Public Economics

Pre-Exam Summary 2019–2020

Artur Obminski

December 30, 2019

You can find the latest version of this document here.

1 Introduction to Public Economics (AB)

1.1 Background

- Adam Smith: Canons of taxation
  - Equality
  - Certainty
  - Convenience of payment
  - Economy of collection
- Musgrave: 3 branches of government
  - Resource allocation to address market failures
  - Income redistribution
  - Macroeconomic stabilization
- Sufficient statistics: Theory used to derive formulas based on empirical estimates

1.2 Public spending and taxation

- Growth of the state
  - 19th century: Public spending $\sim$ 10% of GDP, social largely absent
  - 20th century: Double state size in 1920s, large increases during world wars, acceleration 1960s–1980s
  - Now: Public spending $\sim$ 45% of GDP
    - U.S. 40%, France 55%
  - Higher taxes (and public debt) during war times
  - Social spending 0–5% in 1930, 20–32% now
  - Differences in welfare states
    - Social insurance v.s. means tested benefits
    - Differences in public social insurance (esp. pension) spending
  - Other aspects of public interventions: State ownership, political rights, labor law, monetary regimes
- Theories of state growth
  - Wagner's law: Demand for public goods grows with income (elasticity > 1)
  - Baumol's cost disease: Public services labor intensive, so cost to provide them will increase faster than prices
  - Ratchet effect theory: Wars increase government spending and taxation, not reversed after end of war
  - Leviathan theory: Governments controlled by self-interested politician-bureaucrats
  - Public economy: Public spending changes matches changes in those with voice (democratization $\rightarrow$ demand for redistribution)
  - Technology and enforcement: Development of firm accountancy, computerization $\rightarrow$ third party reporting
- Increasing focus on social protection and health in government spending

- *Tax*: "Compulsory unrequited payments to general government"
  - Not user fee (unrequited)
- *5 main components of government revenue*:
  - Personal income tax
  - Corporate income tax
  - Social security contributions
  - Consumption taxes
  - Property taxes
  - ⋆ Since 1960s, declining consumption and property; increasing social security
- Taxation rising share of government revenue and of GDP ($\sim$ 25% in 1960s $\rightarrow\sim$ 35% now
  - Dispersion since 1970s–1980s: US, Japan 25%; Germany, UK 35%; France, Nordics 45%
  - Differences in taxation on transfers, government expenditure; social insurance schemes

1.3 Normative theories of social justice

- Welfarism
  - Social welfare depends only on individual's utility $u_i(x)$
    - ∗ No concern for distribution
  - $SW(x) = W(u_1(x), \ldots, u_n(x)$
  - Utilitarianism
    - ∗ Bentham: Greatest good for the greatest number of people
    - ∗ J.S. Mill: Protecting individual freedom, distinguishing higher and lower pleasures
  - Concerns
    - ∗ Maximizing sum or per capita
    - ∗ No concern for human rights, minority rights
    - ∗ Measuring (non-monetary aspects of) utility difficult
    - ∗ Application of market solutions to a wider range of moral problems, e.g. skipping queues
    - ∗ Transactions not always voluntary, e.g. poverty inducing them
    - ∗ Risk of corruption
    - ∗ Crowding out of social norms
- Egalitarianism
  - Tobin's specific egalitarianism: Some goods necessary for life and citizenship should be provided with strict equality (e.g. justice, vote), or with guaranteed minimum (e.g. education, health)
  - Rights-based approach: Explaining rise of welfare state
  - Absolute limits to inequality: Avoiding civil disintegration
- Libertarianism
  - Right to do what we want with what we own (labor, wealth, body), provided we respect other people's right
  - Justice required in process generating income distribution, including initial allocation

  - Night-watchman state: Minimal state, limited to the narrow functions of protection against force, theft, fraud, enforcement of contracts
  - Reparation if inherited wealth result of past injustice
  - Policy by unanimous consent only — only Pareto improvements
- Rawlsian approach
  - Veil of ignorance: Need to make social choice free of current status (money, power, ability)
  - 2 principles:
    * Equal right for all to most extensive basic liberty compatible with similar liberty for others — non-welfarist
    * Maximin: Social inequalities arranged to benefit least advantaged in society
  - Rejecting meritocracy, because meritocracy still based on initial distribution of e.g. ability
- Sen's capabilities approach
  - Functionings of human beings: Nourishment, shelter, physical mobility, ability to take part in the life of the community
  - Capabilities: Substantive freedom to achieve functionings
  - Non-welfarist, sometimes making people worse off
- Social Welfare Functions (SWF)
  - Utilitarian (Benthamite): $SWF(x) = \sum_i^n u_i(x)$
  - Maximin: $SWF(x) = \min_i u_i(x)$
  - General: $SWF(x) = \sum_i^n V(u_i(x))$
  - General with preference for equality ε: $SWF(x) = \frac{1}{1-\varepsilon}\frac{1}{n}\sum_i^n u_i(x)^{1-\varepsilon}$ with $\varepsilon \neq 1$
    * Utilitarian: $\varepsilon = 0$
    * Rawlsian: $\varepsilon = +\infty$
  - Individual expected utility with risk aversion: $u(c) = \frac{1}{1-\rho}c^{1-\rho}, \rho \neq 1$
    * Without incentive costs, $\varepsilon > 0$, total equality optimal, otherwise not
  - Utilitarian and Rawlsian SWF Pareto compatible

1.4 Rationales for government interventions

- 2 basic criteria in welfare analysis: Efficiency and equity
- *First theorem*: Any competitive equilibrium is Pareto efficient if:
  - No externalities or public goods
  - Perfect information
  - Perfect competition
  - Rational individuals
- *Second theorem*: Any efficient allocation can be achieved as a competitive equilibrium
  - Requires lump sum transfers, which are generally not possible
- Lump sum taxes: Fixed in amount, no action can reduce their burden
  - Rare, because intrinsic characteristics (e.g. ability) less observable

than e.g. income, and possible lump-sum taxes are usually unfair

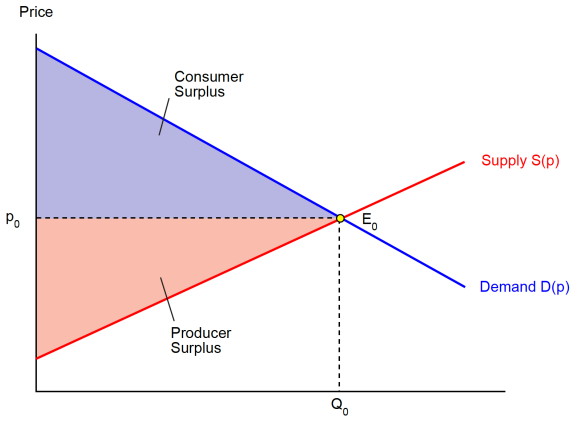
- Non-lump sum taxes distortionary second best

- → Rationales for government intervention:
  - Enforcing contracts, property rights
    - Secure property rights and GDP/capita correlated
  - Externalities → e.g. public good provision
    - Free riding → underprovision of e.g. defense
  - Imperfect or asymmetric information → e.g public insurance
    - Market for "lemons", adverse selection in health insurance, credit constraints in education market
  - Imperfect competition → e.g. regulation
    - Natural monopolies, collusion
  - Agents are not rational → e.g. mandated savings
    - Hyperbolic discounting, overconfidence, inattention
- Even with efficient private market allocation, government can improve equity of distribution
- Limitations of government intervention:
  - Collective choice problems: Aggregating individual preferences
  - Commitment problems: Policies may not be perceived as credible
  - Information constraints

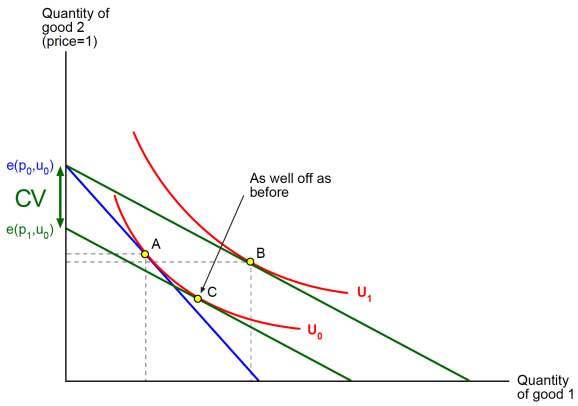
2 Tools of Welfare Analysis (JG)

2.1 Economic Surplus

- Economic surplus: Measure of the amount by which buyers and sellers benefit from participating in the market

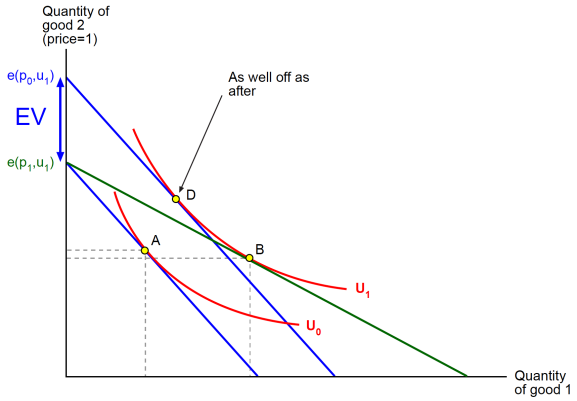


- Consumer surplus: Benefit consumers derive from consuming a good above and beyond what they pay for the good
  - Approximate measure of consumer welfare
  - Requires identifying the entire demand curve, which is empirically challenging
  - Coincides with Compensating Variation (CV) and Equivalent Variation (EV) with quasilinear utility (no income effects) $U(x,y) = v(x) + y; v' > 0, v'' < 0, v(0) = 0$
  - $\max_{x,y} v(x) + y$ s.t. $px + y = m \rightarrow$ FOC: $v'(x) = p$ = Marginal Willingness to Pay (MWTP)
    - Interpretation valid because of quasilinear utility
  - Marshallian demand: $x = D(p)$
  - Willingness To Pay (WTP): Amount of income that the consumer would be willing to sacrifice to enjoy the units of good X instead of zero units
    - $U(x_0, m - WTP) = U(0, m)$
    - $\Leftrightarrow v(x_0) + m - WTP = v(0) + m$
    - $\Leftrightarrow WTP = v(x_0) - v(0) = \int_0^{x_0} v'(x)\mathrm{d}x$
  - Consumer Expenditure (CE) = $p_0 x_0$
  - Consumer Surplus (CS) = WTP - CE = $v(x_0) - p_0 x_0 = \int_0^{x_0} v'(x)\mathrm{d}x - p_0 x_0$
  - Change in consumer surplus $p_0 \rightarrow p_1; p_0 > p_1$: $\Delta CS = \int_{p_1}^{p_1} D(p)\mathrm{d}p$

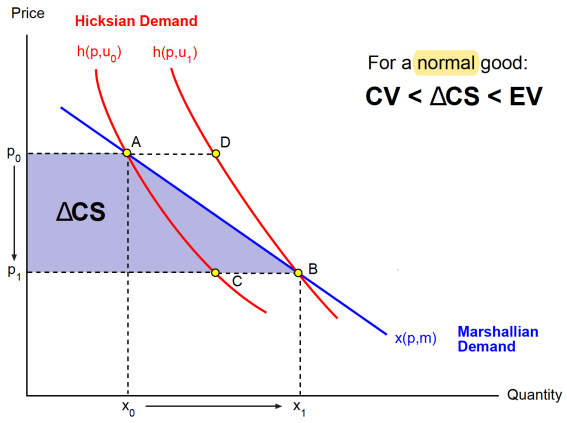


- *Compensating Variation (CV)*: Amount of money taken away from consumer after change to restore original utility level
  - Old utility, new prices
  - $CV = e(p_0, u_0) - e(p_1, u_0) = \int_{p_1}^{p_0} h(p, u_0)\,\mathrm{d}p$
  - $e(p,u)$: Expenditure function

– $h(p, u)$: Hicksian compensated demand



- *Equivalent Variation (EV)*: Amount of money given to consumer before change to leave her as well off as with change
  - New utility, old prices
  - $EV = e(p_0, u_1) - e(p_1, u_1) = \int_{p_1}^{p_0} h(p, u_1)\, dp$
- Comparing estimates of consumer welfare
  - Because utility not observable, CV and EV difficult to measure empirically
  - For normal good and price fall, $CV < \Delta CS < EV$
  - ΔCS exact measure of welfare change only with quasilinear utility, so that $\Delta CS = CV = EV$
  - Changes in CS good approximation of welfare changes if income effects negligible or budget share of considered good is small

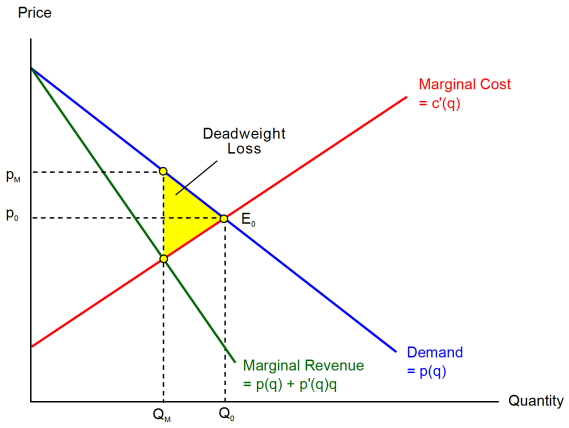


- Producer surplus: Benefit derived by producers from the sale of a unit above and beyond the cost of producing that unit
  - FOC: $p = c'(x) \rightarrow$ Supply curve: $x^S = S(p)$
  - Inverse supply (MC) curve: $p(x) = c'(x)$
  - Willingness To Sell (WTS): Minimum amount of money required to produce given quantity good, = sum of marginal costs of production (total variable costs)
    - $WTS = \int_0^{x_0} c'(x)\mathrm{d}x = TVC$
  - $PS = TR - TVC = p_0 x_0 - \int_0^{x_0} c'(x)\mathrm{d}x$
  - Producer surplus = Profits + Fixed costs
  - Change in producer surplus $p_0 \rightarrow p_1; p_0 < p_1$: $\Delta PS = \Delta\Pi = \int_{p_0}^{p_1} S(p)dp$

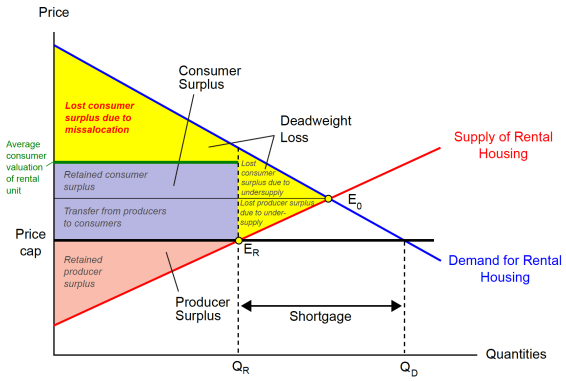
2.2 Measuring inefficiency: Deadweight loss

