14.472 Public Finance II

Redistribution: Tagging and Self-Targeting (V_b)

Amy Finkelstein

Fall 2022

Road Map - Redistribution (Unit V)

- Frameworks (theory): [done]
  - Basic welfare economics (Kaplow)
  - Marginal Value of Public Funds (Hendren)
- Instruments (theory) **[up next]**
  - Tagging (Akerlof)
  - Self-targeting (Nichols and Zeckhauser; Currie and Gahvari)
- Applications (with empirics):
  - Cash vs. In Kind Transfers: Why would we ever redistribute in-kind?
  - Low take-up of benefits: Is it "a problem"
  - Places vs People: Why would we ever redistribute based on place?

Tagging and Self-Targeting

- Central problem in public finance: social planner wants to redistribute (or insure) but has imperfect information about "ability" (or underlying attribute along which want to redistribute (or insure))
  - Concern that may transfer to people whom don't want to, and miss people whom do
    - e.g. is DI going to people who are truly disabled, cash transfers to people who truly have no productive employment etc
  - Concern about distorting incentives (e.g. distort labor supply if transfer based on earnings)
- Diamond-Mirlees optimal non-linear income tax problem (471)
  - Want to redistribute from high ability (high marginal product) to low ability (low marginal product)
  - Key challenge: ability (wage) not observed therefore distribute on the basis of income (wage*hours) which creates distortion in labor supply
- Tagging and targeting:
  - Can we improve on social planner's ability to insure or redistribute above and beyond the optimal non-linear income tax?

Optimal income tax in one slide

- Basic Mirlees (1971) model:
  - High and low ability types have different ability (marginal product / wage)
  - Goal of income tax is to redistribute from high to low ability
  - Ability is not observed
  - Income (= wage x hours) is observed
- Binding IC constraint on high ability type prevents first best redistribution
  - i.e. equalizing consumption across types (w utilitarian swf) not incentive compatible

Tagging and self-targeting

- Terminology:
  - Tagging: using observables to target transfers
  - Self-targeting (or screening): getting "right" individuals to self-select into transfers
- Both are attempts to combat / reduce moral hazard (weaken the binding IC constraint in the optimal income tax problem)
  - Up until now have simply asked: empirically how to estimate the mh costs of a social insurance program and weight those against benefits
  - Now want to ask: are there ways we can design programs to reduce moral hazard?
  - This brings us to: tagging and self-targeting

Tagging

- General vs Targeted Redistribution:
  - Negative income tax: general tax system that redistributes to poor
  - Targeted programs: choose an (identifiable) group to redistribute to
    - Tags may include age, health, family structure, residence etc.
- US has opted for targeted redistribution
  - More targeted allows you to spend less to reach the people you want
  - But may be more costly to administer and/or encourage adverse behavior

Akerlof Tagging Model

- Negative income tax:

$$T = -\alpha Y_{avg} + tY$$

where α is fraction of per capita avg income (Y_{avg}) received by a person with 0 gross income (i.e. minimum support); t is the marginal rate of taxation

- Summing over all individuals and dividing by total income gives:

$$t = \alpha + g$$

where g is the ratio of net taxes collected to total income

- Key points:
  - Tradeoff: higher levels of support (α) come at the cost of higher marginal tax rates (t)
  - Usual distortions: t decreases incentive for labor supply

Akerlof Tagging Model (con't)

- Suppose that we can identify (tag) a group of people that contains only the poor and this group contains only a fraction $\beta < 1$ of the population

Give the minimum support α to only this fraction, funded with same marginal tax rate t:

$$t = \beta\alpha + g$$

vs. general negative income tax:

$$t = \alpha + g$$

- Key point: tagging allows greater support for the poor with less distortion in the tax structure
  - for given amt of support α,t is lower with tagging

Akerlof tagging (cont'd)

- Benefits of tagging: lower tax rate for given amount of transfers to tagged group
- Potential costs of tagging:
  - Higher administrative costs
  - Potential inequity (what if poor but not in tag?)
  - Endogenous tags / Potential behavioral distortions
- Result in paper: if tagging is costless, should always do some redistribution based on tag
  - Intuition: envelope theorem. First amt of tagging generates only second order DWL from distortion in behavior, but first order transfer gain.
- NB: Quantitative (empirical) questions still remain
  - What is the optimal level of a tag?
    - Or (a la Baily!): on the margin should we increase or decrease use of this tag?
  - Another key empirical question: endogeneity of tag

