$
- $\frac{de^*}{d\beta} = \frac{1}{c''(e^*)}$ captures responsiveness of effort to incentives β. Show graph slide 2.10

2.2.2 *Solving Principal Problem*

- Rewrite problem as $\max e^* - [\alpha + \beta(x + e^* + \gamma y)]$ s.t. $U^{CE}(e^*) \geq \underline{U}$ where $\beta = c'(e^*)$

Solve α

- α is set to level that binds ICC: $U^{CE}(e^*) = \underline{U}$ thus $\alpha + \beta e^* - \frac{1}{2}\rho\beta^2 var(x + \gamma y) - c(e^*) = \underline{U}$
    - Intuitively, all rents go to Principal after agent gets utility equal to best outside option
- $\alpha = \underline{U} - \beta e^* + \frac{1}{2}\rho\beta^2 var(x + \gamma y) + c(e^*)$
    - α depends on (β, γ)

Solve β

- Substituting α back in: $\max_{\beta,\gamma} e^* - \frac{1}{2}\rho\beta^2 var(x + \gamma y) - c(e^*) - \underline{U}$
- Taking FOC wrt β gives $\frac{de^*}{d\beta^*} - \rho\beta^* var(x + \gamma y) - c'(e)\frac{de^*}{d\beta^*} = 0$
- Substituting in $\frac{de^*}{d\beta^*} = \frac{1}{c''(e^*)}$ we get $1 = \rho\beta^* var(x + \gamma y)c''(e) + c'(e^*) = 0$
- Rearrange to solve $\beta^* = \frac{1}{1+\rho var(x+\gamma y)c''(e^*)}$

2.3 Principles of Incentive Pay

2.3.1 *Incentive Intensity Principle*

- What does β depend on? $\beta^* = \frac{1}{1+\rho var(x+\gamma y)c''(e^*)}$
    - $\uparrow var(x + \gamma y)$, less precise signal, makes no sense to link incentive to signal, $\downarrow \beta^*$
    - $\uparrow \rho$, more risk averse, more compensation needed via α, less left for incentive, $\downarrow \beta^*$
    - $\uparrow \frac{1}{c''(e^*)}$, more responsive to incentive, more efficient to use signal, $\uparrow \beta^*$

2.3.2 *Informativeness Principle*

- How do we set γ? Note we want to min $var(x + \gamma y) = var(x) + \gamma^2 var(y) + 2\gamma cov(x, y)$
- FOC wrt γ gives $2\gamma^* var(y) + 2cov(x, y) = 0$ thus $\gamma^* = -\frac{cov(x,y)}{var(y)}$
    - $\uparrow Var(y), \uparrow$ signal noise, $\uparrow$ attach less weight to signal, $\downarrow |y^*|$
    - $\uparrow Cov(\epsilon, y)$. Now if y is high ϵ is likely high, so high x may be more due to ϵ, so worker should get less, so this is reflected in negative γ
        - If no correlation signal is uninformative noise so attach no weight $\gamma^* = 0$
- <u>Case for relative performance pay</u>:
    - Assume one manager's output is $z = x + e$ and another is y. Both subject to similar shocks, then y is informative due to association with x. Relative performance pay!
    - Formally, assume two agents A and B exposed to same general shock ϵ_C.
        - Abs: $x_A = e_A + \epsilon_A + \epsilon_C$ thus $\max \alpha + \beta e_A - \frac{1}{2}r\beta^2 Var(\epsilon_A + \epsilon_C) - c(e_A)$
        - Rel: $x_A = e_A - e_B + \epsilon_A - \epsilon_B$ thus $\max \alpha + \beta(e_A - e_B) - \frac{1}{2}r\beta^2 Var(\epsilon_A - \epsilon_B) - c(e_A)$

- In both cases $\frac{dc(e_A)}{de_A}\Big|_{e_A^*} = \beta$ i.e. induce same level of effort
- But relative has less risk premium: Assuming independence $Var(\epsilon_A + \epsilon_C) > Var(\epsilon_A - \epsilon_B)$ if $Var[\epsilon_C] > Var[\epsilon_B]$
  - Lower risk premium → less cost borne due to risk → need to compensate less

2.3.3 *Equal Compensation Principle*

- Assume multiple tasks e_1, e_2 and Principal wants both to be done $e_1, e_2 > 0$
- Agents solves $\max_{e_1,e_2} \alpha + \beta_1 e_1 + \beta_2 e_2 - \frac{1}{2}\rho var(\beta_1 x_1 + \beta_2 x_2) - c(e_1 + e_2)$
- FOC wrt e_i gives interior solution $\beta_1 = c'(e_1 + e_2) = \beta_2$
  - If Principal wants agent to do both tasks positively and their marginal costs are the same, their marginal benefits need to be the same to
  - If $\beta_1 > \beta_2$ then all effort goes to task 1 and we get a corner solution instead

2.3.4 *Ratchet Effect*

- What if employers tend to base next year's targets on last year's performance? Managers haver perverse incentive not to exceed targets, even if can do so easily
- Principal sets $w_t(z_t) = \alpha_t + \beta_t(e_t + x_t)$, observes $z_t = x_t + e_t$ and then updates beliefs of distribution $x_t \sim E(x_t), var(x_t)$ based on realization
- Suppose a good shock occurs in period t where $z_t > e_t + E(x_t)$. Principal will update [Bayesian] beliefs so $E(x_{t+1}) > E(x_t); var(x_{t+1}) < var(x_t)$
- If agent follows static optional choice then $\alpha_{t+1}^* < \alpha_t^*; \beta_{t+1}^* > \beta_t^*$ and gets $\underline{U}$ in both periods
  - where $\alpha_t^* = \underline{U} - \beta\big(e^* + E(x_t)\big) + \frac{1}{2}\rho\beta^2 var(x_t) + c(e^*)$ and $\beta_t^* = \frac{1}{1+\rho var(x_t)c''(e^*)}$
- But if exert less effort in good time $\bar{e}_t < e_t^*$ so that $z_t = e_t + E(x_t)$ then principal will not notice shirk in first period as output may not drop. CEO now gets higher utility $\underline{U} + c(e_t^*) - c(\bar{e}_t) > \underline{U}$

2.3.5 *Principles of Efficiency*

- Consider simple example $w = \alpha + \beta(x + e)$
- If $\beta = 1$ full benefit of higher effort goes to agent: $\frac{dw}{de} = \beta = 1$. Thus most productive efficiency
- If $\beta = 1$ full effect of shock effort goes to agent: $\frac{dw}{dz} = \beta = 1$. Thus lowest allocative efficiency
- Reverses for $\beta = 0$

2.4 Evidence

- Also see Part IIA notes

Do workers respond?