- First Fundamental Theorem of Welfare: The competitive equilibrium maximizes social efficiency

- Inefficiency can come from market imperfections or government intervention
  - e.g. Monopolist: $\max_q p(q)q - c(q)$
  - FOC: $MR = p(q) + p'q)q = c'(q) = MC$



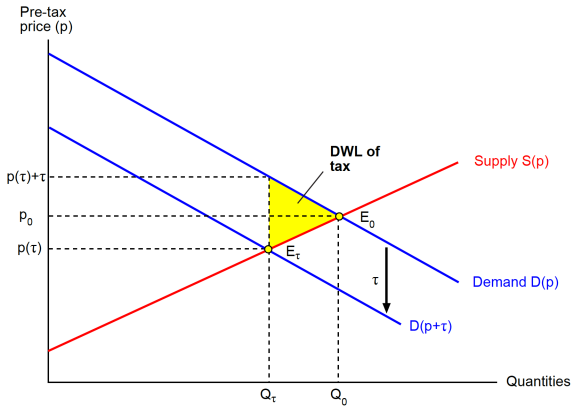
- - e.g. Rent control
    - Undersupply and rationing
    - **and** Misallocation between consumers → Additional DWL



- - Glaeser and Luttmer (2003)
    - Comparing consumption patterns in rent-controlled city (NYC) and free-market cities across demographic groups
    - Predict apartment size using demographics, in cities without rent control, testing if apartment allocations in NYC match predictions

2.3 Efficiency cost of taxation

- 2 reasons for taxation:
  - Financing government expenditure
  - Redistribution
- Partial equilibrium analysis with two goods ok approximation if small market taxed and no close substitutes/complements
- *Excise tax*: Levied on quantity
- *Ad-valorem tax*: Fraction of price
- Reduction in quantity at equilibrium due to tax → Fall in CS and PS, not fully compensated by tax revenues



- *Harberger formula*: $DWL = -\frac{1}{2}\mathrm{d}Q \times \mathrm{d}\tau = -\frac{1}{2}\left(\frac{\varepsilon_S \cdot \varepsilon_D}{\varepsilon_S - \varepsilon_D}\right) pQ \left(\frac{\mathrm{d}\tau}{p}\right)^2$
  - $= \frac{1}{2}\eta_Q(pQ)\left(\frac{\mathrm{d}\tau}{p}\right)^2$ ...where $\eta_Q = -\frac{\mathrm{d}Q}{\mathrm{d}\tau}\frac{p_0}{Q}$
  - Regression equation: $\log Q = \alpha + \beta\frac{\tau}{p_0} + \epsilon$
  - DWL proportional to market size (Q)

  - DWL increases with elasticities → Tax inelastic goods
  - Marginal DWL increases with tax rate τ → Keep broad tax base

- Marginal DWL
  - $\frac{\mathrm{d}DWL}{\mathrm{d}\tau} \approx \eta_Q Q \frac{\tau}{p}$
  - Marginal change in government revenue: $\frac{\mathrm{d}R}{\mathrm{d}\tau} = \frac{d(Q\cdot\tau)}{d\tau} = Q + \tau\frac{\mathrm{d}Q}{\mathrm{d}\tau} = Q + Q\eta_Q\frac{\tau}{p}$
  - $\frac{\mathrm{d}DWL/\mathrm{d}\tau}{\mathrm{d}R/\mathrm{d}\tau} \approx \frac{\eta_Q\frac{\tau}{p}}{1-\eta_Q\frac{\tau}{p}}$
- Marion and Muehlegger (2008): Deadweight loss from taxing diesel fuels
  - Diesel fuel for business purposes (e.g. trucking) taxed, not for residential purposes (e.g. heating homes)
  - 1993 reform: Red dye to residential fuel, decreasing evasion
  - Tax elasticity much higher than price elasticity before reform → Tax evasion likely
  - MDWL = 68 cents/dollar of net tax revenue before reform, 42 cents/-dollar after reform

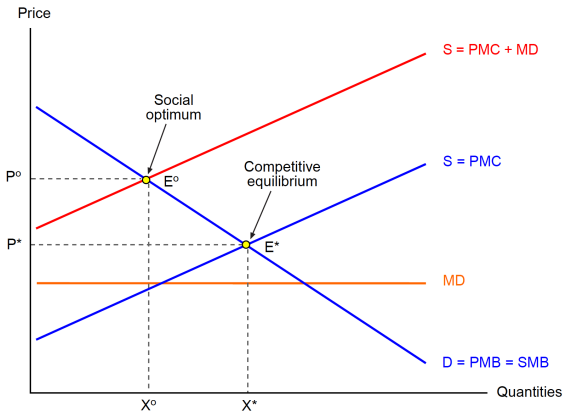
2.4 Tax incidence

- *Legal liability*: What the law says about who should pay the tax (also called statutory or formal incidence)
- *Remittance responsibility*: Who is responsible for remitting tax to authorities
- *Economic/effective incidence*: Who actually bears the tax burden, i.e. who loses utility
- Kotlikoff and Summers (1987) **incidence formula** (on producers): $\frac{\mathrm{d}p}{\mathrm{d}\tau} = \frac{\varepsilon_D}{\varepsilon_S-\varepsilon_D}$
  - Incidence on consumers: $\frac{\mathrm{d}q}{\mathrm{d}\tau} = \frac{\mathrm{d}(p+\tau)}{\mathrm{d}\tau} = 1 + \frac{\mathrm{d}p}{\mathrm{d}\tau} = \frac{\varepsilon_S}{\varepsilon_S-\varepsilon_D}$
  - Legal incidence does not describe economic incidence
  - Legal liability is irrelevant to distribution of the tax burdens
  - More inelastic factor bears more of the tax
- Carbonnier (2007): French VAT reforms
  - Consumer share 57% for car sales, 77% for housing repair services

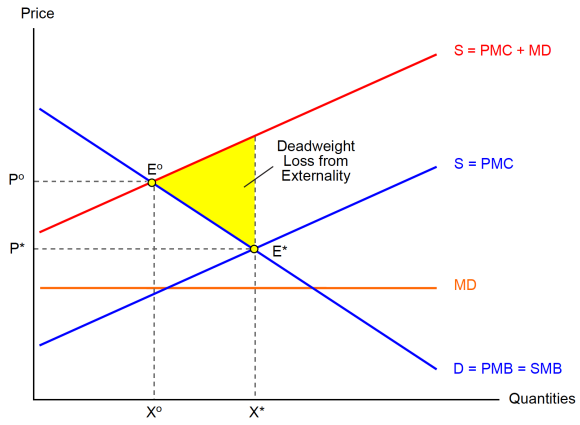
3 Externalities (JG)

3.1 Introduction

- Externality: Whenever the actions of one party directly makes another party worse or better off, yet the first party neither bears the costs nor receives the benefits
  - Directly: Not through prices/market mechanism
  - *Pecuniary externality*: Accounted for in prices, so not true externalities and no need for policy response
- Externalities represent market failures → Public intervention can improve welfare
- First best welfare function to maximize: $W(X) = V(X) + M - C(X) - \mu X$, where μ is marginal damage (MD), e.g. of pollution
  - Same as without numeraire M, $S(X) = V(X) - C(X) - \mu X$ = Aggregate Marshallian surplus, minus pollution cost
  - $X^o = \arg\max_X \{V(X) - C(X) - \mu X\}$ → FOC: $V'(X^o) = C'(X^o) + \mu$
  - Social Marginal Benefit (SMB) = Social Marginal Cost (SMC)
  - Private Marginal Benefit (PMB) = Private Marginal Cost (PMB) + Marginal Damage (MB)
- Competitive equilibrium
  - Firm not internalizing MD: $\Pi = pX - C(X) \rightarrow X^S = S(p)$
  - Consumer taking level of externality as given: $P(X) = \overline{P} \rightarrow U(X, N) = V(X) + N - \mu\overline{P} \rightarrow X^D = D(p)$
  - $p^* = V'(X^*) = C'(X^*) \Leftrightarrow PMB = PMC$
  - Social optimum: $V'(X^*) = C'(X^*) + d$ → Overproduction in competitive equilibrium



- Pertubation argument: Reducing production by $\mathrm{d}W < 0$ at competitive equilibrium
  - $\mathrm{d}W = V'(X^*)\mathrm{d}X - C'(X^*)\mathrm{d}X - \mu\mathrm{d}X = -\mu\mathrm{d}X > 0$
  - First Welfare Theorem violated, because welfare can be improved from moving away from competitive equilibrium



3.2 Measuring externalities

- No direct market to recover WTP to reduce negative externalities / increase positive externalities, because if existed, no longer externalities

- 3 valuation methods:
  - *Direct valuation*: Measure directly physical effects of externalities (e.g. health damages), use market prices (e.g. medical expenses) to estimate monetary value
    - Problem: Individuals may reduce exposure through costly *avoidance behavior*
    - Moretti and Neidell (2011) (ozone pollution on health; Los Angeles, USA; IV using boat arrivals as instrument for pollution): IV results 4x OLS, for respiratory illnesses; Avoidance costs 1/4 of hospitalization costs
    - Direct costs difficult to estimate
      - All channels may not be identified
      - All market prices may not be available
  - *Contingent valuation*: Ask people directly about willingness-to-pay (WTP) (stated preferences)
    - Problems: Survey costs, framing effects (e.g. order of questions), embedding effects (e.g. different magnitudes), strategic responses
  - *Hedonic valuation*: Decompose good (e.g. house) into characteristics (e.g. size, amenities), estimate contribution of each characteristic to overall value (revealed preferences)
    - Typically regression model, including e.g. pollution, controlling for other observable characteristics
    - Problems: Omitted variable bias, selection bias (sorting based in MWTP)
    - Alternative: First differences, but still potential OVB (e.g. recession affecting both housing prices and pollution)
  - Chay and Greenstone (2005) (DiD, US counties, Clean Air Act pollution ceilings, nonattainment instrument for reduction): -10% pollution $\rightarrow$ +2.5% housing prices
    - But maybe incomplete awareness
  - Linden and Rockoff (2008) (North Carolina, sex offender registry, offender moving in instrument for

crime risk): Offender within 160 m → -4% housing prices

* But maybe overestimation of risks, and why effect only within 160 m radius?

3.3 Correcting externalities

- Coasian bargaining (private): Establish property rights to create markets for pollution

  1 With well-defined property rights and costless bargaining, negotiations between parties creating/affected by externality can bring about social optimum

  2 Efficient solution does not depend on how property rights are assigned

  * Assignment of property rights affects distribution, not efficiency

  – Problems:

  * Assignment: Difficult to assign responsibility, value marginal damage
  * Holdout: Shared property rights give each party power over all others, which could lead to breakdown in negotiations
    · e.g. Last person asking for much more compensation for deal to pass
  * Free rider: When investment has personal cost but common benefit, individuals underinvest
  * Transaction costs and negotiating: Hard to negotiate when large numbers of individuals on one or both sides → Government can represent them

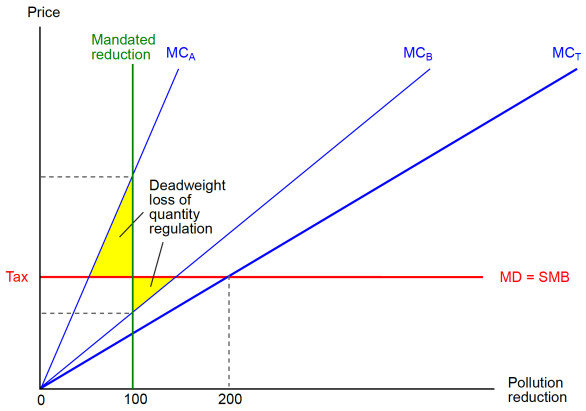
- 2 types of public solutions:
  – *Command-and-control (CaC)*: Set standards on pollution level, controls to enforce standard
    * Restrictions, mandates
  – *Market based instruments*: Rely on market/price mechanisms to give incentives to reduce emissions
    * Pigouvian taxes/subsidies, tradable permits (cap-and-trade)
    * Superior to CaC

- 2 types of regulation:
  – Price regulation: Pigouvian taxes/-subsidies
    * Government can impose a tax equal to the marginal damage $t = MD = d$: $\max_X \Pi = p \times X - C(X) - t \times X \rightarrow$ Shifts PMC to social optimum
  – Quantity regulation: Command-and-control, cap-and-trade

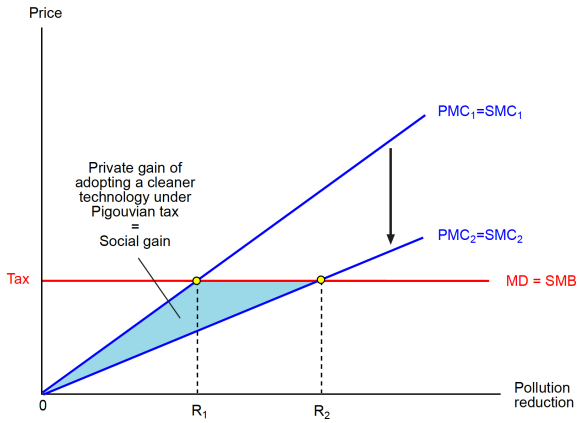
3.4 Choice of instruments

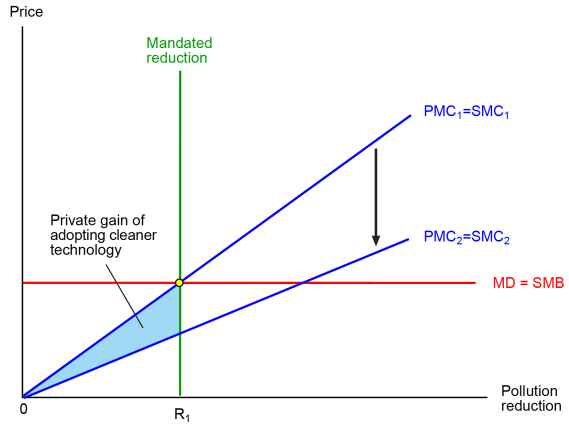
- With homogeneous firms (in terms of reduction costs) and perfect information, instruments equally efficient
- Socially optimal pollution reduction: $C'(R^o) = B'(R^o)$ (MC = MB)
  – Can be achieved by mandating R^o or setting tax t equal to MD μ

- Because firms have different costs of pollution reduction, market-based instruments more efficient than CaC
  - Setting tax t equal to MD μ, pollution reduction costs equalized across firms

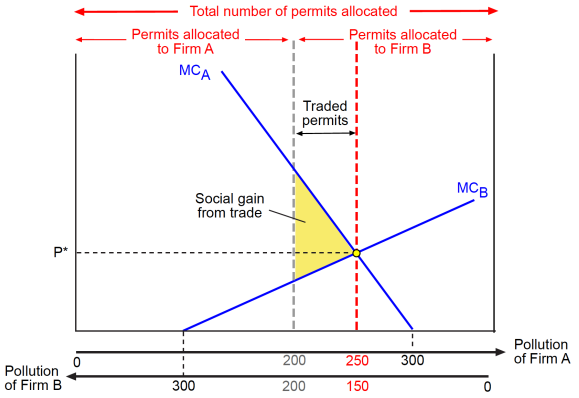


- In addition to allocative efficiency, also dynamic efficiency of Pigouvian tax
  - Private gain for firm of adopting cleaner technology equal to social gain; while with CaC, private gain lower than potential social gain