Tagging (examples)

- Akerlof example: categorical welfare
  - i.e. Cash welfare to poor in female headed households
    - Lower marginal product (i.e. child care costs etc)
    - Endogeneity of tag?
- Disability insurance can also be rationalized / understood as a potential tag
  - Diamond-Sheskinski (1995)
- Place-based policies as a potential tag (Gaubert et al. 2020)

Disability Insurance as a Tag

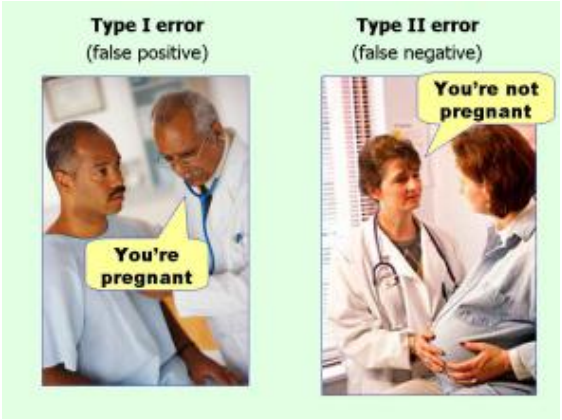
Diamond-Sheshinski (1995)

- People have different disutilities of work
- First best outcome: only work if marginal product of work exceeds disutility from work
  - Consumption fully insured across states (work / not work)
- Issue: don't directly observe "disutility of work"
- The disabled have higher disutility of work
  - Disability as a tag for high disutility of work / want to redistribute income to
- By adding disability insurance to existing income tax system can redistribute with less distortion (Akerlovian tag)
  - optimal disability insurance is non zero (envelope thm)
  - Again though, doesn't tell us what optimal system is or whether on margin should expand or reduce current DI benefits...

Diamond Sheshinski (con't)

- Take optimal social insurance level problem (tradeoff btwn insurance and incentives) and add an imperfect tag
- Key feature of their model: imperfect tag
  - Observed disability is an imperfect screen of true medical condition / disutility of work
  - Type I and Type II errors
  - The villagers in the boy who cried wolf

Don't Make a Type III Error



Diamond Sheshinski (con't)

- Take optimal social insurance level problem (tradeoff btwn insurance and incentives) and add an imperfect tag
- Key feature of their model: imperfect tag
  - Observed disability is an imperfect screen of true medical condition / disutility of work
  - Type I and Type II errors
- Government gets an imperfect signal of disutility of work
  - Standard result that larger benefits provide better insurance but with larger efficiency costs
  - Main new result: optimal insurance rate increasing in how good the screening device is
    - The worse the screening device, the lower the optimal insurance rate

Empirical question: how good a tag is disability

- Type I and Type II errors in screening process
- What empirical literature discusses:
  - Large empirical literature asking how DI affects labor supply
  - But how does this relate to optimal DI? Theory is about disutility of work among marginal enrollee

Despande and Lockwood 2022: Beyond Health

- Deshpande and Lockwood (EMA 2022): Beyond Health: Non-Health Risk and Value of Disability Insurance
- Subsantial academic and policy interest in the question: Are the individuals who get SSDI "not that disabled"?
  - i.e. How bad is their health really? How much would they work in absence of DI?
- But what we actually want to know: how useful is disability as a tag for marginal utility of consumption
  - e.g. looking at "less severe" disability recipients compared to "less severe" non recipients, the former are worse off on a number of consumption proxies (probabiltiy of foreclosure or eviction; consumption level)
  - more severe non recipients are better off on these dimensions than more or less severe recipients

Places vs. People

- A lot of place-based policies
  - eg investment and wage subsidies for firms who locate in poor areas
- Standard economic rationale for place based policies is agglomeration economies
- Generally considered a poor way to do redistribution
  - with perfectly mobile workers and inelastic housing supply, benefits of location-based subsidies capitalized into land rents (transfer to local landowners)
  - without perfect mobility, place-based subsidies can affect utility of inframarginal workers but these may or may not be the high marginal utility of consumption workers (seems indirect)
  - "Help poor people, not poor places"