- Lazear (1996): Impact of switch from fixed salaries ($\beta = 0$) to piece rates ($\beta > 0$) on performance of workers installing auto windshields resulted in 20-36% boost in productivity and 12% total pay
  - 1/3 due to attracting better workers and 2/3 due to more effort
  - But… external validity of applying simple jobs to complex e.g. CEOs
  - But… not a test of agency theory but input to the theory

Do contracts reflect agency concerns?

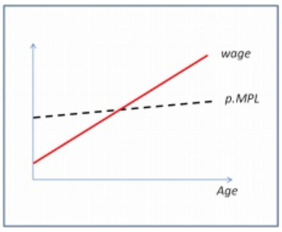
- β: Jensen and Murphy (1990) estimate CEO lose \$3.25 for every \$1000 firm loss in market value, mostly driven by stock ownership
  - May be underestimate as cannot control for unobservable like ρ and x
  - But seems large for many Fortune 500 firms are so large this matters

- More appropriate is to look at % changes to account for size so $w = \alpha + \beta(b(S)e + x)$ where $b(S)$ is function of firm sizes
  - New optimal intensity $\beta = \frac{c'(e)}{b(S)}$ so for same effort level, β is negatively correlated with firm size, explaining why 'intensity' is so low for big firms. See also slide 15
- ρ: Aggarwal and Samwick (1988) classify S&P firms 1993-96 by variation in stock return volatile to test if pay-performance sensitivity decreases as predicted. Validate this claim
- γ: Gibbons and Murphy (1990): Tests informativeness principle to see if relative performance pay matters. Indeed, CEOs are penalized if competitors fare better
  - But… relevant peer groups are entire stock market, not companies in same industry, even though the 'common shock' assumption is less strong here!

2.5 ALTERNATIVES

- Why us use of contracts limited for workers within firms? Prendergast (1999) notes:
  - Cannot write all multi-tasking into contract so agents will always "game" the system
  - Subject evaluations allow for more holistic picture of performance, even if opens to behavioural dynamics
  - Hierarchies creates dangers of rent-seeking activities, where only want to please Principal

2.5.1 DEFERRED COMPENSATION



- Observe that workers are overpaid when old and underpaid when young
- May be because (1) reduces turnover rate for young workers (helps workers acquire firm specific capital) and (2) deals with issue of verification of efforts is significantly delayed

2.5.2 EFFICIENCY WAGES

- Firms overpay workers in order to make them exert more effort, assuming that monitoring is costly to firm
- Set up as follows: CEO gets paid w; decision to exert effort e be binary and marginal cost $\frac{c}{2}$; if shirk then caught with probability p and get nothing
- Thus see CEO exerts effort iff $w - \frac{c}{2} \geq (1-p)w$ so $w \geq \frac{c}{2p}$
- CEO earns rents from insufficient monitoring ($p < 1$) as $w - \frac{c}{2} = \frac{(1-p)c}{2p}$

2.5.3 TOURNAMENTS

- Firms reward workers based on relative performance (rank order) instead of absolute performance (piece rate)
- For example, consider "winner takes all tournament"
  - Wage will no longer reflect marginal product
  - This is common when worker's output is expensive or unreliable e.g. sports
    - Ehrenberg and Bognanno (1990): golfers have better scores when there is larger prize money
    - Becker and Huselid (1992): higher prizes result in faster, though riskier, NASCAR race times
    - Conyon and Peck (1997): return to becoming CEO increases with number of individuals competing at rank below

3 Regulation

3.1 Monopoly Motivation

- Fully competitive market has productive and allocative efficiency since $p = MC = AC$.
- Monopoly is neither:
  - $p^* > MC$ creating DWL so not allocatively efficient
  - Output does not occur at lowest point on AC curve so not productively efficient
- Several ways how monopolies develop:
  - Horizontal integration: M&A between firms at same position of production chain
  - Vertical integration: gain market power by controlling stages of production process
  - Legal monopoly
  - Internal expansion in new industry (blue ocean strategy)

3.2 Regulating Monopolies

- With natural [or legal] monopolies consumer demands is often super inelastic so the firm's price setting power risks exploitation. How do we solve this?

3.2.1 Rate of Return Regulation [US]

- Set up $R = E + s * B$ where R is total revenue, E is expenses, s is regulated rate of return, and B is regulatory asset base
- Company has opportunity to recover costs from investment in B by earning s, whilst cap on s prevents consumer exploitation
  - Irl firm can decide complicated price structure as long as average RoR does not exceed s
  - Regulated firms submit detailed cost breakdown and government separates into un-/reasonable and capital/maintenance parts

Averch-Johnson Effect

- Assume neoclassical production function where $f(k, l); f' > 0; f'' < 0$
- Allows firm to earn RoR above cost of capital