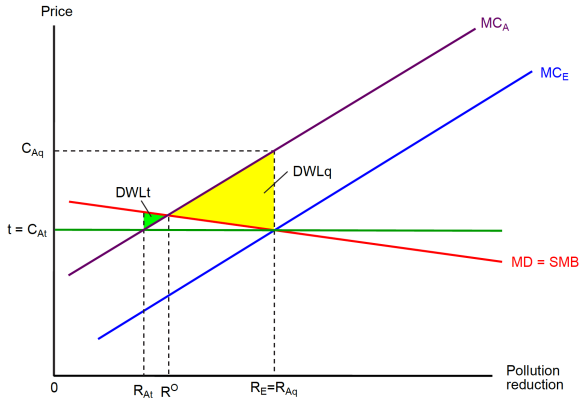


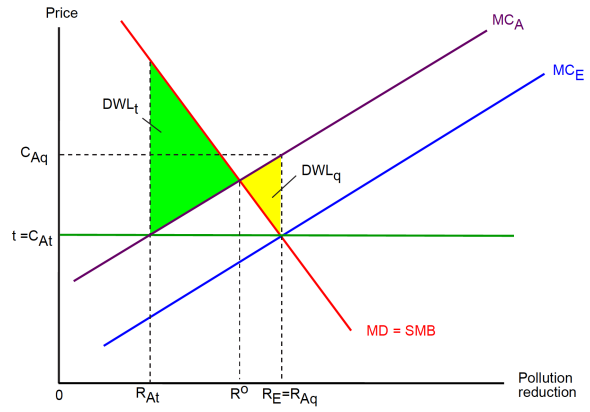


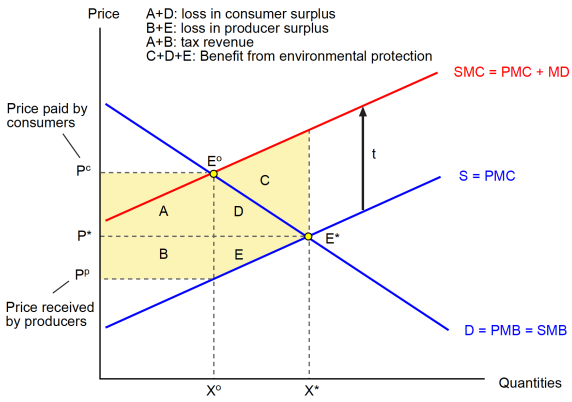
- With *tradable permits*, marginal costs of pollution reduction equalized across firms $P^* = MC_A = MC_B \rightarrow$ Total cost of pollution reduction minimized
  - Same allocative and dynamic efficiency properties as Pigouvian tax



- Because aggregate cost of pollution reduction uncertain, relative efficiency of price vs. quantity regulation depends on steepness of the marginal damage curve
  - MD curve flat (e.g., global warming) $\rightarrow$ More concern for potential cost overruns $\rightarrow$ Tax more efficient
  - MD curve steep (e.g., nuclear leakage) $\rightarrow$ Important to get quantity right $\rightarrow$ Quantity regulation more efficient





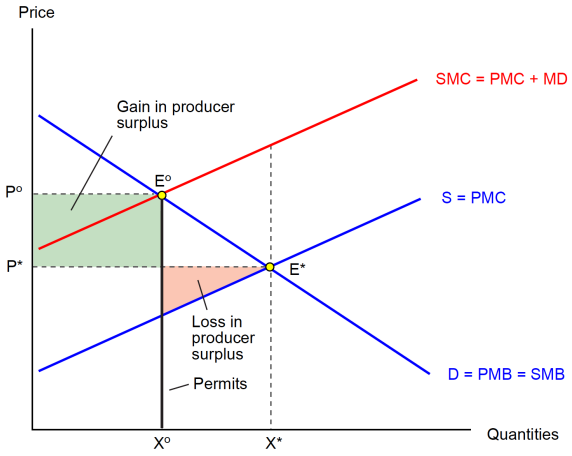


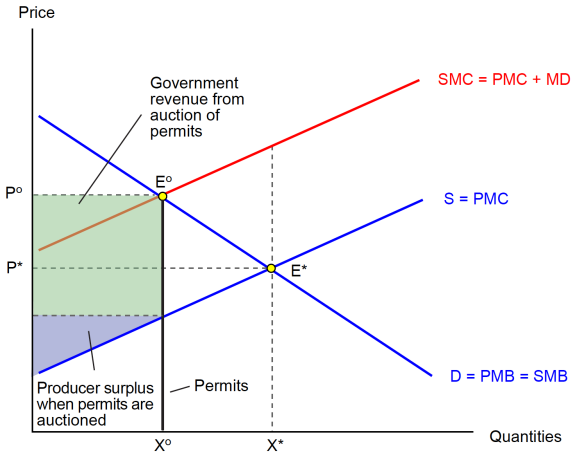
3.5 Distributional concerns

- Some households and sectors might be more affected than others → Effects on political process

- Impact of Pigouvian carbon tax
  - Demand less elastic than supply → Most of tax burden passed onto consumers through higher prices
  - Low-income individuals spend larger share of income on energy → Without redistribution, often regressive

- Impact of tradable permits
  - Auctioning → Government revenue, same distributional consequences as Pigouvian taxes, possibility of redistributing revenues
  - Grandfathering quota allowances → Big rents to firms, overcompensation





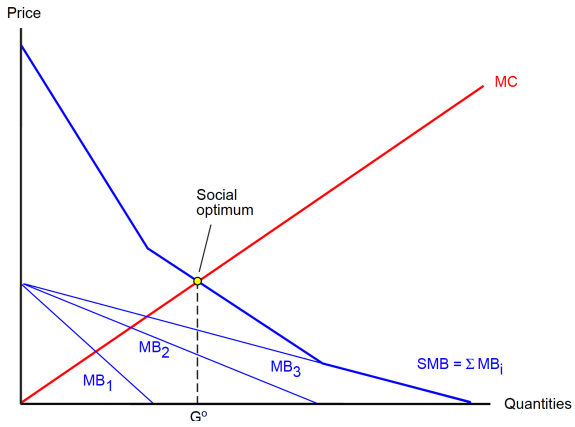
3.6 Case: Global warming

- Chancel and Piketty (2015): Defining international climate justice difficult
  - Major conflicts about country responsibilities: Most countries contribute to climate mitigation efforts, but financing of climate adaptation fund heavily concentrated among a few
  - Unequal historical responsibilities for climate change, unequal ability to pay for mitigation
  - Problems: Emerging countries account for 1/3+ of cumulated historical CO_2 emissions; importance of within-country inequalities
  - Proposed model:
    * Production-based → Consumption-based measure
    * Country-level → Individual-level contributions, through progressive carbon tax
- Recent move CaC → Carbon taxes, cap-and-trade
- Carbon taxation: Current rates $\sim$ € 10–25/ton CO_2, but limited impact
  - Numerous exemptions
  - Tax rates across fuels not reflecting emissions arising from their use
  - No harmonization across countries → "carbon leakage"
- Tradable permits: Europan Union Emission Trading Scheme (ETS)
  - Initially, vast majority of permits distributed freely to firms (overallocation), inability to save permits across phases (banking) → Low permit prices, volatility
  - Grandfathering → Overcompensating polluting firms, punishing early pollution reducing action
  - Initially, coverage only for $\sim$ 50% of emissions
  - Now, movement towards auctioning, prices increasing

4 Public Goods (JG)

4.1 Optimal provision: Samuelson rule

- Pure public goods are non-rival and non-excludable
- Free rider problem → Inefficient private (under-)provision → Justification for public intervention
- If good G is *private* (assuming quasi-linear utility):
  - $\max_{G_1,G_2,\ldots,G_N} \sum_{i=1}^{N} V_i(G_i) - C\left(\sum_{i=1}^{N} G_i\right)$
  - FOC: $V_i'(G_i^o) = C'(G^o) \quad \forall i \Leftrightarrow \text{MB}_i = \text{MC}\ \forall i$
  - Optimal quantity where *horizontal sum* of MB cross MC
  - Decentralized equilibrium reaches this equilibrium (first welfare theorem)
    - $\sum_i D_i(p^*) = S(p^*)$
    - $p^* = V_i'(G_i^*) = C_i'(G_i^*) \quad \forall i$
    - $\text{MB}_i = \text{MC}$
    - Consumers pay same per-unit price, but consume different quantities
- If good G is *public*:
  - Each consumer consumes *total* quantity supplied G
  - $\max_G \sum_{i=1}^{N} V_i(G) - C(G)$
  - Interior solution FOC: $\max_G \sum_{i=1}^{N} V_i'(G^o) = C'(G^o) \Leftrightarrow \sum_i \text{MB}_i = \text{MC}$
  - Optimal quantity where *vertical sum* of MB cross MC



- *Samuelson Rule*: Sum of marginal benefits should be equal to the marginal cost of producing the public good, $\sum_i \text{MB}_i = \text{MC}$
  - Additional unit increases utility of all individuals
  - Excludability not analyzed, only rivalness

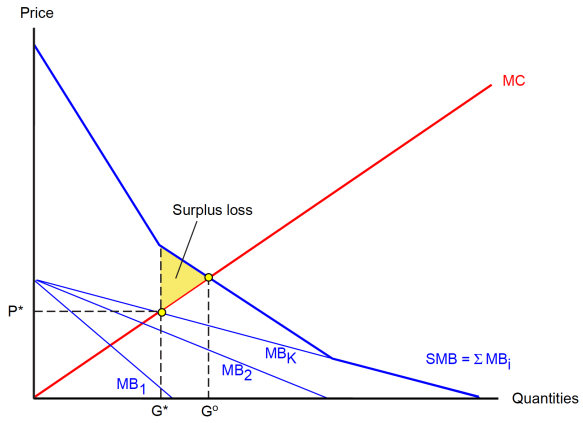
4.2 Private provision of public goods

- In private market, consumers i are price takers and decide on quantity bought G_i from price taking profit maximizing provider, taking other consumers' purchases $\overline{G_{-i}} = \sum_{i \neq j} G_j$ as given
- Consumer i solves $\max_{G_i} V_i(G_i + \overline{G_{-i}} + M_i - p \times G_i \text{ s.t. } G_i \geq 0$
- Kuhn-Tucker FOC: $V_i'(G_i^* + \overline{G_{-i}}) \leq p$ and $G_i^* \left[\left(V_i'\left(G_i^* + \overline{G_{-i}}\right) - p\right] = 0 \quad \forall i$
  - Consumer contributes as long as MB below MC of contribution
- If consumers have different MB_i, $V_i'(G_i^* + \overline{G_{-i}^*}) \leq p$ cannot hold with equality

$(G_i^* > 0)$ for all individuals simultaneously

  - Only holds with equality for consumer K with highest MB, all others free ride

- Equilibrium: $V_K'(G^*) = p^* = C'(G^*) \Rightarrow \sum_i V_i'(G^*) = V_K'(G^*) + \sum_{i \neq K} V_i'(G^*) > C'(G^*) \Rightarrow \sum_i MB_i > MC \Rightarrow$ Underprovision $(G^* < G^o)$



- Where private provision can overcome free riding:
  - Intense preferences: When some individuals have especially high demand for a public good, e.g. USA with NATO in 1960s
  - Altruism: When individuals privately value the benefits and costs of others
  - Warm glow: When individuals care both about the total amount of the public good and about their own contributions

4.3 Empirical evidence on free riding

- Marwell and Ames (1981): Lab experiment, sharing game
  - Nash equilibrium 100% individual, Pareto optimum 100% group
  - Result: 40–60% group, only economics graduate students free ride substantially
- Isaac, McCue and Plott (1985): Increased free-riding over time with repeated interactions
- Andreoni (1988): Partners free-ride more than strangers, but contribute again after restart
  - Non-monetary pleasure from cooperative outcomes? Attempt to enforce social norms about participation?
- Fehr and Gächter (2000): Existence of (costly) punishment opportunities causes large rise in average contribution, and no decrease over time
  - Herrmann, Thöni, and Gächter: Cross-societal differences, anti-social free riders punishing back
- Miguel and Gugerty (2005) (Western Kenya): Ethnic diversity related to lower levels of voluntary contributions to school funding, water well maintenance, fewer sanctions on defaulting parents

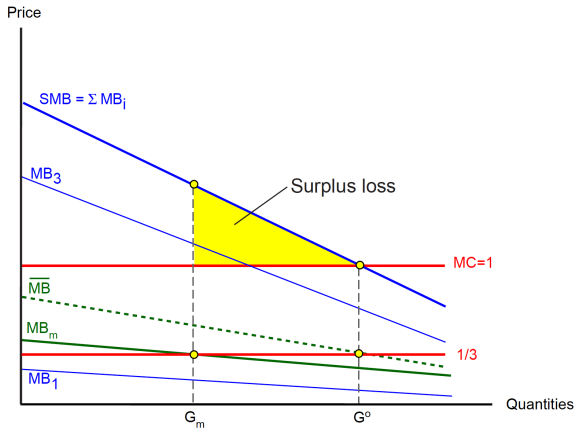
4.4 Public provision of public goods

- 3 main issues of public intervention in public good provision:
  - Determining optimal supply with imperfect information on preferences
    - *Lindahl pricing*: Individuals should pay different prices,,

equal to their own marginal benefit

- Individuals' FOC: $V_i'(G) = \tau_i$
- Share of financing: $G^i = f(\tau_i)$
- Lindahl equilibrium: (1) Fully financed $\sum_i \tau_i = 1$, and (2) All individuals agree on quantity G^* — generally exists
- Satisfies Samuelson rule, Pareto efficiency: $\sum_i V_i'(G^*) = \sum_i \tau_i^* = 1$

* Strategic incentive to underreport preferred amount of good, mechanisms to induce truthful revelation impractical and costly
* → Majority voting
  - *Condorcet Paradox*: Majority voting does not always lead to stable outcome — cycle in pairwise social ordering
  - *Arrow's Impossibility Theorem*: The only social choice rule that (a) always produces a winner, (b) satisfies Pareto condition (unanimity), and (c) satisfies Independence of Irrelevant Alternatives condition is dictatorship
  - *Median Voter Theorem*: If preferences are single peaked, majority voting yields outcome preferred by median voter
  - Stable, but generally not Pareto efficient
  - $\max_G \quad V_m(G) + M_i - \frac{G}{N} \Rightarrow V_m'(G^m) = \frac{1}{N}$
  - Average voter's preference efficient: $\sum_i \text{MB}_i = \text{MC} \Leftrightarrow \sum_i V_i'(G^o) = 1 \Leftrightarrow \frac{1}{N}\sum_i V_i'(G^o) = \frac{1}{N}$
  - Degree of difference = Degree of inefficiency — no account of intensity of preferences