Places as a Tag

- Gaubert, Kline and Yagan (2020) "Place-Based Redistribution"
- Key insight: place (distressed neighborhood) may be a tag for unobserved ability of individuals
- Empirically the key issues are:
  - how good a tag is it (how strong a signal)
  - how large is efficiency cost from migration response (endogeneity of tag)

Self-targeting

- Want to redistribute based on an unobserved characteristic (e.g. ability)
- Self-targeting insight: if a program design feature affects marginal utility differently based on ability, may be able to redistribute more for a given cost
  - Exploit single crossing feature: people of different ability have different marginal utility (disutility) from specific goods
- Example: in kind vs cash transfers
  - General economic view: cash dominates (allow people to optimize unconstrained).
  - But if demand for a specific good is decreasing in ability, in kind may be desirable
- Example: Ordeal mechanism:
  - If stigma, tedious administrative procedures etc imposes a higher disutility on higher ability individuals, may be desirable
- Implication: design of optimal second best transfer policy may involve sacrifice of productive efficiency

Nichols and Zeckhauser 1982

- Toy model illustrates potential role of in-kind transfers (vs cash) and ordeal mechanisms

In-Kind Transfers to Deter Imposters

•Two types: Intended recipients (B) and potential imposters (A). Type not observed.

•There is a pure income tax-transfer scheme in place in which if pre-tax income is restricted to a certain level, receive a cash transfer. Assume B receives transfer, A does not.

•Assume that optimal tax transfer scheme has not fully equalized mu of income (B's is still higher so would like to do more transfers but if so would violate IC constraint). Binding IC constraint: A indiff btwn pretending to be B and not…

• Given his transfer income, B chooses to purchase optimal amt X*(B) of good X.

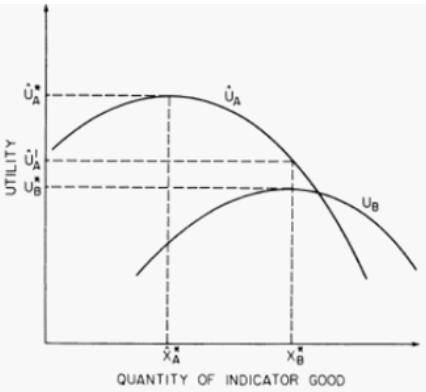


FIGURE 2. IN-KIND TRANSFERS TO DETER IMPOSTORS

• Were A to shirk and receive same income as B, he would buy only X*(A) worth of X.
• Figure shows A's utility as a function of X consumed *if he masquerades as B and gets the transfer intended for B.*

In-Kind Transfers to Deter Imposters

•Key point: when shirking and claiming to be B type, A's optimal consumption of X is less than B's

•Now imagine we convert part of the cash transfer to in kind provision of X

•Setting amt provided below X*(A) has no effect (relative to cash)

•AS we raise the amount provided above X*(A), A suffers increasing losses if he masquerades as B, and B suffers no loss so long as X< X*(B)

•So at a minimum would want to set amt of X provided at X*(B). Providing X*(B) in lieu of cash: B (intended recipient) is no worse off; A is no worse off if he doesn't masquerade. Moreover…

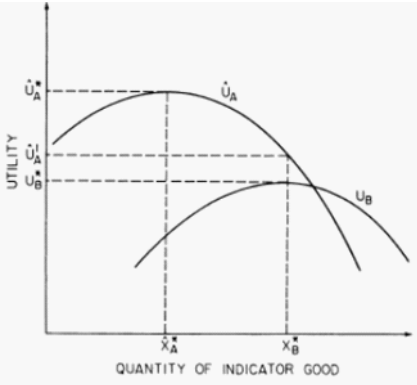


FIGURE 2. IN-KIND TRANSFERS TO DETER IMPOSTORS

If A does masquerade he has lower utility than with cash redistribution (see picture) thus creating opportunity to do more redistribution (beofre A was indifferent btwn masquerading and not, now strictly prefers not to)

In-Kind Transfers to Deter Imposters

•Have just argued that can increase redistribution (which wanted to do given binding IC constraint) by providing X*(B) in kind in lieu of cash

•*Key point:* In general, will be optimal to transfer an amt X larger than X*(B).