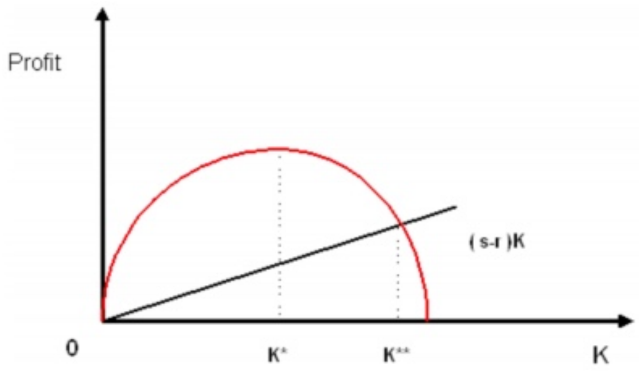
No regulation [Benchmark]

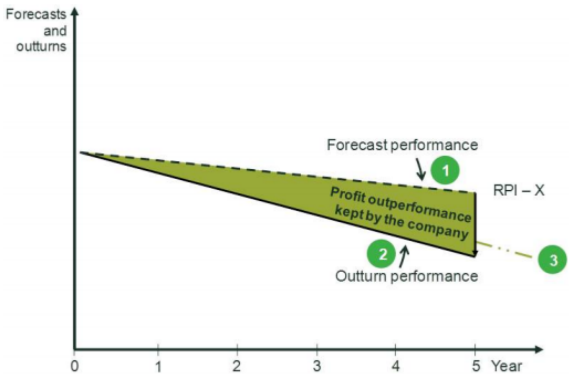
- Monopoly solves $\max_{k,l} \mathcal{L} = pf(k, l) - rk - wl$
- FOC wrt k, l give $pf_k = p\frac{df(k,l)}{dk} = r$ and $pf_l = p\frac{df(k,l)}{dl} = w$
- Thus $MRT^*_{kl} = \frac{\left(\frac{df}{dk}\right)}{\left(\frac{df}{dl}\right)} = \frac{r}{w}$
  - Have productive efficiency as we use k, l in right proportion
  - Do not have allocative efficiency, as too little is produced

Regulation

- Now cap on profit-to-capital ratio so $\frac{\pi}{k} \leq s - r$ where $s > r$ [this will hold to max profits]
  - Sub in π get $\frac{pf(k,l) - rk - wl}{k} \leq (s - r)$ thus $pf(k, l) - \mathrm{w}l - sk \leq 0$
- Monopoly solves $\max_{k,l} \mathcal{L} = pf(k, l) - rk - wl + \lambda[sk - pf(k, l) + \mathrm{w}l]$
- FOC gives $\frac{d\mathcal{L}}{dk} = p\frac{df}{dk} - r + \lambda s - \lambda p\frac{df}{dk} = 0$ and $\frac{d\mathcal{L}}{dl} = p\frac{df}{dl} - w - \lambda p\frac{df}{dl} + \lambda w = 0$
  - Rearranges to $pf_k = p\frac{df}{dk} = \frac{r - \lambda s}{1 - \lambda}$ and $pf_l = p\frac{df}{dl} = \frac{w - \lambda w}{1 - \lambda}$
- Thus $MRT_{kl} = \frac{\left(\frac{df}{dk}\right)}{\left(\frac{df}{dl}\right)} = \frac{r - \lambda s}{w(1 - \lambda)} = \frac{r(1 - \lambda) - \lambda(s - r)}{w(1 - \lambda)} \frac{r}{w} - \frac{\lambda(s - r)}{w(1 - \lambda)} < \frac{r}{w}$ for all $s > r$

- Do not have productive efficiency since overinvest in capital. Intuitively, profits are set by use of capital, thus want to use more than "optimal"
    - When $s = r$ there is no excessive return so no problem!
  - Less allocative efficiency as more is produced now overall
- Larger regulator lag, smaller A-J effect: $s \to r$ larger $\to \lambda$ smaller
- Aside: We can show that $\lambda \in (0,1)$ since $\frac{d\mathcal{L}}{dp} = f(k,l)(1-\lambda) > 0$ as profits must rise if price of inelastic good increases





3.2.2 *RPI-X [UK]*
- Incentive based regulation where regulators sets price control for period (updated every 4-5 years) and leaves everything else to monopoly
- Regulator allows price to adjust over time by RPI minus expected efficiency improvement in costs so $P_{t+1} = P_t + RPI - X$
  - X is proxy for competitive market, reflecting firm and competitor's past/forecast
- Within that time frame, price is independent of cost! Thus stimulates cost reduction, since exceeding X allows to earn profit and falling behind costs
  - Companies are incentivised to save money as this is retained until next review
  - Especially early earnings are valuable
- Pros
  - companies control within price cap
  - simpler than RoR cap
  - flexibility to adjust prices
- Cons
  - Dependent on estimating efficient cost accurately, which is hard forecast
  - Underinvestment if updated too frequently since will retain profits for shorter period and this may make long-term savings not worthwhile
- Does it work? See slide 4.25

4 In-Kind vs. Cash Transfers
- Basic theory says cash dominated as in-kind places constraints on recipient's behaviour and is associated with more admin (although less risk of corruption)
- Consider three cases how they may still be justified…

4.1 Asymmetric Information and Self-Selection
- May have several obstacles prevent fist-best redistribution
  - Government is potentially unable to distinguish between types
  - Lack of information on preferences and income

- Allocation problem solved via in-kind benefits so that [e.g.] only the sick use health care
    - Good must not be easily tradeable but… still need to worry about use i.e. rationing
- Why do [e.g.] healthy still support in-kind transfers (additional tax)? Assume people are risk averse and uncertain about type. Redistribution is thus insurance against health risk

4.1.1 *Besley and Coate (1991): Set Up*

- Assume two types of individuals: low- and high income $J = y_l, y_h$
- Individuals consume numeraire z and a one-unit good x that varies in quality q, with the price of improving marginal unit of quality being p
    - Get quasi-concave utility function: $v(p, y) = \max_q u(z, q)$ s.t. $z + pq \leq y$
- Government provides x for free but offers low quality q_g. Individuals thus choose between
    - Public provision: $u(y, q_g)$
    - Private purchase: $v(p, y) = \max_q u(y - pq, q)$