- – Crowding out private provision
  * Individuals solve $\max_{G_i} V(G_i + \overline{G_{-i}}) + M_i - G_i \Rightarrow$ FOC: $V'(G_i^* + \overline{G_{-i}^*}) = 1 \quad \forall i$
  * Tax $T = \sum_i t_i$
  * Total contribution $Z_i = G_i + t_i$
  * New optimization: $\max_{Z_i} V(Z_i + \overline{Z_{-i}}) + M_i - Z_i \Rightarrow$ FOC: $V'(Z_i^* + \overline{Z_{-i}^*}) = 1 \quad \forall i$
  * Isomorphic program → $Z^* = G^*$ → Full crowding-out

* 3 assumptions for full crowding-out:
  - All individuals contribute to public good when privately provided
  - Taxes collected from each individual do not exceed their voluntary contribution in absence of government supply
  - Individuals care only about total amount of public good provided, not own private contribution
  - c.f. Warm glow
  - ⋆ Otherwise, partial crowding out
* Kingma (1989): Individual contributions to US public radio stations
  - Crowd-out rate 13.5%
  - Problem: Government support may depend on individual contributions (reverse causality)
* Hungerman (2005): Church-provided welfare, Clinton 1996 reform reducing/eliminating welfare for non-US citizens
  - Government -$1 → Church increase +$0.4, but not through higher donations
* Andreoni and Payne (2003): Charities may decrease fundraising efforts in response to increasing government grants
  - +$1,000 government grants → -$250 fundraising
  - Matching grants beneficial

– Financing through distortionary taxes
* *Pigou Conjecture*: Total cost of providing public good higher than production cost when financed by distortionary taxation
* Atkinson and Stern (1974): $U(X, L, G) = V(G) - \frac{L^{k+1}}{(k+1)} + X$
  - First best, with lump sum taxation: $N \times V'(G^{FB}) = 1$
  - Second best, with labor taxation: $N \times V'\left(G^{SB}\right)\left[1 - \frac{\tau}{1-\tau}e\right] = 1$
  - $G^{SB} < G^{FB}$

5 Taxation of Goods and Services (AB)

5.1 Background and definitions

- Expenditure (purchases) $\neq$ Consumption (regardless of who purchased)
- *Indirect taxation*
  - Traditional definition: Whether tax is paid indirectly though goods purchases
  - But: Direct taxes can also be remitted by third party, and shifting can be partial of both direct and indirect taxes
  - Modern definition: Indirect taxes levied on transactions irrespective of the characteristics of buyer and seller

Indirect	1	Differentiated sales tax
Transitional	2	Uniform sales tax
	3	Proportional expenditure tax
Direct	4	Linear direct tax
	5	Non-linear direct tax

Table 1: Typology of taxes

- Indirect taxation opposed by Left because regressive, Right because effective
- History of taxation dominated by indirect taxation, and still prevalent in developing countries
- IMF, EU advocating adoption of VAT
- *Excise tax*: Function of quantities, $q = p + \tau$
- *Ad valorem tax*: Function of prices, $q = p(1 + \tau)$
- Value Added Tax (VAT): A broad-based tax on commodity sales with systematic offsetting of the tax charged on inputs
  - Applies to all sales to private consumers and other businesses (B2C and B2B)
  1. No taxation of intermediate goods
  2. Remittance is fractional
  3. Third party reporting
  4. Tax collection earlier (cash flow benefit)
  - *Reduced rate*: Goods faced with lower VAT rates, but for which sellers can still get credit for input VAT expenditures
  - *Zero rate*: As reduced rate, with no output VAT
  - *Exempted goods*: Sales not subject to VAT but seller also cannot reclaim input VAT
- Retail Sales Tax (RST): Tax on value of sales to final consumers, with sales to other businesses (B2B) untaxed
  1. No taxation of intermediate goods
  2. Tax remittance at final sale only
  3. RST requires end user distinction, between sales to businesses (untaxed) and sales to final consumers (taxed)

5.2 Incidence

- Standard textbook model: Partial equilibrium, perfect competition, perfect information, no compliance cost

  1 Legal incidence differs from economic incidence

  2 Invariance of tax incidence

  3 More inelastic factor bears more of tax

  4 Symmetry of tax increases/decreases

  – General equilibrium case:

    * Tax shifting impacts other markets
    * Factor prices affected
    * Demand very inelastic → Full incidence of VAT on consumers, except with close substitutes (e.g. border trade)

- Empirical evidence

  – Kerschbamer and Kirchsteiger (2000) (ultimatum game, varying remittance responsibility): Net amount differs based on remittance → Invariance assumption violated

  – Besley and Rosen (1999) (US cities, commodity prices): Full-shifting for some commodities, over-shifting for others

    * Evidence of retailers' market power

  – Kosonen (2015) (Finland, experimental VAT cut, DiD): 50% pass-through, higher for large firms, no quantity change

  – Benedeck et al. (IMF, 2015) (EU VAT changes): Full shifting for standard rate goods, lower for reduced rate goods

- Salience: Visibility of taxes might affect behavioral responses, $\frac{dx}{dp} \neq \frac{dx}{dt}$

  – Chetty et al. (2009): Degree of underreaction to tax θ: $\theta = \frac{\partial \log(x)}{\partial \log(1+\tau)} / \frac{\partial \log(x)}{\partial \log(p)}$

  – Grocery store experiment, tax-inclusive price tags → -8% demand

  – State-level changes in excise and sales taxes on alcohol → $\theta = 0.06$

  – Incidence formula with salience: $\frac{dp}{d\tau} = \theta \frac{\varepsilon_D}{\varepsilon_S - \varepsilon_D}$

- Asymmetry of VAT pass-through

  – Standard theory predicts symmetry

  – Small literature, mixed results

  – Benzarti et al. (2019): VAT reforms in Finland, France, EU

    * Evidence of asymmetric pass-through
    * Story: Fear of consumer antagonism; adjustment shock to increasing but not to decreasing prices, firms accumulate stock of shocks not transmitted to posted prices

- Is indirect taxation regressive?

  - Difficult to estimate life-time consumption, permanent income using cross-sectional data
  - VAT slightly progressive as share of consumption
  - Excises (e.g. tobacco) regressive
  - Reduced rates progressive

5.3 Optimal commodity taxation

- First best: Lump sum tax
  - $(1+\tau)q_1x_1 + (1+\tau)q_2x_2 + (1+\tau)wL = wT$
  - $q_1x_1 + q_2x_2 + wL = \frac{wT}{1+\tau}$
  - $\rightarrow$ Tax on full income (exogeneous)
- Government does not observe leisure, but hours worked, so can subsidize labor supply
  - $(1+\tau)q_1x_1 + (1+\tau)q_2x_2 = (1+\tau)wh$
  - $1+\tau$ cancels out, so no revenue collected unless distorting relative prices
- *Inverse elasticity rule*: Assuming cross-effects to zero, optimal taxes inversely proportional to elasticities
  - $\min_{t_1,\ldots,t_n} \sum_i^n DWL_i \quad \text{s.t. } \sum_i R_i = R$
  - $\frac{\partial DWL_i}{\partial \tau_i} / \frac{\partial R_i}{\partial \tau_i} = \lambda$
  - $\rightarrow \frac{\tau_i}{P} = \lambda \left(\frac{1}{\varepsilon_S^i} + \frac{1}{\varepsilon_D^i} \right)$
  - Each commodity should have different tax rate
  - Optimal tax rate depends on elasticity of demand and supply
  - Elastically demanded goods should be taxed less than inelastically demanded goods
  - BUT: Assumes no cross-price effects
- *Ramsay rule*
  - Assumptions: Identical households, only commodity taxes, competitive economy, pre-tax prices p_i fixed and tax-included prices $q_i = p_i + t_i$, government needs to raise revenue R
  - $\max_{t_1,\ldots,t_n} V(q_1 \ldots q_n, w, l)$ s.t. $R = \sum_{i=1}^n t_i x_i$
  - $\rightarrow \sum_{i=1}^n t_i S_{ik} = -\left[1 - \frac{\alpha}{\lambda} - \sum_{i=1}^n t_i \frac{\partial x_i}{\partial l}\right] x_k = -\theta x_k$, with S_{ik} = derivative of compensated demand curve
  - Optimal tax system should have (i) same reduction in compensated demand for each good, and (ii) limited distortions in quantities (not prices)
  - But with identical individuals, lump sum taxation optimal, no commodity taxation
- If heterogeneous compensated demands, goods that are complementary with leisure should be taxed at higher rate; if homogeneous, uniform taxation
- Rates differentiation
  - Motivations: Redistribution, externalities, optimal taxation (discussed)

  - Problems: Poor targeting (rich also spend on food), direct taxation more efficient for redistribution, administrative costs, opportunity for industry lobbying for reduced rates

- New Zeeland *Goods and Services Tax (GST)*: Comprehensive base, single rate, low threshold registration

5.4 Direct vs. indirect taxation

- Balanced view
  - Direct taxation better for redistribution
  - Indirect taxation more efficient to raise revenues
  - Compliance is higher with VAT third party reporting
  - Lower disincentives effect on labour supply
- Superiority of direct taxation
  - Indirect taxation historical remnant from time with insufficient administration/information
  - Direct taxation better for all objectives
- Uniform commodity taxation equivalent to linear labor income tax (if same taxation on inheritance)
  - $(1+\tau)q_1x_1 + (1+\tau)q_2x_2 = wh$
  - $q_1x_1 + q_2x_2 = \frac{wh}{1+\tau} = \left(1 - \frac{\tau}{1+\tau}\right) wh$
- Atkinson and Stiglitz (1976): Differentiation of commodity taxation depends on the relationship between labor and the marginal rate of substitution between commodities
  - Assuming weak separability (separate indirect utility of consumption of all goods and of leisure), no need for indirect taxation
- Tax compliance: Third party reporting creates paper trail, incentives for information gathering
  - Pomeranz (2015) (Chile, deterrence letter and announced audit): Significant effects

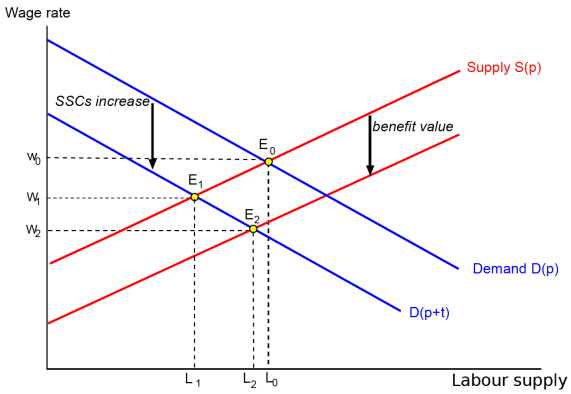
6 Labor Income Taxation (1) (AB)

6.1 Introduction

- Large increases in income tax rates with with war efforts: Top rates reached 40–70% in WWI, 70–97% in WWII
- Types of labor income taxation:
  - Income tax: Taxation of labor and capital income
  - Social Security Contributions (SSCs): Confer entitlement to receive future social benefit
    - Taxation of earnings (not capital income)
    - Nominally split between employee and employers
    - Usually capped at threshold
  - Means-tested benefits: e.g. Child benefits, minimum income
    - Since negative tax payment, removed with increasing income, analysis similar to labor taxation

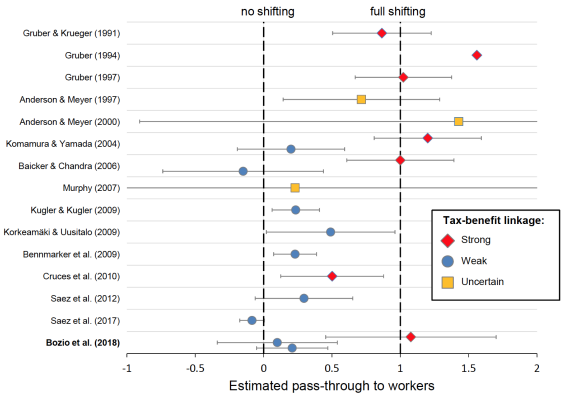
6.2 Incidence

- Conceptual framework (Feldstein, 1994)
  - 2 types of workers T and C
  - Posted wage w_K
  - Payroll tax on employers τ_K
  - Elasticity of substitution between workers σ
  - Extent to which employees value employer contributions q
  - Elasticity of labor supply η^S
  - Perceived wage of workers of type k: $\tilde{w}_k \equiv w_k\,(1 + q\tau_k)$
- Pass-through ρ of employer SSCs to wage of treated workers relative to control workers
  - $\rho = \frac{\mathrm{d}\ln\left(\frac{w_T}{w_C}\right)}{\mathrm{d}\ln(1+\tau_T)} \approx -\frac{\sigma+\eta^S \cdot q}{\sigma+\eta^S}$
  - ⋆ $q = 1 \rightarrow$ Full shifting to workers
  - ⋆ $q = 0$ and $\sigma >> \eta^S \rightarrow$ Full shifting (usual assumption)
  - ⋆ $q = 0$ and $\eta^S >> \sigma \rightarrow$ No shifting



- Empirical estimates
  - Labor income shares fairly stable
  - Gruber (1997) (Chile, privatization of pension system, cut in SSCs): No employment effect and full-shifting of SSCs to wages

- Saez et al. (2012) (Greece, uncapping of SSCs for new cohorts, RDD): No labor supply effect (intensive and extensive), incidence of SSCs similar to nominal incidence
- Saez et al. (2019) (Sweden, cut in SSCs below age 25, RDD): Large impact on employment, no shifting at individual level to wages, some evidence of shifting to wages on firm level
- Bozio, Breda and Grenet (2019) (France, uncapping reforms, DiD): Incidence of SSC on employers for reforms with no tax-benefit linkage (health, child), incidence on employees in reform with strong linkage (pension)
- For income tax, limited evidence but suggests at leasts partial shifting



6.3 Labor supply

- Labor supply elasticity: $\varepsilon = \frac{\partial \log L}{\partial \log w}$
- Leisure $L = T - H$ is time not spent in market work
  - Household production, unpaid activities, human capital accumulation, pure leisure
  - Trade-off between untaxed activities (home production) and taxed activities (work and market consumption)
- Baseline model: Static, pure intensive margin choice, no frictions or adjustment costs, linear tax system
  - $\max_{c,l} u(c,l)$ s.t. $c = wl + R$
  - Uncompensated (Marshallian) elasticity: $\varepsilon^u = \frac{w}{l}\frac{\partial l^\mu}{\partial w}$
  - Income effect: $\eta = w\frac{\partial l}{\partial R}$
  - Compensated (Hicksian) elasticity: $\varepsilon^c = \frac{w}{l}\frac{\partial l^c}{\partial w}$
    - Determines DWL, since income effect can be used to compensate through transfers
  - Slutsky equation: $\frac{\partial l}{\partial w} = \frac{\partial l^c}{\partial w} + l\frac{\partial l}{\partial R}$
    - $\varepsilon^u = \varepsilon^c + \eta$
  - Net effect of income tax ambiguous
    - Income effect: Lower unearned income, lower after-tax wage $\rightarrow$ Increased labor supply
    - Substitution effect: Lower after-tax wage $\rightarrow$ Decreased labor supply
- Estimation issues
  - Unobserved heterogeneity: w likely correlated with taste for work