•Intuition: envelope thm: marginal increase in X above X*(B) has only second order welfare loss to B but first order welfare loss to A if masquerades.

•*Optimal in kind transfer scheme forces B to consume "too much" X*

•*Sacrificing productive efficiency to increase targeting efficiency!*

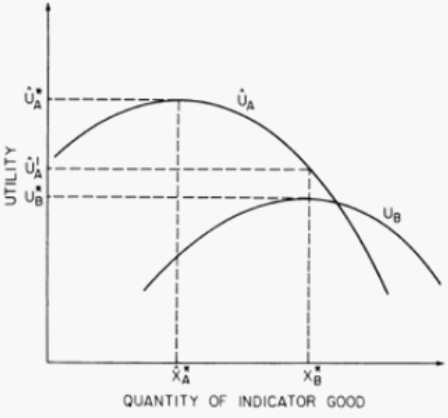


FIGURE 2. IN-KIND TRANSFERS TO DETER IMPOSTORS

Relationship to optimal income tax theory (471)

- In-kind transfers can weaken IC constraint and allow more redistribution if hurts the would-be mimicker (high ability) less than the mimicked (low ability)
  - i.e. if high and low ability want to consumer different levels of the in-kind good
- In-kind transfers can improve the efficiency of the income tax system via impacts on labor supply
- Relates to literature on benefits of commodity taxation in presence of optimal income tax (Currie and Gahvari 2008 JEL is nice overview)

Relationship to optimal income tax theory

- Atkinson-Stiglitz (1976): no role for in-kind transfers
  - assumes: preferences are weakly separable btwn labor supply and consumption goods, and identical for all consumer types (only heterogeneity is in skills)
  - Pareto efficient allocations (constrained by self selection) can be implemented through a non-linear income tax
  - Commodity taxes are not needed (and therefore in-kind transfers as well) in presence of optimal income tax
  - Key intuition: consumption taxes are redundant bc MRS between any two goods is same for the mimicker and the mimicked
- Saez (2002): rationale for in-kind transfers
  - Allow for heterogeneity in preferences across types
  - Then differential commodity taxes can be useful for redistribution if consumption patterns provide additional information about ability (correlated preference heterogeneity)

Ordeals

- Nichols and Zeckhauser analysis also suggests may be optimal to have "ordeals" in transfer programs: i.e. pure deadweight cost e.g.
  - Tedious administrative procedures; stigma; etc
- May enhance target efficiency if benefits from transfers vary across potential recipients
  - Suppose intended get 100 utils from transfer
  - Suppose imposters get 10 utils
  - Then ordeal that imposes an 11 util loss in order to qualify for the transfer would be an effective screening device
- Example: make people on Medicaid (which pays for long term care) get care in nursing home rather than in home
  - People tend to prefer home care & nh more expensive
  - Nevertheless, may be a good screen for those who would buy private insurance in absence of public program...
- Will return to and consider some opposing theories and empirical evidence when we get to take-up

Behavioral Economics Take on Ordeals

- Ordeals may have exactly the opposite targeting effect as that conjecuted by neo-classical theory (e.g. Nichols and Zeckhauser 1982)
  - screen out precisely those applicants the social planner would most likely enroll
  - e.g. poverty imposed "bandwidth" tax on poor individuals, making them less likely to undertake high net-value activities like enrolling in transfer programs for which they are eligible (Mullainathan and Shafir, 2013)
- This raises two questions:
  - Empirically: who is screened out by ordeals?
  - Conceptually: how do the self-targeting properties of the ordeals relate to its welfare implications?

NZ and Empirical Opportunities

- Theory:
  - Nichols and Zeckhauser vs. "Behavioral Economics"
- Empirical question: are screens screening out the "right" people?
  - Application I: In-kind vs cash transfers
  - Application II: "Ordeals" / Take-up of benefits

Road Map - Redistribution (Unit V)

- Frameworks (theory): [done]
  - Basic welfare economics (Kaplow)
  - Marginal Value of Public Funds (Hendren)
- Instruments (theory) [done]
  - Tagging (Akerlof)
  - Self-targeting (Nichols and Zeckhauser; Currie and Gahvari)
- Applications (with empirics): **[up next]**
  - Cash vs. In Kind Transfers: Why would we ever redistribute in-kind?
  - Low take-up of benefits: Is it "a problem"