4.1.2 *Solving*

- Government wants to choose q_g so only poor consume x?
    - $u(y_l, q_g) \geq \max_q u(y_l - pq, q) = v(p, y_l)$: Optimal choice of q_h is far from q_g (i.e. allocation distortion≥ benefit from in-kind transfer)
    - $u(y_h, q_g) \leq \max_q u(y_h - pq, q) = v(p, y_h)$: Optimal choice of q_h is near q_g (i.e. allocation distortion ≤ benefit from in-kind transfer)
- In-kind benefits are an imperfect instrument for poverty relief. Only low quality benefits provided so that rich do not consume them
    - Same amount of cash transfer $v(p, y_l + pq_g)$
    - In-kind transfer: $u(y_l, q_g) \leq v(p, y_l + pq_g)$ equal when $q_l = q_g$

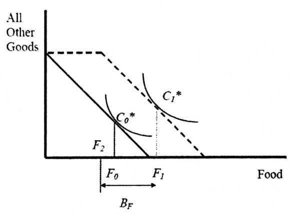
4.2 Domain Specific Altruism

"Rich" may care about "poor" so redistribution can be actually Pareto-improving
But assume that rich care do not care about utility of poor per se (else cash dominates):

- Merit goods (Musgrave, 1959): want to encourage consumption of certain types of goods (e.g. health care) rather than utility level
- Paternalism: view that individual consumption choices fail to maximize own utility so override these

Thus in-kind transfers means rich are willing to support more redistribution

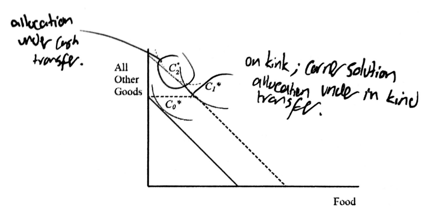
Predictions for inframarginal recipients



- C_0^* is the optimal consumption without food stamps and C_1^* is the optimal consumption with food stamps
- Both food and nonfood consumptions increase (food: $F_0 \to F_1$)
- Out of pocket food spending decreases ($F_0 \to F_2$)
- Overall food consumption goes up by less than food stamp benefits

Predictions for constrained recipients

rw do u-lines a



- Compared with the desired allocation without in-kind restriction C_2^*, the constrained allocation C_1^* chooses more food and less other good consumptions

- Behaviour depends on type of recipient:
    - Inframarginal: Food consumption with in-kind = with cash transfers

- Constrained recipients: Food consumption with in-kind > with cash transfers

[How do u-lines cross?]

- Consider evidence from SNAP (food stamp) program:
  - SNAP led to much more spending on food than cash (4-10x)
  - Few causal estimates show: Most households are inframarginal as SNAP leads to increases in food spending that are similar to cash

[How is this relevant?]

4.3 Price Effects

- Proposed by Coate et al (1994): cash transfers increase demand for normal goods, increasing price. In-kind transfers also increase supply, keeping price low

Supported Cunha et al (2014): Re-examine 2003 RCT in rural Mexico where 200 villages either received boxes of food, equivalent value in cash transfer, or nothing.

- Find that food prices significantly lower under in-kind (4% relative to control)
- Cash transfers had small positive effect
- Supply effect was largest in remote villages

5 Health Care

5.1 Motivation

- Individuals are uncertain whether they require insurance care. Since they are risk averse, they are willing to pay a premium to insure against financial risks
  - Premium reflects value of certainty provided by full insurance, reflecting degrees of risk aversion as well as magnitude and probability of loss
  - Catastrophic health risks are low probability but large, thus exactly what this need!
- A market for health insurance may exist if
  - (1) Future health states are uncertain (not routine care or pre-existing conditions)
  - (2) Incidence of sickness is independent (not epidemics)
  - (3) Value of risk premium is large / consumers are risk averse
- Private company can profitably pools risks (i.e. employer based insurance) if has large number and risks are idiosyncratic

Summary Points

- Experience rating
  - If insurers can identify risk types, different markets develop for different types (personal characteristics, medical history etc.) (closer to first best)
  - But… what happens to high-risk that cannot be insured (elderly, chronically ill); may be 'efficient' but also inequitable
- Screening / rationing
  - Lower risk types offered cheaper but limited coverage (see R&S 1976)
  - But… under provision

5.2 Adverse Selection (Rothschild & Stiglitz, 1976)

- Use lecture 6, Vas, and supo

5.2.1 Set Up

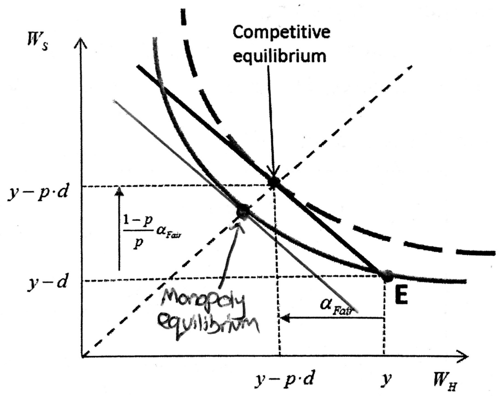
- Two types of insurers $i = \{L, H\}$ with insurance firms offering contracts that are contingent on states realisation Bad or Good
- Consumer maximizes welfare $V(p_i, \alpha) = (1 - p_i)u(y - \alpha_G) + p_i u(y - \delta + \alpha_B)$
    - o α_G is amount paid out if insurer is healthy (good)
    - o α_B is premium paid out if insurer is sick (bad)
    - o p_i is probability of falling ill and differs for the two types: $p_L < p_H$
    - o δ is damage from bad state
    - o The slope of an IC is thus $-\frac{1-p_i}{p_i}\frac{u'(w_G)}{u'(w_B)}$
- Individuals start with endowment E. Graphically, x-axis represents wealth in state G and the y-axis in state B
- Firm maximizes profits $\pi = (1 - p_i)\alpha_G - p_i\alpha_B$. Under perfect competition zpc holds
    - o This gives a line with slope $-\frac{1-p}{p}$ where $p = \frac{n_L p_L + n_H p_H}{n_L + n_H}$
- We sketch these three components