  - Non-participation: Low wage earners must have very high unobserved propensity to work
  - Non-linear budget set: Tax and benefit systems non-linear, taste for work correlated with tax rate
  - Intertemporal substitution: Temporary/permanent wage/tax change have different impacts on labor
  - → Sufficient statistics: Develop formulas for welfare analysis not depending on structural parameters

- Empirical evidence
  - Intensive margin: Primary earners (used to be men) have low elasticities (around 0.1), secondary earners (typically married women) much higher (0.5–1)
  - Extensive margin: Highly educated men have very low participation elasticities, married women much higher (if taxation/benefits at household level), lone mothers very high
  - Blau and Kahn (2007): Married female labor supply elasticity falling sharply over time
  - *Negative Income Tax (NIT)*: Guaranteed income payment to all poor households, gradually reduced with earnings (US, 1960s)
    * Ashenfelter and Plant (1990): Male elasticity 0.1; female 0.5, concentrated along extensive margin
    * BUT: Self-reported, selective attrition, no separation of income and substitution effects
  - Cesarini et al. (2017) (Sweden, lottery winners): Effects on both extensive and intensive labor supply margin, persistent over time; significant but small income effects ($\eta = -0.1$, effects on spouse but not as large as on winner
  - Long-run macro elasticities: 0.02–0.03
    * Maybe utility depends on relative consumption, long-run estimates misleading
  - Prescott (2004) (7 OECD countries, GE model): $u(c,l) = c - \frac{l^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}$
    * $\varepsilon = 0.7$

- Macro calibrated models need high labor supply elasticities, micro evidence suggest small ones
  - Omitted variables: Labor market regulations, culture
  - Extensive vs intensive margin: Microeconomists missing large extensive margin effects
    * Chetty et al. (AER, 2011): In Hicksian steady state, possible to reconcile micro and macro
  - Frictions: Macro estimates larger because long-run, capturing frictions

  - Other programs: Pension, education, child care

6.4 Taxable income

- Deadweight loss of taxation should depend on all behavioural responses, not just hours worked
  - Effort on the job, career choice, form and timing of compensation, tax avoidance (legal shifting of income, tax evasion (illegal underreporting of income)
- Elasticity of taxable income (ETI): $e = \frac{1-\tau}{z}\frac{\partial z}{\partial(1-\tau)}$
  - $dDWL = dW - dR$ (with dW = utility loss, dR = revenue increase)
  - $dR = dM + dB$ (with dM = mechanical effects, dB = behavioral revenue effects)
  - Envelope theorem: $dW = dM$
  - $dDWL = dM - (dM + dB) = -dB \rightarrow$ Function of ETI
- Saez et al. (2012): ETI not sufficient if tax shifting (between personal/corporate income, over time) or fiscal externalities (e.g. charitable giving)
- Older studies with higher ETI, newer with lower
- Feldstein (1995) (US 1986 tax reform, top MTR down and broadening of tax base): ETI 1.1–3.1 (very large), but...
  - Mean reversion $\rightarrow$ Underestimation
  - Non-tax related changes in inequality, as T and C groups not from same part of distribution $\rightarrow$ Overestimation
  - Very small sample
  - Heterogenous elasticity, e.g. increasing in income
  - Behavioral effect confounded with mechanical effect of broadened tax base
  - Some responses short-term shifting, not long-term
- Gruber and Saez (2002) (US, all tax changes 1979–1990): Broad income 0.12, taxable income 0.4, higher for top income groups
  - Imprecise, sensitive to exclusion of low income, mean reversion
- Kleven and Schultz (2014) (Denmark, census data over 25 years): Labor elasticity 0.05–0.2, capital 0.1–0.3, larger for large tax changes and top incomes

7 Labor Income Taxation (2) (AB)

7.1 Optimal tax and transfer systems

- SWF $= \int_i \mu_i u^i$, with μ_i = social weight on individual i
- Simple optimal taxation model
  - $u(c)$ identical for all
  - No behavioral responses, so pre-tax income z fixed
  - With utilitarian SWF: $\int_0^\infty u(z - T(z))h(z)dz$, s.t. $\int T(z)h(z)dz \geq E(\lambda)$ (revenue requirement)
  - FOC: $u'(z-T(z)) = \lambda \rightarrow$ After-tax income same for all
  - Utilitarianism with decreasing marginal utility and no behavioural responses leads to perfect egalitarianism
  - BUT: Behavioral responses, utilitarian 100% redistribution not in line with people's perceptions of fairness
- Tagging: Government can observe characteristics X, correlated with endowments/ability and immutable $\rightarrow$ Redistribution based on X can be efficient (no DWL loss)
  - BUT: Stigmatization of tagged individuals, degree to which characteristics reflect needs, administrative costs (e.g. medical tests for disability)
  - Suggestions: Gender, height
- *Mirrlees model*: $T(.) < 0$ at the bottom (transfers), $T(.) > 0$ further up
  - Non-negative MTR $T'(.) > 0$ (rules out EITC/working tax credit), MTR should be 0 at the top if the skill distribution is bounded
- Optimal linear income taxation
  - Individuals earn z and consume $c = (1 - \tau)z + T(0)$
  - Maximize $u(c, z)$ to get labor supply choice $z(1 - \tau, R)$
  - Revenue maximization FOC: $R'(\tau) = Z - \tau^* \frac{dZ}{d(1-\tau)} = 0$
  - $\tau^* = \frac{1}{1+e}$, where $e = \frac{1-\tau}{Z}\frac{dZ}{d(1-\tau)}$
- Maximization of general SWF
  - $\overline{g}$ average normalised social marginal welfare weight weighted by pre-tax income
  - Optimal linear rate: $\tau^* = \frac{1-\bar{g}}{1-\bar{g}+e}$
  - No taste for distribution ($\overline{g} = 1$): $\tau^* = 0$
  - Rawlsian SWF ($\overline{g} = 0$): $\tau^* = \frac{1}{1+e}$
  - Higher inequality $\rightarrow$ lower $\overline{g}$
- Empirically, willingness to give large sums to poorest, but costly, so fast withdrawal $\rightarrow$ High MTR

7.2 Taxing top incomes

- Notation:
  - $\overline{z}$: Threshold for top marginal tax rate τ
  - z: Average income for taxpayers above $\overline{z}$
  - g: Social value of additional income of high earners
  - Elasticity $e = \frac{1-\tau}{z}\frac{\partial z}{\partial(1-\tau)}$
- Effects of small increase $d\tau$:
  - Mechanical effect: $dM = N(z - \overline{z})d\tau$
  - Behavioral response: $dB = -N \times e \times z\frac{\tau}{1-\tau}d\tau$
  - Welfare effect: $dW = -g \times N(z - \overline{z})d\tau$
  - Net effect: $N(z-\overline{z})d\tau\left[1 - g - e\frac{z}{z-\overline{z}}\frac{\tau}{1-\tau}\right]$, where $a = \frac{z}{z-\overline{z}}$ is the thinness of the income distribution
  - At optimum, net effect $= 0 \rightarrow \tau^* = \frac{1-g}{1-g+ae}$
- Saez et al. (2012) (US time series): Top 1% income share increasing with decreasing top MTR, but not the reverse
- Piketty et al. (2014) (pre-tax top 1%, 18 OECD countries, 1960–2010): Elasticity close to zero –1980, 0.6 onwards, with no effect on GDP
- Kleven et al. (2013)
  - Football, foreign player increase: High elasticities, 1.2–1.5
  - Denmark, top tax rate decrease for 3 years: Elasticities $> 1 \rightarrow$ Tax competition matters
- Sources of top income inequality:
  - Technology favoring skilled workers
  - Supply side: Rich people work more, especially with lower MTR $\rightarrow$ Lower top MTR
  - Tax avoidance: Less avoidance, international mobility with lower MTR $\rightarrow$ Broaden tax base, coordinate internationally
    - Increase in top incomes may be overestimated, if tax avoidance was more prominent before
    - Explains small effect of MTR changes, but MTR with narrow tax base measure yields similar results
  - Rent-seeking: Top earners extract more pay when top rates are low $\rightarrow$ Increase top MTR
    - But no correlation between MTR and growth

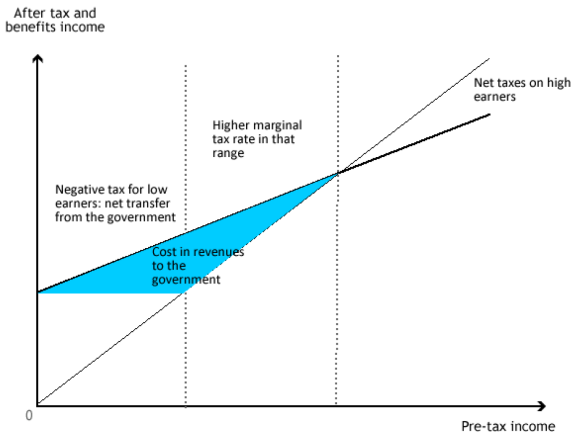
7.3 Transfers to the poor

- Traditional welfare
  - Tagging: e.g. Disability, old age, lone mothers
  - Means-testing with 100% tax rate $\rightarrow$ No incentives to return to work $\rightarrow$ Poverty trap

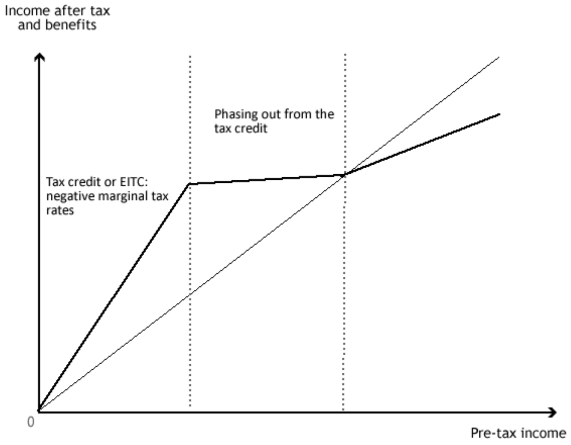
- Negative Income Tax (NIT)/Basic Income (BI)
  - Replacement of all welfare benefits by guaranteed income paid by government
  - Each additional income taxed at marginal rate below 100%
  - Concern about unit of taxation: Household or individual
  - Administrative costs higher than for tagging
  - Lower marginal tax rates for low incomes → Positive effects on extensive margin
  - Higher marginal tax rates higher in the income distribution → Negative effects on intensive margin

Iron triangle of redistribution: Can only choose two

- Redistribution to the poor (high replacement income)
- Incentives to work (low marginal tax rates)
- Low cost to the government



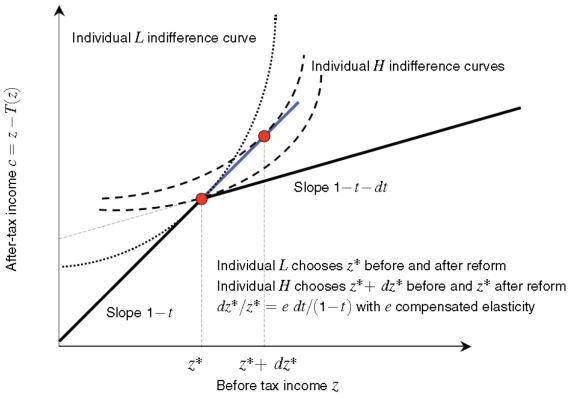
- Workfare: Removing high marginal tax rates on low incomes



- 2 approaches:
  - Mirrlees: Intensive margin, lump sum grant to those with no earnings, high MTR at bottom
    - Low cost, limited intensive response
  - Diamond and Saez (2011):
    - g_0: Social marginal weight on zero earners
    - e_0: Elasticity of fraction non-working to bottom net-of-tax rate
    - Optimal bottom MTR with intensive margin only: $\tau_1 = \frac{g_0 - 1}{g_0 - 1 + e_0}$
- Empirically, bigger labour supply elasticities at extensive margin (fixed costs of work)
- Saez (2002): Starting with positive phase-out rate (tax) $\tau_1 > 0$ at low earnings levels, increasing transfers (reducing the phase-out rate) is desirable, as long as the intensive response (disincentive to work) is small

  - Redistribution + increased labor force participation (lower expenditure on benefits)

- → Trade-off between intensive and extensive margin effects of MTR
- Earned Income Tax Credit (EITC) in USA: Based on family income, number of children
  1. Increasing subsidy part
  2. Constant amount
  3. Taper rate, withdrawing benefits with income
  - Consistent positive employment effects for single mothers
    - Hoynes and Patel (2017): Extensive margin elasticities 0.32–0.45
  - Small intensive margin effects (e.g., clustering at the kink)
  - BUT: Low amount to childless, little increase with more than two children, marriage penalty, complexity
- Saez (2010): Bunching at kinks provides non-parametric estimates of intensive elasticity
  - $\varepsilon^c = \frac{dz/z}{dt/(1-t)} = \frac{\text{excess mass at kink}}{\%\Delta\text{net of tax rate}}$
  - Implied intensive elasticity of EITC: 0.25 — fueled by self-employment



- Chetty et al. (2013) (US EITC, local information effects, zip codes):
  - Places with high self-employment EITC bunching display wage earnings distribution more concentrated around plateau
  - Significant intensive margin effects larger than extensive margin effects
  - → Importance of information

8 Normative and Intertemporal Theories of Social and Fiscal Justice (TP)

8.1 Inequality and beliefs

- Conflicts about inequality and role of government from different interests, values/objectives, and belief systems
  - Income and wealth correlated with political attitudes but not perfect predictors
  - Policies broadly argued for in terms of universal objectives, rather than self-interest
    - Arguments not based on specific SWF, but fairness, helping poor, etc.
  - Belief systems important explanation, and positive/constructive
- Piketty (1995): Modelling politics as based on beliefs
  - Upwardly and downwardly mobile individuals (from parental to own income) have probability to vote for left-wing parties intermediate between those of permanently low-income and high-income groups
    - No longer the case for level of education, but for income
  - Income levels: $y_0 < y_1$
  - Individual effort: e_i
  - Index of importance of inequality in social origins: $\Delta\pi = \pi_1 - \pi_0$
  - $P(y_i = y_1) = \pi_0 + \theta e_i$ if parental income $= y_0$
  - $P(y_i = y_1) = \pi_1 + \theta e_i$ if parental income $= y_1$
  - Redistributive taxation: $c_0 = (1-t)y_0 + ty$, $c_1 = (1-t)y_1 + ty$, with $y =$ Average pre-tax income
  - Individual utility: $U_i = c_i - C(e_i)$, with $C(e) = e^2/2a$
  - $\max_{e_i}(1-p_i)_0 + p_i c_1 - C(e_i)$, with $p_i = \pi_i + \theta e_i \rightarrow$ FOC: $e_i = a\theta(1-t)(y_1 - y_0)$
  - $\rightarrow e_i$ decreasing in redistribution, but magnitude depends on relative importance of θ and $\Delta\pi$
  - Assuming Rawlsian objective (maximizing with $p = \pi_0 + \theta e_i$), $t^* = \frac{H\Delta\pi}{a(y_1-y_0)\theta^2}$, where $H =$ Pop. share with high-income parents
  - $\rightarrow t^*$ increasing in $\Delta\pi$ (parental origins more important), decreasing in θ (effort less important)
  - $\Rightarrow$ Politics conflict over beliefs about $\Delta\pi$ and θ
- Explaining differences in beliefs
  - Difficult to learn about $\Delta\pi$ and θ
  - Self-serving beliefs play a role, but not necessary for analysis