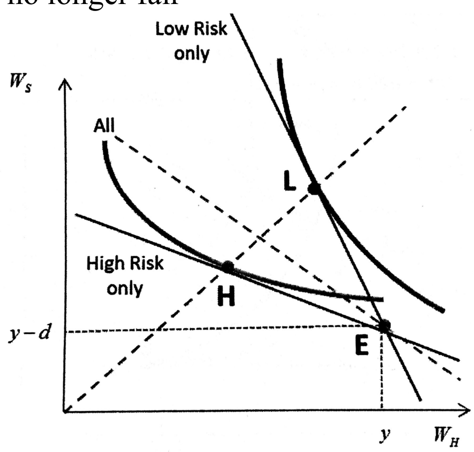
- Define equilibrium as set of contracts x^J offered by $J \geq 2$ firms and purchased by consumers s.t.
    - o consumers prefer no contract offered over the one purchased
    - o no firm can offer contract β where consumers strictly prefer β and firms make profit

5.2.2 First Best Benchmark – Full Information

Homogenous Consumers

- Pooling equilibrium: All consumers purchase same contract
- Note difference based on market:
    - o Competitive Equilibrium: Consumers fully insured, allocative efficiency (utility same in both states), fair prices (equal expected loss)
- Monopoly Equilibrium: Same but prices are no longer fair





Heterogenous Consumers

- Separating equilibrium: Different consumers purchase different contracts

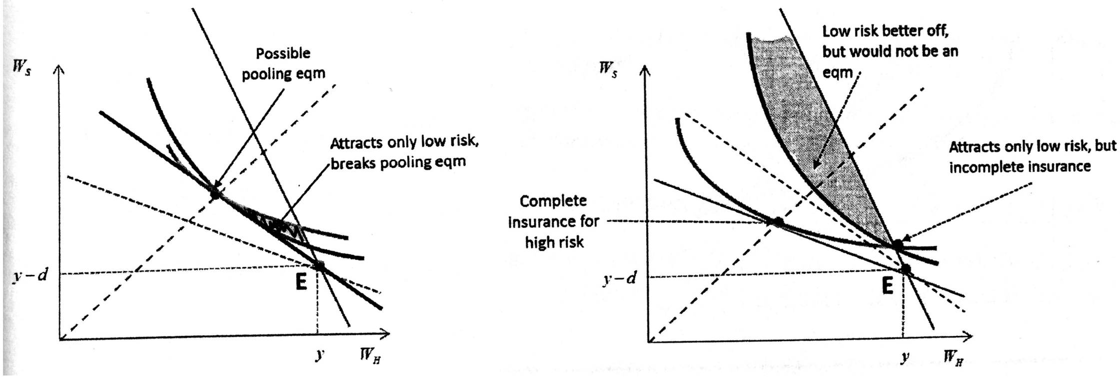
5.2.3 Second Best Benchmark – Information Asymmetry

Failure of Pooling Equilibrium

- Market can provide full and "fair" pooling insurance but this is not sustainable
- IC's cross where high-risk IC is flatter than low-risk one. Thus exist partial insurance contracts that are preferred by L but not H and this profitable for firm. Thus "cream skimming" deviation

Separating Equilibrium / Adverse Selection

- Issue of Adverse Selection: Full insurance cannot be sold at price fair for low-risk since it would attract high risk. Only incomplete coverage can be sold at price that is fair for low risk.
    - H types offered complete insurance contract at their fair price
    - L types offered best possible coverage subject to constraint that H will not select into contract. Thus only offered partial insurance



5.2.4 *Empirical Evidence (Cutler & Zeckhauser, 1997)*

- 1995-96 Harvard University offered two healthcare plans: PPO (most generous) (20% share) and HMO (tightly managed) (80% share). PPO cost $500 more and Harvard mostly covered this
- Adverse-selection response: young and healthy groups left PPO (now 15% share) and its premiums had to rise $500 further to cover higher costs. Vicious cycle until failed in 1997

5.3 Government Provision

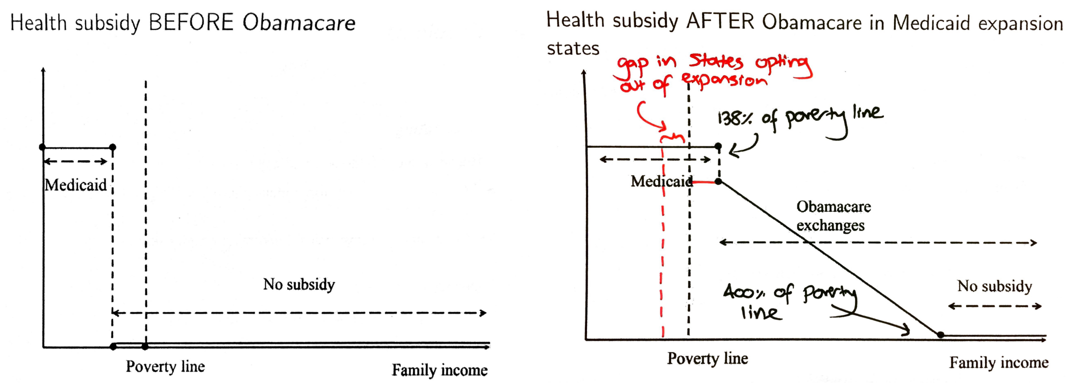
Government either directly controls hospitals (e.g. UK NHS) or reimburses private providers (e.g. France) and thus provide universal healthcare funded by taxation
Has three advantages:

- Prohibits experience rating
- Makes insurance compulsory
- Provides insurance to those who couldn't afford it otherwise

5.3.1 *Obamacare Example*

In 2013 16% had no insurance b/c pre-conditions, over-confidence, or budget constrained
Solution was to force a pooling equilibrium:

- Compulsory to accept all applicants under same rate and fine/individual-mandate those who opt-out
- Extended subsidies those below 400% of poverty line and Medicaid below 138%



- In 2017 fell to 9% uninsured. People left are undocumented immigrants, low income in non-expansionary states, and people who still decided to opt-out/pay fine

5.4 Moral Hazzard

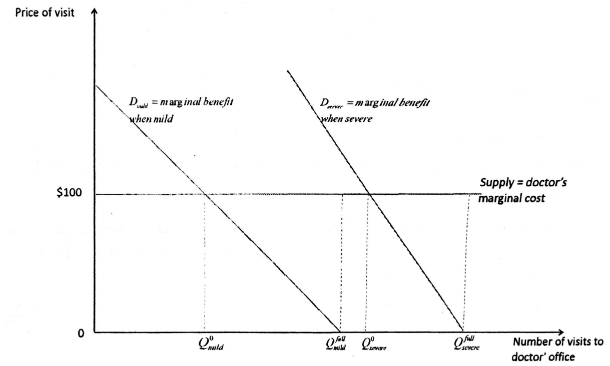
States that probability of becoming sick p and cost of treatment d are not exogenous

- Ex-ante: may take less care in avoiding risk (e.g. pregnancy) and such services are usually excluded from private insurance
- Ex-post: Patients may choose more expensive, higher quality care. Doctors may recommend unnecessary treatment

5.4.1 Model

Set Up

- Assume basic elements:
  - Two diseases: Mild and Severe
  - MC of vising doctor per visit Q is fixed at £100;
  - MB of visiting doctor is diminishing in Q i.e. downward sloping

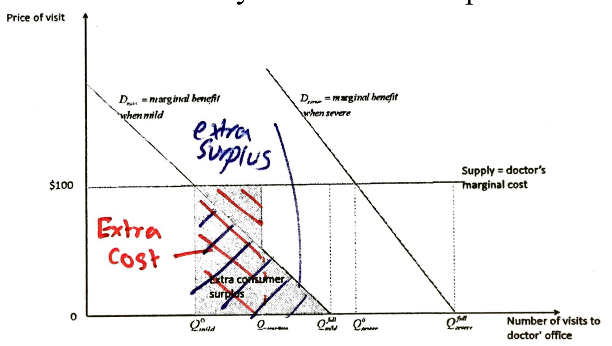


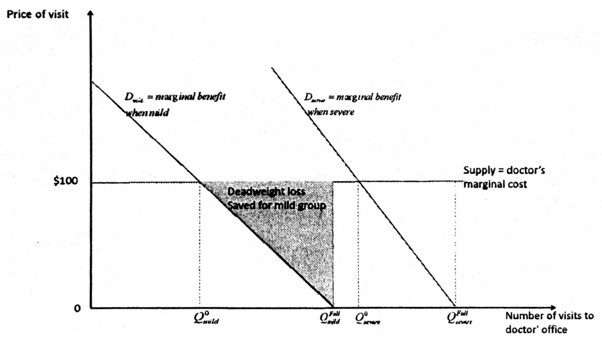
- Observe that full insurance leads to more health cost: $Q^{full} > Q^0$
- Consequently, a larger premium is required to cover additional costs. Some may thus prefer no insurance to full insurance!

Solution: Deductible

- Consumer has to pay $D = MC \times Q_I$ before making claim. Equivalent to choosing Q_I
- Consumer thus chooses either (1) buy an insurance with D or (2) no insurance
  - If S: still prefer deductible as $Q^0_{severe} > Q_{insurance}$
  - If M: ambiguous if surplus < extra payment

- Deductible discourages claims when costs of treatment are small and thus DWL of M
  - Consumer bears mild risks but is still insured against severe losses
  - May allow for lower premiums so could actually prefer this to full insurance

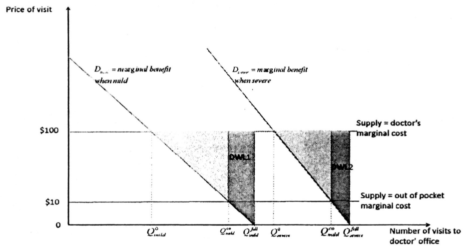




Solution: Co-Insurance

- Consumer is charged price for health care but price is lower than market. Choose Q_{mild}^{co} so MB = money out of their pocket
- Let β denote fraction of market price that individual must pay
    - Allocative efficiency when $\beta = 0$
    - Productive efficiently (i.e. no moral hazard) when $\beta = 1$

Assume $\beta = 0.1$, the payment out of pocket is \$10



Social planner: deductible insurance helps to partially reduce the dead weight loss both in mild diseases (DWL1) and severe diseases (DWL2)

5.4.2 Evidence: Consumer Side

Manning et al. (1986)

- RAND Health Insurance Experiment in US 1970s tracked 6000 people over 3 years
- Randomized assignment to different cost sharing schemes
    - Control: reimburse all covered expenses in full
    - Treatment: 12 different fee-for-service insurance plans over two dimensions
        - Coinsurance (% paid out of pocket)
        - Deductibles (up to max of \$1,000)
    - All families given \$1000 to participate so none made worse off (no opportunity cost)
- Results:
    - Found demand somewhat price sensitive: free care plan used +46% than 95% coinsurance. Price elasticity is approximately 0.2
    - But… no measurable effect on health outcomes except for chronically ill who do not have enough income to easily cover co-payments

Finkelstein et al (2012)

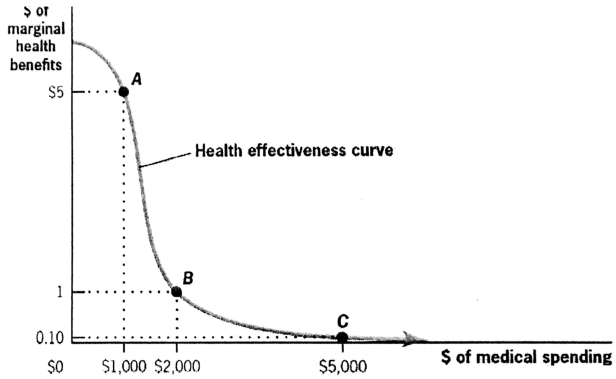
- In 2008 Oregon had limited Medicaid budget so picked 30,000 lottery winners out of 90,000 participants
- Winning the lottery increased probability of having health insurance by 29%-points. Use lottery as instrument to estimate causal effect of insurance coverage
- Note we can estimate two coefficients:
    - ITT: $\beta_{ITT} = E(Y_i|Z_i = 1) - E(Y_i|Z_i = 0)$
    - LATE: $\beta_{LATE} = \frac{E(Y_i|Z_i=1)-E(Y_i|Z_i=0)}{P(D_i=1|Z_i=1)-P(D_i=1|Z_i=0)}$ [what is difference?]
- Results that lottery…
    - Increases health care utilization (e.g. 2%-point or 30% increase in probability of having a hospital admission
    - Decreases out-of-pocket medical expenditures (e.g. 20%-point or 35% decline in having any out-of-pocket medical expenditure)
    - Better self-reported physical and mental health (averaging two-tenths of a standard deviation improvement)

Card et al (2008, 2009)

- Notes that medicare becomes available when turn 65. Hence take discontinuity there
- Results
  - Increase in coverage, especially for disadvantaged groups (i.e. ethnic minorities):
  - Increase in healthcare utilization:
  - Decrease in mortality after admission for conditions requiring ER immediate hospitalization
    - "Nearly 1% -point drop in 7-day mortality for patients at age 65, equivalent to a 20% reduction in deaths for this severely ill patient group. The mortality gap persists for at least 9 months after admission."

Overview

- How do we rectify that Manning finds no effect but others do?
- Studies examine different parts of the "medical effectiveness curve"! Moving from uninsured to some insurance has positive effect; adding to existing insurance does not



5.4.3 EVIDENCE: PRODUCER SIDE

- So far assumed doctors provide treatment that patient requests. Know that irl it is closer to the opposite way around!

Framework

- Payment for physician services: $p = \alpha + \beta c$ where α is fixed cost payment and β proportional cost payment

Retrospective or "fee-for-service"

- $\alpha = 0, \beta = 1$ e.g. US (although now moving away)
- Allocative efficiency: Doctors paid according to costs they incur
- Productive inefficiency: do not produce at minimum cost as full pass through of costs

Prospective or "diagnosis fixed payments"

- $\alpha > 0, \beta = 0$ e.g. UK
- Allocative inefficiency: little risk borne by insurance; price paid can diverge from true cost
- Productive efficiency: Strong incentive to minimise costs but also [unobservable] quality

Cutler (1993)

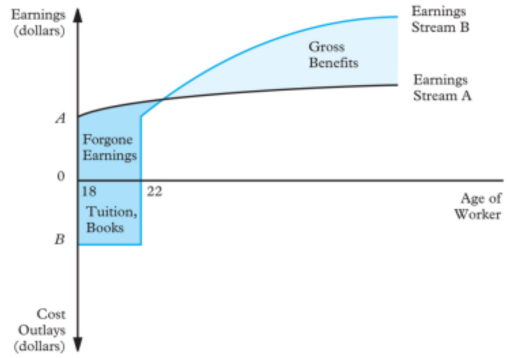
- In 1983 Medicare moved to PPS based on nationally standardized payments for specific diagnoses. Thus have natural regression

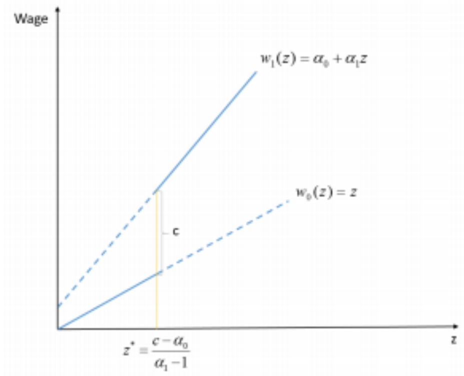
- Results:
    - Reduction in treatment intensity: avg. hospital stay for elderly fell by 1.3 days
    - No adverse impact on patient outcomes
    - Suggests they had been practicing "flat of the curve" medicine

6 Education

6.1 Human Capital Theory (Mincer, 1958; Becker, 1964)

- Treat education as investment where people balance costs of acquiring human capital (direct expenses; forgone earning; psychic losses) and expected returns (future earnings, satisfaction)
- Graph appears to be from Becker





6.1.1 *Mincer Model (1958)*

Set Up

- Workers are ex ante identical
- Different occupations require different lengths of schooling
- Costs of education consists of foregone earnings only
- Perfect information and credit markets
- Workers live infinitely

Solving

- Present Value of lifetime earnings for no schooling is $PV_0 = \sum_{t=0}^{\infty} \delta^t w_0 = \frac{w_0}{1-\delta}$ and with s schooling is $PV_s = \sum_{t=s}^{\infty} \delta^t w_s = \frac{w_s \delta^s}{1-\delta}$
    - Where $\delta = (1+i)^{-1}$ is discount rate and i market interest rate
- <u>Compensating Difference Principle</u>: Since all workers are identical, individuals must be indifferent among all occupations in equilibrium. Thus lifetime earnings equal across all
- Thus $PV_0 = PV_s$ so $\frac{w_0}{1-\delta} = \frac{w_s \delta^s}{1-\delta}$ so $w_s = w_0 \delta^{-s} = w_0 (1+i)^s$
- Assume that % increase in earnings from a year of education is r: $w_s = w_0 (1+r)^s$
    - In equilibrium internal rate of return should be equal to interest rate!
    - Returns $r > i$ indicate under-investment in education. IRL many studies estimate rate of return 7-12%
        - But… ability bias: Do not distinguish between contribution that innate ability makes to higher earnings and the contribution by schooling
        - But… returns to whom? [ATE vs. LATE] People who seek more schooling may have lower psychic costs of learning