  - Effort and upward mobility → Belief in importance of effort → Exerting more effort → Increasing income

- Even if family origins to not matter, redistribution desired due to risk aversion
  - Ex-post Rawlsian social objective: $\max c_0 - C(e) \rightarrow t^* = \frac{2}{3} + \frac{\pi}{3a(y_1 - y_0)\theta^2}$

8.2 Political conflict and belief systems about inequality

- Extensive evidence that beliefs about luck v.s. effort, income mobility prospects matter explain attitudes towards inequality
- Basic model: Left pessimistic about mobility (belief in luck and redistributive taxation), right optimistic about mobility (belief in effort)
- Introducing individual ability parameters β_i
  - $P(y_i = y_1) = \pi_0 + \theta\beta_i e_i$ if parental income $= y_0$
  - $P(y_i = y_1) = \pi_1 + \theta\beta_i e_i$ if parental income $= y_1$
  - → Beliefs in $\beta_1 > \beta_0$ (ability depending on family origin) → Conservative, pessimist right → Belief in low taxation and low mobility
  - → Beliefs in high $Var(\beta_i)$ with limited correlation with parental origins → Eugenist right → Beliefs in low taxes, innate talent
  - → Beliefs in strong effects of education on β_i → Positive left → Beliefs in high taxation, high mobility
- Leftist belief in addressing primary inequality through education, worker co-determination, etc, not just ex-post redistribution
- France, UK, US 1950s–Now: Highly educated individuals now voting for left, high-income, and especially high-wealth, individuals still for right
- Model 1: Domestic inequality v.s. external inequality
  - In addition to low tax–high tax dimension, also openness/migration dimension, with left more pro-openness ($O_L > O_R$)
  - Views on openness/migration correlated with education and income
  - Extra-European migration increase in EU, civil rights movements in US → Salience of openness dimension → Some poor vote for political right
  - Globalization → Easier tax evasion → Policy conflict about taxes/redistribution disappears → Policy conflict between globalists and nativists
  - Alesina et al. (2001): Less demand for redistribution in US v.s. EU because more racial prejudice
- Model 2: Education inequality v.s. wealth inequality
  - Education by nature inegalitarian

- Rise of higher education → Differential chances of admission to selective institutions x_1 → Higher probability with highly educated parents → Belief in education related effort among highly educated
  - $P(x_i = x_1) = \alpha_0 + \Phi f_i$ if parental education = x_0
  - $P(x_i = x_1) = \alpha_1 + \Phi f_i$ if parental education = x_1
  - Income: $P(y_i = y_1) = \pi_0 + \theta e_i + \mu s$ if $x_i = x_1$ and parental income = y_0
- Multiple elites:
  - Brahmin left: Believes in education related effort Φ, prioritizes funding for top education
  - Merchant right: Believes in business effort θ, lower taxes
  - Those with effort in neither tend to abstain, or sometimes vote for nativists

8.3 Intertemporal justice

- Inequality between generations (rather than within) arguably the most pressing issue today, especially including natural capital
- Phelps (1961): *Golden rule of capital accumulation*: Re-invest the full capital share, until marginal product of capital r equals growth rate g
  - Only reasonable assuming zero productivity growth and exogenous and permanently positive population growth
  - EU: Max 0.5% structural deficit, not taking into account public assets, investment in e.g. education
  - Population N growth rate: $n \geq 0$
  - Productivity growth rate: h → Efficient labor supply L growth rate: $g = n + h$
  - Labor participation rate λ
  - Production function $Y_t = f(K_t, L_t)$ with constant returns to scale
  - Closed economy, capital accumulation determined by saving: $\frac{dK_t}{dt} = I_t = s_t Y_t$
  - Representative agent approach, assuming away intragenerational inequality
  - At balanced growth path, $\frac{dK_t}{dt} = gK_t = I = sY_t \rightarrow \beta_t = \frac{K_t}{Y_t} = \frac{s}{g}$
  - Maximizing long-run per capita consumption $c = (1-s)f(k) = f(k) - nk$: $f'(k^*) = n$
  - → With $h = 0$, $g = n > 0$, $f'(k^*) = r^* = g$
  - No point in accumulating more than capital share, since k only meant to increase c – but $r^* = g$ a maximum (to avoid dynamic inefficiency), can accumulate less
  - With $f(k) = \alpha k, n = 1\%, \alpha = 30\%, r^* = 1\% \rightarrow \beta^* = \frac{k^*}{y^*} = \frac{\alpha}{r^*} = 3,000\%$ (more than actual ∼ 500–600%
  - If $h > 0$, $n = 0$, or n is endogenous, no simple answer

  - If $n = 0$, $r^* = n = 0\% \rightarrow$ Infinite capital accumulation

- Modified golden rule: $r^* = \delta + \gamma g$
  - Not so useful
  - $U(c) = \frac{c^{1-\gamma}}{1-\gamma}$
  - Maximizing $V = \int_{t>0} e^{-\delta t} U(c_t)$
  - δ: Time preference; γ: Concavity of SWF
  - Assuming $h > 0$ and specifying intertemporal SWF
  - Controversy about choice of parameters
  - Sacrifice today in order to avoid future damage: $\mu Y_0 = e^{-r^* T} \lambda Y_T$, with $r^* = \delta + \gamma g =$ Social discount rate
  - Stern v.s. Nordhaus on climate change:
    - $\delta = 0.1\%$, $g = 1.3\%$
    - Stern: $\gamma = 1 \rightarrow r^* = 1.4\% \rightarrow$ Worth spending 9% of GDP to avoid 10% GDP loss in 2080
    - Nordhaus: $\gamma = 3 \rightarrow r^* = 4.0\% \rightarrow$ Worth spending 1.5% of GDP to avoid 10% GDP loss in 2080
  - If future growth certain, high γ makes sense $\rightarrow$ Rawlsian, maximizing consumption of poorest generation (current generation)
    - BUT: Growth endogeneous, long-run relative price of environment may be infinite

8.4 Aggregating interests/values/beliefs

- *Arrow's Impossibility Theorem*: Ruling out interpersonal comparisons of utilities, no consistent collective rule to aggregate individual preferences and take collective decisions
  - $\rightarrow$ Economics about creating new value, politics about dividing the pie in a more or less chaotic manner $\rightarrow$ Nihilist view of politics
  - OR: Need for minimal agreement about common values

- *Majority cycle*: In general, one can find policies A, B, C such that a majority prefers A to B, a majority prefers B to C, and a majority prefers C to A
  - *Condorcet Paradox*: With multidimensional political conflicts, majority cycles pervasive

- *Condorcet Jury Theorem*:
  - Assumption: Shared objective function (same values and preferences), but different beliefs and information about optimal policy given these
  - Choice between policies A and B, and everyone receives signal about optimality with common probability $p > 0.5$
  - Law of large numbers $\rightarrow$ With large enough population, probability of majority rule yielding optimal decision approaches 1

  - With different signal qualities, >2 policies, etc, indirect democracy may be preferred

- If politics about beliefs and information, public deliberation and communication critical

9 Capital Income, Inheritance and Wealth Taxes over Time and across Countries (TP)

9.1 Overview and definitions

- National income: $Y = F(K, L) = Y_K + Y_L = rK + vL$, where r = Average rate of return, and v = Average wage rate
  - Capital K both physical and immaterial
- In rich countries:
  - $\beta = K/Y = 5 - 6$
  - $\alpha = Y_K/Y = 25 - 30\%$ (net of depreciation)
  - $r = \alpha/\beta = 4 - 5\%$
  - $Y \approx €30,000$
  - Assets: Bottom 25% cash deposits; 25%–95% real estate; top financial assets
  - Income: Bottom 95% labor income; top capital income
- Capital in Europe $\sim$ 700% of GDP before WWI $\rightarrow$ 300% until 1980s $\rightarrow$ 500–600% now
  - US more stable around 400–500%
  - Importance of colonial assets for UK, France 1850–WWI
  - Net public wealth 25–30% 1950–1980, $\sim$0 now — but China $\sim$30%
- Capital taxation 8–9% of GDP in US, EU
  - $\approx$30% on capital flows or $\approx$1.5% on capital stock
  - Inheritance $<$1%; Annual wealth and property 1–2%; Corporate profits 2–3%; Personal capital income 2–3% of GDP

9.2 Choice of taxation

- With perfect capital markets, flow based and stock based taxation theoretically equivalent
  - But capital markets imperfect
  - Taxing stock incentivizes getting high returns on capital
- Progressive taxation in rich countries since $\sim$1900, more progressive after WWI, until 1980s
  - Not created on wealth in France, UK, US, because existing proportional real estate taxation difficult to reform
  - Germany, Sweden: Valuation problems $\rightarrow$ Suspension of wealth tax
  - Fisman et al 2016: US demand for relatively high progressive wealth taxation
  - US: Historically high income progressivity, but now, because of tax evasion, abolished

  - Progressive income tax important for financing public goods, while inheritance tax more to limit perpetuation and concentration of wealth

- Tax revenues/GDP in rich countries ∼10% until WWI, increase until 1980, then stable
  - 55% in Sweden, 45% France, 30% USA

- Inheritance taxes:
  - *Estate taxes*: Depend on total wealth left (real estate and personal estate, incl. financial), irrespective of how split between successors
    * US and UK, complete testamentary freedom
  - *Inheritance taxes*: Depend on wealth received by each successor and kin relationship
    * France and Germany, limited testamentary freedom
    * France: $\frac{1}{n+1}$ to each child; assets acquired during marriage and not inherited by one spouse split equally
    * Importance of exemptions, e.g. allowing for untaxed inter vivos gifts to children

- Limited evidence of capital flight
  - Zucman (2008): Outward and inward mobility seem to balance each other
  - Garbinti et al. (2016): Top financial wealth increased faster than real estate wealth (and much faster than national income) since 1980s-1990s, little evidence of capital flight

- Better to tax capital stock than transactions (e.g. UK mansion tax), but at time of transaction, available liquidity

10 Optimal Taxation of Capital Income, Inheritance and Wealth (TP)

10.1 Taxation and capital mobility

- Atkinson-Stiglitz: If full information on capital income flows, 100% life-cycle wealth (zero inheritance), and perfect capital markets, no reason to tax capital
- 4 main reasons for taxing capital
  - Fuzzy frontier: If frontier between labor and capital income flows not clear (e.g. for self-employed), better to tax both income flows at comparable rates
  - Fiscal capacity: If income flows difficult to observe for top wealth holders, wealth stock may be a better indicator of capacity to contribute than income
  - Incentives: By taxing capital stock rather than income flow, agents given incentives to get higher returns
    - Implicitly assumes imperfect capital markets, that rates of return depend on individual effort
    - Higher average returns for higher wealth levels (larger portfolios)
  - Meritocracy: Individuals not responsible for inherited wealth, so should be taxed more than labor income, imperfect capital markets imply that part of ideal inheritance tax should be shifted to lifetime capital tax
- Without fiscal coordination (automated exchange of bank information, unified corporate tax base, etc.), all forms of capital taxation difficult
- >8% of rich country output held in tax havens (10% EU, 4% US)
  - Effective tax rate paid by US firms substantially lower than nominal rate, and decreasing, in 2013 35% v.s. 20%

10.2 Motivations for taxing capital

- Assuming closed economy or perfect international coordination, to find economic optimum with coordination
- Atkinson-Stiglitz result on no capital taxation very strong conditional on strong assumptions
  - If 100% of capital accumulation from life-cycle savings, taxing capital or capital income equivalent to using differential commodity taxation (current consumption v.s. future consumption)
  - If separable preferences, differential commodity taxation undesirable → Tax labor income directly

- Formal model:
  - 2 time periods
  - Labor income $y_L i = v_i l_i$ in $t = 1$
  - Optimization: $\max U(c_1, c_2) - V(l)$ s.t. $c_1 + \frac{c_2}{1+r} = y_L$
  - Savings $s = y_L - c_1 \rightarrow c_2 = (1 + r)s = y_K$
  - $\rightarrow$ Taxing y_K equivalent to taxing relative price of consumption in period 2
- Atkinson-Stiglitz 2: Infinite-horizon dynasties (infinite elasticity of long run capital supply) $\rightarrow$ Optimal linear capital tax is zero (but optimal progressive tax positive
  - Dynastic capital stock high (k_t^A) or low ($k_t^B < k_t^A$), here assumed to be zero (0)
  - Proportion of high-wealth dynasties λ exogenous
  - High-wealth dynasties maximize standard $U_t = \sum_{t \geq 0} \frac{U(c_t)}{(1+\theta)^t}$
  - Competitive equlibrium $f'(k^*) = f'(\lambda k^A) = r^* = \theta$
  - Any taxation τ will yield $(1 - \tau)f'(k^*) = (1 - \tau)f'(\lambda k^A) = (1 - \tau)r^* = \theta \rightarrow$ Unsustainable disaccumulation
  - Worker income $y_\tau^* = v_\tau^* + s_\tau^* = f(k_\tau^*) - r_\tau^* k_\tau^* + \tau r_\tau^* k_\tau^* = f(k_\tau^*) - \theta k_\tau^* \rightarrow$ Maximum when $f'(k_\tau^*) = \theta \Leftrightarrow \tau = 0\%$
  - Result breaks down if long run elasticity of capital supply finite
  - With progressive tax, middle-class capital accumulation will compensate for the rich decline in capital accumulation

10.3 Optimal capital taxation

- Piketty and Saez (2013)
  - Bequest to next generation $b_{i,t+1}$ stochastic function of bequest received $b_{i,t}$ and shocks on bequest taste parameters, rates of return, wage rates, etc.
  - Always positive probability to move between wealth levels across generations, but mobility/inequality varies with shocks and economic parameter
    - Mobility decreases and inequality increases in $r - g$
  - $\rightarrow$ (Macro) Optimal Inheritance Tax Formula: $\tau_B = (1 - \frac{(1-\alpha-\tau)s_{b0}}{b_y}/(1 + e_B + s_{b0}$
    - τ_B increasing in macro bequest flow b_y; decreasing in elasticity e_B and average bequest taste s_{b0}
    - Trade-off between taxing rich cohort successors and taxing own children
    - Behavioral responses not so important as long as elasticity is reasonable
  - Realistic calibrations: $\tau_B = 50 - 60\%$