6.1.2 *One Factor Model*

Set Up

- Not realistic to assume that everyone is identical. Thus…
- Individuals distinguished by type z distributed uniformly $z \sim u[0,1]$ observed by all
- Obtaining education is dummy choice with cost c
- Return to no education is $w_0(z) = z$ and with education $w_1(z) = \alpha_0 + \alpha_1 z$
    - $\alpha_1 > 1$ implies education is complementary to ability
    - $c > \alpha_0 > 0$ so least talented individual will not choose education

Solving

- Choose education iff $w_1(z) - w_0(z) > c \implies z > z^* = \frac{c-\alpha_0}{\alpha_1 - 1}$
- Find out equilibrium wages as follows:
    - No educated group (avg. wage) $\bar{w}_0 = \frac{(0+z^*)}{2} = \frac{c-\alpha_0}{2(\alpha_1-1)}$ [because uniform take halfway]
    - Educated group (avg. wage) $\bar{w}_1 = \alpha_0 + \alpha_1 \frac{(1+z^*)}{2} = \alpha_0 + \alpha_1 \frac{c-\alpha_0+\alpha_1-1}{2(\alpha_1-1)}$
    - Thus average difference: $\bar{w}_1 - \bar{w}_0 = \alpha_0 + \frac{c+\alpha_1-\alpha_0}{2} > \alpha_0$

- Reflects two components: Note that $\overline{w_1} - \overline{w_0} = \alpha_0 + (\alpha_1 - 1)\left[\frac{c-\alpha_0}{2(\alpha_1-1)}\right] + \frac{\alpha_1}{2}$
    - First two terms give true return to education evaluated at the mean ability of the uneducated group $\frac{z^*}{2}$
        - Return to education is not α_0 but $\alpha_0 + \alpha_1 z$. Therefor for a group of mean ability $\bar{z}$ the return to education is $w_1(\bar{z}) - w_0(\bar{z}) = \alpha_0 + (\alpha_1 - 1)\bar{z}$
    - Third term is the additional effect that results from the selection effect: High ability individuals are selected into education increasing the wage differential

Becker Comments:

- And since "abler" persons tend to invest more than others, the distribution of earnings could be very unequal and even skewed, even though "ability" were symmetrically and not too unequally distribute
- Not just schools but also on the job training
    - Some investments in human capital do not affect earnings because costs are paid and returns are collected by the firms,

6.2 Signalling Theory

- See Lecture 10, Vas, and supo L2 Q1 (no need to repeat), especially for first-best second-best def

6.3 Evidence Debate

Seems sensible to write essay plan for this

6.3.1 Motivation

- Note signalling shares many predictions, or is at least consistent
    - People with more years of schooling are more productive and receive higher wage
    - People will attend school before they enter workforce
    - Consistent with rate of return to schooling being roughly equal to rate of interest
- Pure tests are hard irl:
    - Measure productivity before and after education – pure test of human capital
    - Assign people with identical ability diploma randomly – pure test of signalling

6.3.2 Sheepskin Effect

- People possessing an academic degree earn a greater income than people who have an equivalent amount of studying without possessing an academic degree
- Jaeger and Page (1996): Find sizeable premia for high school, bachelor, but not Master or PhD.

6.3.3 *Curriculum Effect*

- Human Capital implies useful curriculum should matter - graduate earn more in jobs that use specific course trainings
  - Effect of additional courses is small for high school (Altonji, 1995)
  - Large math effects on earning in high school (Rose and Betts, 2001)
- Speed of completion of education:
  - Brodaty et al (2008) school openings in France where grade repetions is freueqnt (45% amongst boys). Impact of delay is negative on wages. A year of delay causes 9% decrease in wages

- College proximity as exogenous variation
  - Bedard (2001) looks at high school dropout rates as function of cost of attending college
  - University exists locally – low college attendance costs
- Testable implications: Signalling: higher high school dropout rates in regions contain a university (lower attending cost); higher skill pool among dropouts in regions contain a university
- Areas with higher university access have higher high school drop out rates 4-31% and college participation 10-15%

- Tyler et al (2000)
  - General Education Development certificate (GED)
  - By 1996 9.8% those aged 18-24 completed High School with GED versus HS diploma
  - In 1996 0.76m HS droupouts attempted GED and 0.5m passed
  - Monetar cost of taking GED is \$50 and average of 20h studying
- Simple comparisons with dropouts, GED, HS tells us nothing about causal effect of GED
  - GED would have earned less than HS regardless and more than dropouts
- Use variation in GED passing standards across states (Texas lower than NY, Florida, Oregon, Connecticut)
  - Quasi experiment effectively randomly assigns GED signal to some but not others
  - If we could determine who these marginal people are we can identify pure signalling effect
- Ideal estimate $T = E[Y_1|\eta = k] - E[Y_0|\eta = k]$
  - Where Y_1, Y_0 are earnings of GED and non-GED holders; η abililty and k GED score
- But also location effect. Could be controlled by comparing workers passing GED in both
  - $\hat{T} = T + \delta = E[Y_1|\eta = k, NY] - E[Y_0|\eta = k, TX]$

Large signalling effects for whites, estimated at 20% earnings gains after 5 years
GED is taken as positive signal eventually, but takes five years

- For GED to be positive signal it must (1) GED holders are on average more productive (2) GED more costly to obtain for less productive and (3) Employers are unable to perfectly distinguish between productivity directly

- Results do not say anything about education being unproductive
- Concern on self-selection: High ability people who were extremely unlucky on date of test will retake and pass it.
- 40-44 score groups in different states not directly comparable

- Low-passing (TX): students whose true abilitly is 40-44 + high abilitly but bad luck
- High passing (NY): students whose true abilitly is 40-44
- Only rely on scores from first attempt at GED

- Jespen et al (2016) do this and find no effect of GED on employment and earnings