- Top inheritance tax rate rapidly increasing since WWI, peak in WWII 60–90%, now 30–50%
- (Micro) Optimal Inheritance Tax Formula: $\tau_B = \frac{1-(Gb^*/Ry_L^*)}{1+e_B}$
  - b^*: Average bequest left by zero-bequest receivers as a fraction of average bequest left
  - y_L^*: Average labor income earned by zero-bequest receivers as a fraction of average labor income
  - G: Generational growth rate; R: Generational rate of return
  - Does not require estimating average bequest taste, but instead mobility of resources across generations
  - Pure distribution effect: $e_B = 0$ & $G = R \to \tau_B = 1 - \frac{b^*}{y_L^*}$
  - Fiscal Golden rule: $e_B = 0$ & $b^* = y_L^* = 1 \to \tau_B = 1 - \frac{G}{R}$
- Equivalence between bequest tax and capital income tax violated if (a) tax enforcement constraints, (b)life-cycle savings, or (c) uninsurable risk in rate of return
  - Capital market imperfections can justify shifting one-off inheritance taxation toward lifetime capital taxation — although this creates distortions in inter-temporal incentives
  - Appropriate trade-off difficult, formulas complicated
- Taxation of immaterial capital
  - If copy costs zero, social optimum should involves free use, but need for incentivizing production of new ideas
  - In practice, both public production and private production with patents
    - Patents equivalent to gradual capital tax
- Capital usage rights: Maybe temporary private property, 100% tax at the end of period?
- Need to consider inequality in analysis of temporary property
  - Simple high flat wealth tax unlikely to be optimal

11 Corporate Taxation (AB)

11.1 Overview and definitions

- Firms are largely absent of tax theory, but remit 90% of taxes in OECD countries
- 2 opposing views:
  - Equity highly concentrated in top incomes/wealth, corporate income tax reduces tax avoidance through shifting and mitigates inequalities
  - Corporate income tax largely shifted to workers, hinders investment decisions and therefore growth, so cutting CIT is beneficial
- *Corporation*: Legal entity separate from the persons that form it; owned by *shareholders*
  - Corporate firms: Limited liability → Profit subject to corporate income tax
  - Non-corporate firms: Full owner liability for outstanding firm debt → Profit subject to personal income tax
- Reasons for taxing corporations:
  - Limited liability status important benefit, corporations also benefiting from public goods provision → Tax benefits
  - Backstop for personal income taxation: To escape income taxation, individuals could accumulate earnings tax-free within corporation → Limit tax avoidance
  - Taxation of pure profit/rents (returns exceeding returns to capital and labor): No distortion of investment decisions, so limited DWL
- *Investment tax credit (ITC)*: Tax credit amounting to percentage of firm's qualified investment expenditures, equivalent to accelerated depreciation
- *Research tax credit (RTC)*: Based on R&D spending
- CIT tax base:
  - Current costs C: Employee compensation, input costs
  - Depreciation costs Dep on capital investments (allowances legally specified)
  - Financing costs: Interest payments I and opportunity cost of equity OCE
- 3 dimensions of corporation taxes:
  - Income included in tax base
    - Full return to equity (most common): $Y = R - (C + Dep + I)$
    - Full return to capital: $Y = R - (C + Dep)$
    - Economic rent: $Y = R - (C + Dep + I + OCE$
  - Location of tax base

        * Source-based: Corporate income earned in country where productive activity takes place, tax on investment (most common)
        * Residence-based: Corporate income earned in residence country of corporate headquarters or shareholders, tax on savings (e.g. UK)
        * Destination-based: Corporate income earned in country where goods and services are consumed
    - Relationship with personal income taxation
        * Classical system: Tax liability of companies completely separated from those of shareholders; no relief for distributed profits (dividends), so they are taxed twice
            · Strong incentives for share buybacks and retained earnings
        * Imputation system: Shareholders receive credits for corporation tax paid on distributed profits
- Taxation of after-tax profits
    - Dividends: Taxed with personal income tax
    - Share buybacks: Capital gains tax
    - Retained earnings: Taxed only by CIT
- Trends in corporate taxation:
    - Decrease in statutory corporate tax rates, especially in 1980s — tax competition
    - Decrease in depreciation allowances: Broadening of the tax base while reduction in rates
        * But increase in R&D allowances
        * *Effective tax rates (ETR)*: $\frac{r^g - r^n}{r^g}$ (gross and net returns, with credits reducing difference)
    - Little evidence of decrease in tax revenues: High pro-cyclical volatility 1.5–3.5% of GDP, decrease mainly in US due to decreased profitability
    - Increase in tax avoidance and evasion: Use of transfer pricing and tax havens

11.2 Incidence

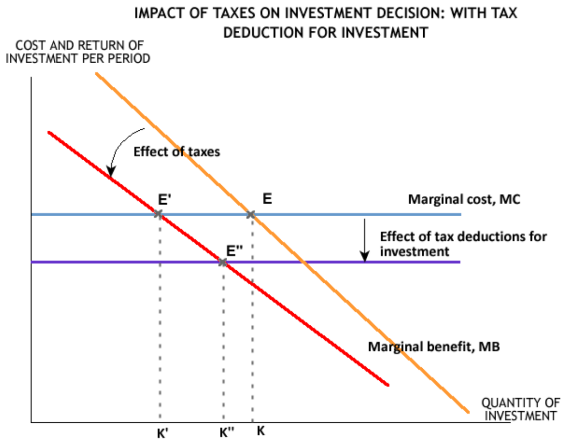
- Corporations remit taxes, but economic incidence is about changes in welfare, of individuals
- Individuals potentially paying CIT
    - Capital owners (through lower profits)
        * Old theory, support in Saez and Zucman (2019)
        * CIT very progressive, since capital ownership very concentrated (US top 0.01% wealth owning 45%)

    - ∗ BUT: Different share classes, with different rights to profits; indirect holding of equity (through e.g. mutual funds, life insurance – 60% of all US equity)
  - Workers (through lower wages)
  - Consumers (through higher prices)
- *Harberger model*
  - Static GE in **closed economy**
  - Corporate X and non-corporate Y sectors
  - Assumptions: Fixed supply of factors (short-run); free factor mobility across sectors; full employment of factors; constant returns to scale in both sectors; perfect competition
  - Increase in CIT $d\tau$
  1. Factor substitution effect: Capital bears tax
     - ∗ Depending on elasticity of substitution between capital and labor
     - ∗ Tax shifts production in sector X away from capital → Aggregate demand for capital decreases → As K fixed, r decreases
  2. Output effect: Mixed incidence
     - ∗ Shift in demand towards sector Y → If X capital intensive, reduced demand for capital and lower r ; If X labor intensive, increased demand for capital and labor bears some or all of tax
  3. Substitution + output: Overshifting effects
     - ∗ If X capital intensive → Potentially >100% incidence on capital
     - ∗ IF X labor intensive → Potentially 100% incidence on labor
  - → Incidence depends on elasticities
  - With realistic calibrations, *all* capital bears entire CIT, not only corporate sector, so progressive tax but less so than with full incidence on corporate shareholders
- Small open economy case (internationally mobile capital, immobile labor)
  - Sector 1: Small open economy; Sector 2: Rest of world
  - After-tax capital returns equalized: $r^* = F_{2K} = (1-\tau)F_{1K}$ → Capital moves until holds → Full tax incidence on labor
  - This case now more realistic than Harberger, more for EU than for US
- Empirical evidence
  - Limited evidence, few sources of variation

- Arulampalam et al. (2012) (firm-level data, 9 countries, 1996–2005, dynamic panel model with FD): Elasticity of wage bill w.r.t. CIT -0.1, 50–60% incidence of CIT on wages
  - 2 channels for CIT to affect wages: Directly through reducing post-tax profits on which workers and firms bargain; and indirectly through pre-tax profits through investment or output prices → Focus on direct channel
  - Tax liability instrumented for with country and year specific EMTR and ATR, lagged firm specific variables (e.g. fixed assets)
- Suáres Serrato and Zidar (2016) (local US markets (open economy), firm-level): 30–35% incidence on workers, 40% on shareholders
  - Allow for monopolistically competitive and heterogeneously productive firms
- Fuest et al. (2018) (changes in German municipality-level business tax, panel data, event study): +€1 tax → -30–70% wage bill, driven almost entirely by settings with collective wage bargaining
  - Corporate and non-corporate firms
  - Smaller effects for large firms

11.3 Efficiency costs

- Investment decisions
  - Investment important driver of growth, so taxation that affects investment matters for growth
  - Model: Firm decides on capital accumulation K_t, Profit function $F(K_t)$ concave, Price of capital goods q_t, Depreciation rate δ, Required rate of return ρ
  - NPV of new capital dK_{t+1}: $(1-\delta)q_t + \frac{F'(K_{t+1}+q_{t+1}}{1+\rho}$
  - $MB = MC \rightarrow F'(K_{t+1}) = q_t\left[(1+\delta)(1+\rho) - \frac{q_{t+1}}{q_t}\right] \approx q_t\left[\delta+\rho-\frac{q_{t+1}-q_t}{q_t}\right]$
  - With constant investment prices ($q_{t+1} = q_t$), UCC equals required rate of return plus depreciation: $\frac{F'(K_{t+1})}{q_t} = \delta + \rho$
  - With CIT:
    - NPV of depreciation deductions D_t: $\Gamma_t = \sum_{z=t}^{\infty} \frac{\tau_{Dividends} D_{z-t}}{(1+r)^{(z-t)}}$
    - $F'(K_{t+1}) = q_t \frac{1-\Gamma_t}{1-\tau_{CIT}} \times \left[\delta+\rho+\frac{q_{t+1}(1-\Gamma_{t+1})-q_t(1-\Gamma_{t+1})}{q_t(1-\Gamma_{t+1})}\right]$
    - Common CIT: Partial expensing ($D_0 < 1$), partial deductibility of financing cost ($\rho'(\tau_{CIT}) > 0$) → Need for higher r to justify investment → Reduced investment
  - If all costs deductible ($D_0 = 1$), → $\Gamma_{t+1} = \tau_{t+1} \rightarrow F'(K_{t+1})$ independent of CIT



- – Djankov et al. (2010) (PwC survey, 85 countries, 2005–2006, OLS): +10p.p. CIT → -2p.p. I/GDP
  - – House and Shapiro (2008) (US, changes in accelerated depreciation 2002–2003, DiD by asset category): Cost-of-capital intertemporal substitution elasticity of investment between -6 and -14
    - ∗ Accelerated depreciation generates large effective subsidy if firm is liquidity constrained
- Payout decisions
  - – Dividends typically taxed higher than capital gains and retained earnings
  - – Reasons for paying out dividends:
    - ∗ Agency problems: Managers may misuse cash, equity holders prefer tax inefficiencies to reduce manager control over firms assets
    - ∗ Signaling theory: Investors have imperfect information about firm, managers paying dividends signal that firm has cash to burn
  - – Chetty and Saez (2010) model:
    - ∗ Baseline cash holdings X, equity E, investment I with returns $f(I)$ in next period, dividends $D = E + X - I$
    - ∗ Net-of-tax payout in next period: $(1 - \tau_{Dividend})[(1 - \tau_{CIT})f(I) + X - D] + E$
    - ∗ Firm value $V = (1 - \tau_{Dividend})D - E + \frac{(1-\tau_{Dividend})[(1-\tau_{CIT})f(I)+X-D]+E}{1+r}$
    - ∗ Traditional view: Cash constrained firms
      - · Marginal value of paying dividends negative, pre-tax return above interest rate
      - · Firms should not pay dividends, should fund investments through equity
      - · Dividend tax similar to CIT, dividend tax cuts stimulate equity issues and investment
    - ∗ Modern view: Cash rich firms
      - · Marginal value of issuing equity negative, marginal investments funded out of retained earnings or riskless debt
      - · Firms should not emit equity, should split

cash between dividends and investment: $(1 - \tau_{CIT})f'(X - D) = r$

· Higher CIT lowers investment, change in dividend tax rate has no effect on investment

– Poterba and Summers (1984) (UK, 1955–1981, policy changes): Taxes on dividends impact substantially dividend payout (old view)

– Chetty and Saez (2005) (US 2003 dividend tax cut, simple diff): Elasticity of regular dividend payments w.r.t. MRT on dividend income -0.5

∗ Largest response from firms with strong principals whose tax incentives changed → Agency issues

∗ Dividend response too fast for old view, temporary tax cut effects also in line with new view

∗ Results consistent with +/-/0 effects on investment

– Yagan (2015) (US dividend tax cut in 2003, DiD by corporation type): Zero effect on investment, zero effect on wages (new view, or temporariness of reform blocking effects)

∗ No clear effects on investment

– Alstadsæter et al. (2017) (Sweden, 2006 dividend tax cut, DiD by cash constraint and close/wide holding): Cash-constrained firms increase investment relative to, cash-rich firms, closely held +32% and widely held +18%; no aggregate impact on investment (only reallocation)

∗ Lower dividend taxation → Higher payout from cash rich firms → More funds for cash constrained firms

• Elasticity of corporate taxable income

– Response of corporate taxable income to 1% change in statutory CIT rate

– Devereux et al. (2014) (UK, 2001–2008, bunching at kinks): Low elasticities, 0.15 for small firms, 0.50 for very small firms

– Firms maximize net-of-tax profit: $\pi = y - c(y) - T$

– Tax $T = t_c(B_c - A_c) + E$, with A_c lowest point of relevant bracket, E taxes paid in lower brackets

– Tax base $B_c = y - \alpha c(y)$, with α share of deductible costs

– → FOC: $c'(y) = \frac{1-t_c}{1-\alpha t_c}$

– Impact of CIT on total welfare: $dW = \frac{t_c B_c}{1-t_c} e d(1 - t_c)$, with $e = ECTI$

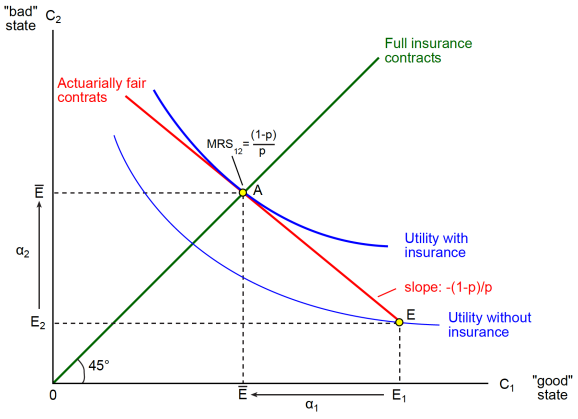
– Mechanical change in tax burden: $dM = -(B_c - A_c)d(1 - t_c)$

– Excess burden of CIT: $\frac{dW}{dM} = -\frac{B_c}{B_c A_c}\frac{t_c}{1-t_c}e$ = MDWL of tax increase

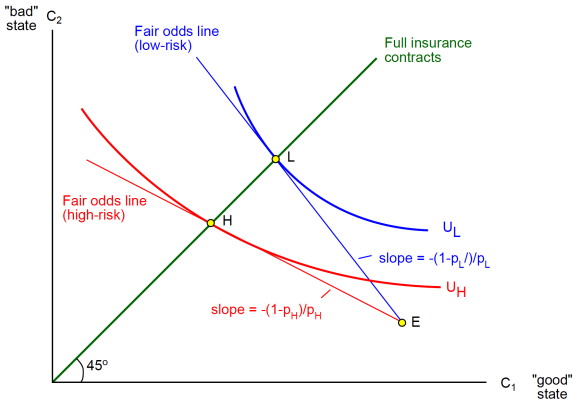
12 Social Insurance (JG)

12.1 Overview and motivation

- Social insurance: Government provision of insurance against adverse events
  - Unemployment, disability, workers compensation (against on-the-job accidents), health, social security (retirement, largest)
  - Mandatory participation
  - Eligibility and benefits depend on contributions, tied to events (e.g. unemployment) but not means tested
  - Largest and most rapidly growing part of government expenditure
  - 30+% in France; 20% OECD average, US, UK
- Trade-off between benefits of consumption smoothing and cost of distorting incentives, DWL
- Motivation for insurance: Reduce cost of adverse events for *risk-averse* individuals
- Sources of market failures justifying public intervention:
  - Asymmetric information (adverse selection)
  - Externalities: Negative externalities especially from underinsurance w.r.t. health
  - Individual optimization failures (myopia/improper planning)
  - Administrative costs: Large economies of scale in administrative costs → Mandated pooled insurance provides efficiency gains
    - Spinnewijn (2015): Overestimation of probability of finding a job
- Basic model
  - 2 states: Good (1) and bad (2)
  - Probability of bad state p
  - Income in good state E_1, in bad state $E_2 < E_1$
  - Insurance: Premium α_1, payout α_2
  - $EU = (1-p)u(E_1-\alpha_1)+p\times u(E_2+\alpha_2)$
  - *Jensen's inequality*: $U(E(c)) > EU \Leftrightarrow u[(1-p)C_1 + p \times C_2] > (1-p)u(C_1) + p \times u(C_2)$
  - $MRS_{12} = \frac{1-p}{p} \times \frac{u'(C_1)}{u'(C_2)}$
  - Perfect competition → Actuarially fair premium → $\alpha_2 = \frac{1-p}{p}\alpha_1$
  - $\max_{\alpha_1}(1-p)u(E_1-\alpha_1) + p \times u\left(E_2 + \frac{1-p}{p}\alpha_1\right)$
  - FOC: $MRS_{12} = \frac{1-p}{p} \rightarrow u'(C_1) = u'(C_2)$ → Full insurance
  - Insurance premium: $C_1 = C_2 \rightarrow \alpha_1^* = p(E_1 - E_2)$ → Risk-averse individuals and actuarially fair insurance → Efficient outcome is full insurance

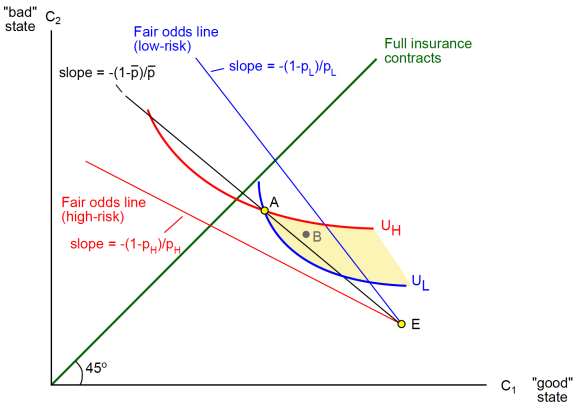


- Adverse selection (Rothschild and Stiglitz, 1976)
  - Individuals with heterogenous risk, information asymmetry
  - 2 types: Low risk (L) with probability P_L of bad state, and high risk (H) with probability $P_H > P_L$ of bad state
  - Fraction θ high risk
  - First best: Separating equilibrium with different contracts $(\alpha_{1i}, \alpha_{2i})$, such that $\alpha_{2i} = \frac{1-p_i}{p_i}\alpha_{1i}$
  - $\rightarrow$ Full insurance for both groups, higher premiums for H

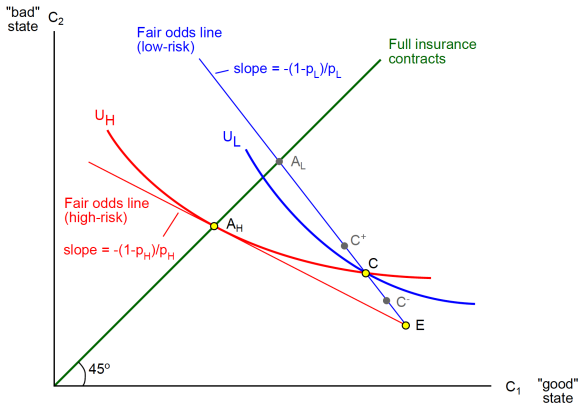


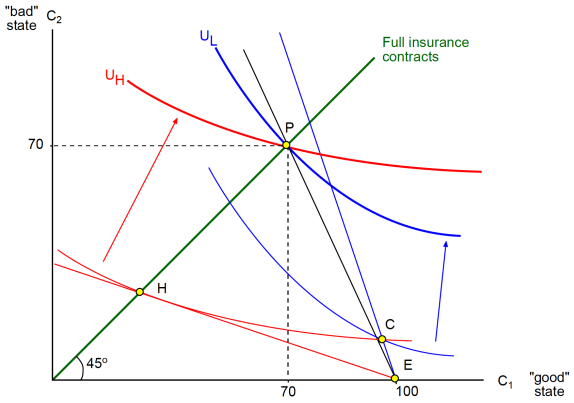
-
  - If not able to distinguish types, H types buy L type insurance $\rightarrow$ Negative insurer profits $\rightarrow$ Market collapse

- Second best 1: Pooling equilibrium cannot exist
  - Average risk is $\overline{p}$, with $p_L < \overline{p} < p_H$ $\rightarrow$ Opportunity for market entrant to offer less insurance at lower premium only to L types (cream skim) $\rightarrow$ Original contract only attracting H, so unprofitable

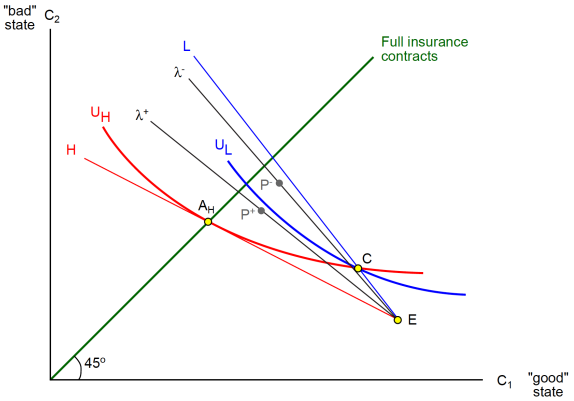


- Second best 2: Separating equilibrium can exist but is not efficient
  - For H, no cost to insurer of providing full insurance, but for L, full insurance creates incentives for H to join, resulting in negative profits $\rightarrow$ L will be underinsured
  - L harmed while H not gaining $\rightarrow$ Pareto inefficient

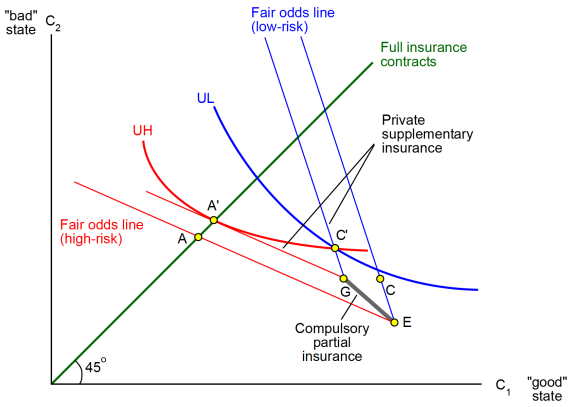




- Separating equilibrium may not exist
  - With high share of L, pooling policy close to actuarially fair line for L → Attracts both types, undermining separating equilibrium, but can itself be undermined → No equilibrium exists
  - With low share of L, pooling policy attracts only H types → Separating equilibrium



- Gains from mandated insurance
  - → L cross-subsidizing H
  - If no private insurance equilibrium (high proportion of L) → Mandate full insurance
  - If separating equilibrium exists, proportion of H not too large → Mandate partial insurance
    - Full insurance would only be preferred by H



- Empirical evidence on adverse selection
  - Test: Do risker types self-select into contracts with better coverage? → Indirect test: Are those who buy more insurance more likely to file claims?
    - But confounding adverse selection with moral hazard → Find cases with little moral hazard, e.g. death
  - Finkelstein and Poterba (2004): Evidence of asymmetric information

    - UK annuities
    - *Backloading*: Indexing and escalation of payments over time — L want more, H less → L buy more
    - *Payment to estate in event of death*: Guarantees and capital protection — H want more → H buy more

12.2 Moral hazard

- Occurs for both public and private insurance
- Dimensions:
  - Reduced precaution against entering adverse state — e.g. reducing preventive action
  - Increased probability of entering adverse state — e.g. claiming more injuries
  - Increased expenditures when in adverse state — e.g. overconsumption
  - Supplier responses to insurance against adverse state — e.g. fewer precautions
- Unemployment insurance → Unemployment duration
  - *Net replacement rate* of unemployment insurance: $r = \frac{\text{Net benefit}}{\text{Net wage}}$
  - Confounding effect in estimating impact of RR on labor supply: UI benefits capped, so RR falls with previous wages → Since high wage earners tend to have better employment prospects, spurious correlation → Use DDD from policy changes (time, space, income group)
  - Meyer (1990): Elasticity -0.9
    - *Hazard rate* h_t = No. of unemployed who find a job in week t divided by total no. unemployed for t weeks
    - *Cox proportional hazard*: $h_t = \alpha_t \times \exp(X\beta)$
    - Assumption: Effect of covariates proportional across weeks
  - Krueger and Meyer (2002): Consensus elasticity -0.5
- Disability insurance → Labor force participation
  - Parsons (1980) (Cross-sectional variation in replacement rates, OLS): Elasticity of 0.6; Increase in DI can completely explain observed decline in elderly US LFP 1950s–1970s
    - Problems: No variation in laws on RR, counterfactual LFP lower for low-wage earners
  - Bound (1989): At most 1/3 of male LFP decline explained by DI
    - Replication of Parsons with workers never applied to DI, with similar result → Decline in LFP driven by low-wage workers leaving labor force rather than high RR ratio

    - If rejected applicants don't work, neither would accepted ones → Using rejected as upper bound of LFP absent DI
    - Only 30% of rejected applicants return to work → Individuals on DI not able to work, would have left labor force even absent DI
  - Gruber (2000) (Differential DI law in Quebec and rest of Canada, 1973–1987, DiD): Implied elasticity 0.28–0.36
    - Only short-run response

12.3 Optimal social insurance

- Fundamental trade-off: Reducing market failure inefficiency (through increased coverage rate) increases moral hazard inefficiency
- Baily-Chetty Model: Second-best, including behavioral responses
  - h wage, c_H consumption in high state H; $b < w$ benefits, c_L consumption in low state L
  - Agent starts unemployed, exerts search effort e at cost $\Psi(e)$, with probability of employment $p(e) = e$
  - Lump-sum tax $t(b)$ in high state
  - Government budget constraint: $e \times t = (1-e)b \Rightarrow t(b) = \frac{1-e}{e}b$
  - $EU = e \times u(w-t) + (1-e)u(b) - \Psi(e)$
  - First best: $u'(c_H) = u'(c_L) \rightarrow$ Full insurance
  - Second best: Social marginal benefit of work w, but private $w - b$
    - Agent takes b and t as given: $MB = MR \rightarrow \Psi'(e) = u(c_H) - u(c_L) = u(w-t) - u(b)$
    - $\rightarrow$ Optimal effort level as function of b and t: $e^*e(b,t)$
    - Government budget constraint: $t \times e(b,t) = (1 - e(b,t))b \rightarrow t$ function of b $t = t(b)$
    - $V(b) = e(b) \times u[w - t(B)] + [1 - e(b)]u(b) - \Psi(e(b))$
    - FOC: $\left.\frac{dV(b)}{db}\right|_{b=b^*} = 0$
    - $\frac{dV(b)}{db} = -\frac{dt}{db}e \times u'(c_H) + (1-e)u'(c_L) + \underbrace{\frac{de}{db}[u(c_H) - u(c_L) - \Psi'(e)]}_{=0 \text{ (agent FOC)}}$
    - $\frac{dt}{db} = \frac{1-e}{e}\left(1 + \frac{\varepsilon_{1-e,b}}{e}\right)$
    - $\varepsilon_{1-e,b} = \frac{-b}{1-e}\frac{de}{db}$ is the elasticity of the probability of unemployment w.r.t. benefits
    - $\rightarrow$ Optimality condition: $\frac{u'(c_L)-u'(c_H)}{u'(c_H)} = \frac{\varepsilon_{1-e,b}}{e}$
      - LHS: Consumption smoothing; RHS: Moral hazard
    - Taylor expansion: $u'(c_L) - u'(c_H) \approx u''(c_H)(c_L - c_H)$
    - Risk aversion: $\gamma = \frac{-u''(c)c}{u'(c)}$
    - $\rightarrow \frac{u'(c_L)-u'(c_H)}{u'(c_H)} \approx \gamma\frac{\Delta c}{c}$

    - $\Rightarrow \gamma \frac{\Delta c}{c}(b^*) = \frac{\varepsilon_{1-e,b}}{e}$ (sufficient statistics that can be estimated)

- Gruber (1997) (panel data on food consumption): Consumption -23% without UI; +10p.p. RR → -2.8% in consumption drop
  - Optimal benefit rate varies a lot with risk aversion
  - $\gamma < 2 \rightarrow$ Replacement rates should be much lower than observed (typically 0.5–0.8)
  - Problem: γ poorly identified, consumption difficult to adjust because of fixed adjustment costs (e.g. housing)

- Chetty and Looney (2007) (US, with large UI system, v.s. Indonesia, with none): Consumption drops induced by adverse shocks of comparable magnitude in developing and developed countries
  - Maybe (i) easy to insure privately, so no need for public UI, or, more likely (ii) agents very risk-averse, so costly consumption smoothing and hence large potential benefits of public UI
    - Limited enforcement → Moral hazard reducing optimal benefit rate